2025. Proc Ling Soc Amer 10(1). 5930. https://doi.org/10.3765/plsa.v10i1.5930.

Orthographic effects on vowel contrast in language acquisition

Julian Vargo[1]

Abstract. This study measures whether differing letter shapes affect the production of /ɯ/ for L1-English learners of Turkish or Korean, which use a Latin-based and featural alphabetic orthography, respectively. Native speakers and L1 English learners of Turkish and Korean were recorded reading word lists and performing a picture-naming task in their target languages to compare whether read speech is significantly different from more naturalistic speech. Acoustic analysis of the high vowels /i, ɯ, u/ (n=1056) reveals that visual similarity between Turkish <ı> and <i> causes /ɯ/ to move towards the front of the vowel space, suggesting these letters' vertical line shape is interpreted as [+front]. The visual similarity of Korean /ɯ/~<으>, /u/~<우> and /o/~<오> results in learners analyzing /ɯ/ as having an intermediary vowel height between /u/ and /o/ rather than being contrastive via the [±round] feature. These findings demonstrate that speech production models must account for visual input, highlighting the need for further multimodal linguistic research in language acquisition.

Keywords. L2 acquisition; orthography; phonetic production; phonology

1. Introduction. In recent decades, reading as a medium of speech has been interpreted as a type of formal register (Labov 1972), but the exploration of iconicity between letter shapes and speech production has only minimally been researched. Furthermore, research on writing systems and reading processing has primarily been the focus of cognitive science research, but is often disregarded in phonology. Writing systems around the world have an incredible amount of diversity. In recent millennia, several alphabetical, featural, abjad, abugida, syllabary, and logographic writing systems have been developed and continue to adapt into other languages. Several orthographies, including English, are derived from the Latin script, which has gained widespread usage on a global scale and has been adapted into several languages outside of the Romance language family. Alphabetic scripts can contain modified extensions of the traditional Latin alphabet, such as the <ı ö ç ğ ş ü> letters seen in Turkish, which alter the alphabet by adding new graphemes or by modifying pre-existing graphemes through diacritical markings.

Alternatively, other alphabetic systems, like Korean Hangeul (한글), are completely unrelated to the English or Latin alphabet. Despite Hangeul and Latin sharing the status of being alphabets, they bear little resemblance to one another. Notably, Hangeul is a featural alphabet, meaning that its graphemes were intentionally made to represent several phonological features (Sampson 1985). Latin orthography, on the other hand, has somewhat arbitrary grapheme shapes and is mostly disregarded as a featural system, with only limited systematicity between grapheme shapes and sound correspondences (eg. English <p> and <b> appear visually similar and correspond to bilabial plosives) (Jee et al. 2022).

---

[1] I would like to thank Dr. Darya Kavitskaya, for her support and mentorship on this paper. Thank you to Dr. Justin Davidson for reviewing this paper and providing detailed feedback. Thank you to the Sociolinguistics Lab at Berkeley for providing incredible feedback for this project. Finally, I'd like to express my gratitude to all of the participants in this study.
Author: Julian Vargo, University of California, Berkeley (julianvargo@berkeley.edu).

Another key feature of Hangeul is that it uses a block-based system that clusters letters together into morphemic or syllabic blocks, and the shapes of these blocks depend on the features of the vowel in each syllable. Vowel nuclei that are [+back, -low] contain horizontally aligned blocks, while other vowel nuclei default to a vertically aligned orientation (see Figure 1). These orthographic differences between Korean and Latin-based orthographies may be of significance for speakers trying to acquire Korean, because of the orthographically-derived reinforcement of different phonological contrasts. Of particular interest to the present study, Korean phonemes like the high back unrounded /ɯ/ may be affected by the orthographic patterning with [-low, +back] vowels /o/ and /u/. The aforementioned Korean /ɯ/ contrasts with the Latin-derived orthographic system for Turkish, which also contains the high back unrounded /ɯ/ but expresses this phoneme with a non-featurally-based grapheme <ı>. Visually, there is no difference between this grapheme and the English <i> grapheme, besides the base part of the letter containing a dot. Ultimately, while the /ɯ/ phoneme visually aligns with the [+back, -low] vowels of Korean, the Turkish grapheme for /ɯ/ aligns with the [+high, +front, -round] vowels of both English and Turkish (orthographically represented as <i>).

These varied writing systems have generated previous research questions regarding how orthography plays a role on cognition and lexical or morphosyntactic retrieval (Pae & Wang 2022; Winskel 2022), but phonetic variables have still not been discussed in the context of letter shape. This difference in orthographic alignment provides the background for the present study's main research question: How do letter shapes and differing orthographic systems contribute to the acquisition of a new phoneme? More specifically, the present study will focus on the production and acquisition of /ɯ/ for Turkish-English or Korean-English bilinguals.

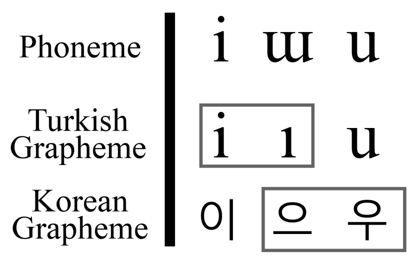

Figure 1. Visual alignment differences between Turkish and Korean grapheme-phoneme correspondences

1.1. ORTHOGRAPHY AND SPEECH PRODUCTION. A limited number of previous studies have examined how phoneme-grapheme correspondences more generally affect one's phonological system. Previous research demonstrates that orthography has significant effects on the production, perception, and cognitive processing of language learners (Hanna et al. 1966; Pulgram 1951; Bisani & Ney 2008; Warner 2004; Cerni et al. 2019). Cerni et. al. (2019) detail how Italian L1 speakers produce English consonants for longer durations if they are spelled with a doubled letter, because Italian writes doubled consonants to represent geminate phonemes. This shows a direct relationship between orthographic effects and a speaker's pronunciation, because the only reason for Italian speakers to transfer their L1 phonology would be through spelling, as English contains no phonemic geminates. Another way in which orthographic effects can affect phonology comes in the form of incomplete neutralization. Warner et. al. (2004) discuss how homophonous words actually contain subphonemic patterns which differentiate the two, and that this difference can be attributed to different spellings of the homophones. In particular, the duration of vowels in various Dutch homophones varies depending on whether a given word is written with word-final <t> or word-final <d>, despite Dutch's word-final devoicing rule. Because preceding vowel duration often serves as a cue to [±voice] (Lisker 1986), and vowels lengthen preceding orthographic <d> in Warner et al. (2004), this subphonemic difference supports the idea that orthography can interfere with other phonological processes.

In addition to pre-existing associations between graphemes and certain phonological features (eg. <d> being associated with [+voice]), previous work shows that graphemes can also be measured in terms of visual distance, to see if there is any kind of iconicity between grapheme shapes and sound patterns (Jee et al. 2022). Describing graphemes in terms of grapho-phonemic systematicity, as shown by Jee et al. (2022), we can posit that in Turkish, the vertical line shape of Turkish and English corresponds to the features [+high] and [-round] for the graphemes <i> and <ı>. In Korean, the features [+back] and [-low] correspond to a horizontal line for the graphemes <으>, <우>, and <오>. Because /ɯ/ demonstrates grapho-phonemic correspondences for different features in the two languages, speaker productions may be affected.

1.2. READING PROCESSING. Taking reading processing literature into account is especially important to understand how to isolate the variable of orthographic effect whilst collecting linguistic data. Reading processing can be understood as being processed at the word level or at the grapheme level (McCandliss et al. 2003). When readers learn a new writing system, reading aloud is done by "sounding out" each grapheme to construct a word (McCandliss et al. 2003). This is in direct contrast to word-level processing, where speakers are able to perform "rapid reading" through the recognition of entire words instead of individual graphemes (McCandliss et al. 2003). Because reading can reflect grapheme-level processing (GLP) or word-level processing (WLP), this creates the potential for the effect of grapho-phonemic interference to affect familiar words and familiar writing systems less. For novel words or for learners of new writing systems, GLP, and the impact of grapho-phonemic interference may occur more often.

In addition to understanding reading style in terms of WLP and GLP, read and not-read speech can be thought of as a function of reliance on previously stored exemplars. Exemplar-based phonological modeling posits that categories are stored as "cloud[s] of remembered tokens of that category" and can be attached to sociolinguistic factors (Pierrehumbert 2001). When speech production occurs, an exemplar-based model is probabilistic in nature, using these stored clouds of previously stored tokens to form a faithful and socially-appropriate reconstruction of what the speaker has heard in the past. In read speech, reliance on previously stored exemplars would not be predicted to be as important, as it provides speakers with some degree of explicit phonological information. An extreme case of this would be when speakers are presented with novel words, where they have zero stored utterances.

In the context of the present study, if speakers utilize GLP during a reading task, the reliance on previously stored exemplars is likely lessened as a speaker's attention is drawn towards each individual letter (see Figure 2). If WLP is utilized during a reading task, speakers are more likely to draw from previously heard exemplars of a certain word than they would through GLP. A picture-naming task, by contrast, could elicit speech with higher dependence on exemplar retrieval and minimal reading-processing. An example of WLP would be '*linguistics*', which, despite being 11 words, is read wholesale by most linguists without individual letter parsing. Compare this to the 11-letter 'echelonment', which requires individual letter parsing for readers who are not already acquainted with this word.

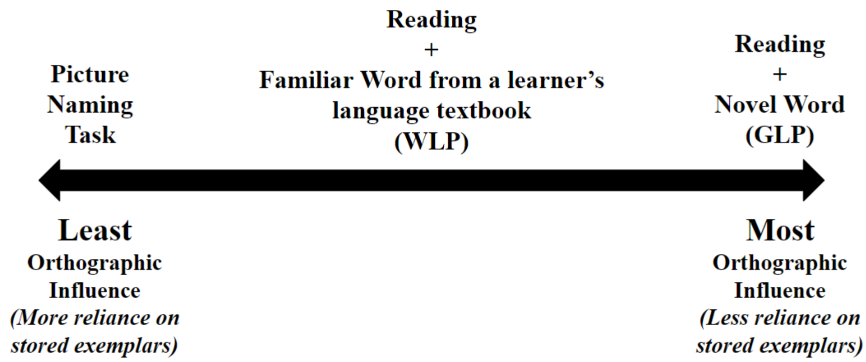

Figure 2. A model of different speech styles that correspond to different types of reading processing and varying reliance on stored exemplars.

1.3. TURKISH & KOREAN PHONOLOGY. Both Korean and Turkish contain the high back unrounded vowel, which is not a phoneme of Mainstream American English. In Californian English, where the study was conducted, this phone appears as an allophone of /u/, alongside the fronted and rounded realization [ʉ]. Because /ɯ/ is not a phoneme of English, this begs the question of how L1 English speakers have mapped this phoneme into their preexisting phonology — will Turkish or Korean /ɯ/ merely map to English /u/, or will speakers' varied pronunciations show patterning with orthographically related phonemes?

It is important to note that the effect of orthography alone is unlikely to cause an L1 Turkish speaker to have a more fronted pronunciation of <ı> despite it visually aligning with <i>. This is because, as an L1 Turkish speaker begins to learn how to read, an association between the <ı> grapheme and previously stored exemplars of /ɯ/ begins to develop at the same time as a speaker learns the <i> grapheme. This is to say, a Turkish learner has no motivation to front <ı> because there is as much association between the shape of "vertical bar" to [+front] as there is to [+back]. In contrast, L1 American English learners only correlate the vertical bar shape of <i> with frontness, subsequently biasing the acquisition of /ɯ/ in the direction of /i/. De Jonge et. al (2022) support the idea of Turkish learners mapping /ɯ/ to English L1 /i/, as L2 Turkish L1 English speakers produce the back high unrounded vowel with more front articulations. Ultimately, it is likely that English L1 phonology will have some effect on the acquisition of novel phonemes and provide a promising baseline for the idea that native speakers will have significantly different vowel spaces compared to their L2-acquisition counterparts.

A necessary consideration regarding Turkish vowel phonology is that Turkish uses vowel harmony. In Turkish Vowel Harmony, all of the vowels in a word must all be [+front] or must all be [-front]. There is also a second rounding harmony rule, where the high vowels /i, ɯ/ become [y, u] when following a rounded syllable. Note that this rule is iterative, and the domain is a phonological word (Arik 2015). If English L1 speakers are successful in implementing harmony rules for /ɯ/, then surrounding back vowels should reinforce /ɯ/'s status as a back phoneme, and any strong fronting of /ɯ/ would be an indication of orthographic cues working against harmony.

For English L1 learners of Korean, previously constructed associations between Korean graphemes and the high back unrounded phoneme aren't as robust as in Turkish because of its non-Roman alphabet, and learners are likely acquiring the /i/ <이>, /ɯ/ <으>, and /u/ <우> graphemes simultaneously. It is worth noting that there is a potential for the vertical line shape of <이> to be interpreted as being visually similar to English <i>. This is difficult to confirm, however, because <이> and <i> both correspond to high front unrounded vowels, making any visually-induced bias of <이> result in minimal changes to a speaker's formant structure.

A final consideration for Korean is its large number of English-derived borrowings, which are often adapted to fit Korean phonotactics. One such example is the Korean word for 'class', which is [kʰɯɭesʰɯ]. This borrowing modifies English [kʰlæs] by inserting [ɯ] to maximize the number of CV syllable sequences. For L1 English speakers who are learning Korean, it is possible that these epenthetic vowels are treated differently by learners, meaning that English borrowing status will be treated as an additional factor in the present study.

1.4. CUE-BASED PHONOLOGY IN ORTHOGRAPHY. If grapheme shape shapes phonetic production, then one way to model speaker phonology is through cue-based features. Cue-based features are a way of modeling how several sub-phonemic characteristics feed into how a certain phone's features are interpreted (Pfiffner 2021). Previous work has described cueing as being either phonetically derived or visually derived. For [ɯ], a low F1, low F2, and high F3 would serve as intermediary links to the [+high], [+back], and [-round] features, respectively. When looking at visual cueing, being able to see a speaker's lips being visually unrounded can fortify the perception of /ɯ/ as underlyingly [-round] (McGurk & MacDonald 1976). During the construction of a phonological category, a combination of "auditorily similar acoustic properties" and gestures is learned by listeners in groups (Kingston et al. 2009; Liberman et al. 1967; Parker et al. 1986). Kingston et al. (2009) refer to the mutually reinforcing combinations of these cues as an "intermediate perceptual property" (IPP). It is possible that the integration of these cues also interacts with orthography, through an intermediate property that feeds into a single feature.

2. Research Questions. Is Turkish <ı> /ɯ/ fronter for English L1 speakers as orthographic influence increases, due to alignment with <i> /i/? Is Korean <으> /ɯ/ backer and rounder for English L1 speakers as orthographic influence increases, aligning with <우> /u/ and <오> /o/?

3. Methodology. Five Korean speakers and six Turkish speakers of varying proficiencies participated in the present study (see Table 1). All of the speakers were native English speakers or fluent in American English from an early age. To measure participants' levels of bilingualism and to understand each participant's unique language-learning background, participants filled out an adapted version of the Bilingual Language Profile (Birdsong et al. 2012). The participants spanned a wide range of various linguistic backgrounds, varied kinds of language input (classroom, heritage speaker, speaking with friends, language apps, etc). The primary dependent variables being analyzed in the present study are the vowel formants F1, F2, and F3 for the high vowels /i, ɯ, u/. The participants were recorded using the Voice Memos application on an iPhone 14, iOS 14.4.1. using the iPhone's "lossless" m4a recording feature. Speaker F conducted the experiment through Zoom but recorded their audio in person, using a Shure KSM27 microphone.

The experiment consists of three parts, conducted in order: a reading task, a picture naming task, and a sociolinguistic questionnaire. For the sociolinguistic questionnaire, participants listed their comfort level reading, writing, speaking, and listening to their target language on scales from one to ten. Additional questions about speaker region and years speaking Turkish or Korean were also asked, but not analyzed in the present study.

The reading task consisted of a single-page word list, which participants were instructed to read aloud. For Turkish, this word list was 75 words long, consisting of 13 words containing /ɯ/ from the first three chapters of UC Berkeley's Elementary Turkish textbook (Özel 1995). 18 words from the first three chapters of Özel (1995) contained the /i/ phoneme, and ten words contained /u/. The target tokens were nouns with minimal inflection, to ensure that speakers were reading distinct roots, instead of repeating the same suffix several times over. 13 non-textbook words containing /ɯ/, ten non-textbook words containing /i/, and 11 non-textbook words containing /u/ were selected for the reading task as well.

Turkish	Korean
Speaker A: 40	Speaker 1: 38.5
Speaker B: 18	Speaker 2: 11
Speaker C: 19	Speaker 3: 30
Speaker D: 11.5	Speaker 4: 18
Speaker E: 17	Speaker 5: 18
Speaker F: 5	

Table 1. Proficiency scores of each participant, 40 being the highest proficiency, were calculated by summing reading, writing, speaking, and understanding ability

For the Korean word list, textbook words were also selected from the first three chapters of the textbook used by UC Berkeley's Elementary Korean course (Cho et al. 2000). 22 textbook words contained the phoneme /ɯ/. Eight of these 22 words also contained the other [+high] phonemes /i/ and /u/. 12 words from the first three chapters of the Korean textbook contain the /i/ phoneme, and one of these words (/mikuk/ 'United States') contains the /u/ phoneme as well. Four textbook words were selected for having the /u/ phoneme. One of these words, 부모님 (/bumonim/ 'parents'), also contained a high front vowel. For non-textbook Korean words, seven words containing /ɯ/ were selected, with four of these words also containing other high vowels. Six words were selected to contain /i/, with one of the words, 끼우다 (/k̤iuta/ 'insert') also containing the /u/ vowel. Four words were selected for containing /u/, and three of the four contained other high vowels. The one word from this category which did not contain other high vowels was 두뇌 (/tunwe/ ~ /tunø/ 'brain'). Although this word contains the graphemes <우> and <이>, this vowel token will not be analyzed because it is rarely realized as a high vowel for native speakers. Instead, the 위 can be thought of as a digraph representing either the diphthong /we/ or the monophthong /ø/, which alternates depending on the dialect of Korean being spoken (Lee 1999).

In the picture task, speakers were presented with ten images (for Turkish) or 11 images (for Korean) on a slide presentation. They were prompted to say aloud the word that each image represented. Given the L2 participant pool, who have widely differing vocabulary sizes, the picture task is constrained to L2 elementary-level words, and the list of images was constructed based on vocabulary from UC Berkeley's Turkish or Korean courses, which were both easy to represent as an image and were taught within the first three units of the course. However, using a picture task instead of another format, like a sociolinguistic interview, allows for the observation of non-conversational speakers in non-read speech.

After all the participant data was collected, each occurrence of a speaker producing the high vowels /i/, /u/, or /ɯ/ was marked by hand on a TextGrid in Praat by the author (Boersma & Weenink 2024). A handful of vowel tokens were realized as voiceless and lacked a clear formant structure on Praat, and were excluded from the analysis. In particular, most of the word-medial vowel tokens in 이스트 (/isɯtʰɯ/ 'east') were realized as voiceless [isɯ̥tʰɯ]. Vowel formant midpoints for F1, F2, and F3 were extracted using a Praat script (Elvira-García 2017). A speaker's language level was calculated by taking the sum of the speaker's self-identified comfort level for speaking, reading, writing, and listening to either Turkish or Korean.

After the raw data were collected, the remainder of the present study's analytical procedure follows the suggested order of operations in sociophonetic analysis, as proposed by Stanley (2022), where allophonic classification precedes outlier removal, followed by normalization. Allophonic classification was particularly challenging for this dataset, as allophonic distributions face much wider variation across such a diverse range of native, heritage, and L2 speakers. Speaker E in particular produced a number of [ɯ] tokens as a clear allophone of /i/, and produced [ɯ̟] (fronted [ɯ]) or [ɪ] as allophones of /ɯ/, creating a kind of [+high, +round] archiphoneme. The original workflow of the study was to simply remove any tokens that were more than two standard deviations from the mean formant values, but unexpected allophony complicated this, because the formant values for these realizations were valid, and were difficult to distinguish from true formant tracking errors. To keep unexpected allophones as part of the data, while removing tokens with formant values that fell more than two standard deviations from the mean, the author manually checked whether there were tracking errors or not. Ultimately, 17 data points in the study were removed out of 1070 tokens: four [i] tokens, five [ɯ] tokens, and eight [u] tokens, all of which were removed for having extraordinarily high F1 values, which were deemed as formant tracking errors. The true F1s of these tokens blended with F0 and could not be reliably distinguished with Praat's LPC algorithm. After removing errors, all of the speakers' vowels were normalized using the hyperbolic arcsine function, as described in Schroeder et al. (1979): B = 7 * arcsinh (formantHz / 650).

After the aforementioned data were entered into the spreadsheet, six separate fixed-effect linear regression models[2] were coded in R version 4.3.1, with an F1, F2, and F3 model separately conducted for both Turkish and Korean using the R packages car, emmeans, and lmtest (R Core Team 2024; Fox & Weisberg 2019; Lenth 2024; Zeleis & Horthorn 2002). Direct comparisons between Korean and Turkish will not be made, because it is probable that standard Korean /ɯ/ has different baseline formant values from standard Turkish /ɯ/, independent of any orthographic effects. Rather, for each language, each model will test whether the normalized formant value differs depending on the data being in a read versus non-read context, being a textbook word versus a non-textbook word, whether a word is an English borrowing or not (for Korean only), and each participant's composite language comfort score.

4. Results. Vowel shifting models show that reading causes /ɯ/-fronting in Turkish and /ɯ/-backing in Korean. Read contexts also cause some degree of /ɯ/-lowering and /ɯ/-rounding in Korean. Figure 3 highlights the key vowel shift trends under orthographically-influenced speech. These effects significantly diminish as language proficiency increases.

For F1 values in Turkish, the intercept, which represents /ɯ/ while reading textbook words, had a predicted 4.369 Bark ($p<.001$, 590 d.f., AIC=1262.9). Both the /i/ and /u/ vowels differed significantly from the intercept, with /i/ being a predicted .065 bark less than /ɯ/ ($p<.001$), and /u/ being a predicted .682 bark more than /ɯ/ ($p=.001$). Neither the picture task, the reading of non-textbook words, the language level, nor relevant interaction terms yield significant changes to /ɯ/ in Turkish. For F2 values in Turkish, the intercept was predicted by the model to be 12.770 bark ($p<.001$, d.f. 590, AIC=2147.3). /i/ is modeled to be 1.590 bark higher than the predicted /ɯ/ intercept ($p<.001$), and /u/ is modeled to be 2.944 bark less than the /ɯ/ intercept ($p<.001$). The picture task had a significant effect on speakers' pronunciations of /ɯ/, with the model predicting an F2 which is 1.644 bark lower than the non-read context ($p=.026$).

[2] Summaries of the linear regression models and supplementary materials are available on GitHub (https://github.com/julian-vargo/lsa_2025)

Language level also had a significant effect on Turkish /ɯ/ F2, with each increase in language level points corresponding to a predicted .059 bark decrease (p<.001). The reading of textbook words did not elicit a significant F2 difference, nor did the relevant interaction terms. For F3 values in Turkish, the intercept was predicted to be 15.515 bark (p<.001, d.f. 590, AIC=1191.2). /u/ differed significantly from /ɯ/ in F3, being a predicted .666 bark less than the intercept (p<.001), and /i/ did not differ significantly in F3 from /ɯ/. Language level was also a significant predictor of F3, with each increase in language level score resulting in a predicted increase in F3 by .018 bark. No other variables were significant predictors of Turkish F3.

Regarding more impressionistic results, another notable phenomenon observed was that speaker E produced some, but not all, tokens of /i/~<i> with a noticeably backed pronunciation. This effect was found when individually searching each vowel with outlier formant values for formant tracking errors. After consulting the spectrograms for these data and listening to the audio, these tokens were found to contain no tracking errors and were coded as /i/.

For F1 values in Korean, the intercept, which represents /ɯ/ while reading non-borrowed textbook words, had a predicted 4.561 bark (p<.001, 420 d.f., AIC=884.03). The F1 of /i/ differed significantly from the /ɯ/ intercept, being a predicted .909 bark lower than /ɯ/ (p<.001). /u/, on the other hand, did not differ significantly in F1 from /ɯ/. During the picture task, the predicted F1 was .940 bark lower than in read contexts (p=.018). The interaction between the picture task and language level was also significant, with every unit increase in language level resulting in an estimated .034 bark increase for the F1 of /ɯ/ (p=.018). No other factors or relevant interaction terms were found to be significant for the F1 model. For F2 values in Korean, the intercept had a predicted value of 10.260 bark (p<.001). Both /i/ and /u/ differed significantly from /ɯ/ for F2, with /i/ being a predicted 3.616 bark higher than /ɯ/ (p<.001) and /u/ being a predicted 1.361 bark lower than /ɯ/ (p=.006). Language level was the only other factor that resulted in significantly different F2 values of /ɯ/, where every unit increase in language level results in a predicted .059 bark increase in F2 (p<.001). For F3 values in Korean, the intercept had a predicted F3 value of 14.556 bark (p<.001). /i/ has an F3 which is significantly higher than /ɯ/ by .431 bark (p=.042), while /u/ did not differ significantly from /ɯ/. Language level was a significant predictor of F3 in Korean, with each unit increase in language level resulting in a .016 increase in bark (p=.005). English borrowed words were also a significant factor in predicting F3, with English borrowed words having a .234 bark increase in F3 from the intercept (p=.018).

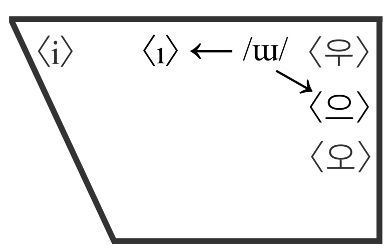

Figure 3. Vowel shifting of /ɯ/ in read contexts. Grapho-phonemes are listed as reference points. <ı> becomes closer in proximity to <i> and that <으> has a lower vowel height than <우>.

5. Discussion. Turkish /ɯ/ significantly differs from its other high vowel counterparts in general for F1 and F2. As expected, the F3 of Turkish /ɯ/ also does not differ significantly from /i/, while /u/ has a significantly lower F3 in Turkish. Much like how /ɯ/ is described in previous work on Turkish phonology, /ɯ/ fits between /i/ and /u/ on a two-dimensional Euclidean vowel plot, and has an F3 which aligns more with [-round].

Other social and linguistic factors complicate these baseline comparisons within Turkish. With regards to Turkish F1, a lack of significant effects due to sociolinguistic factors is not surprising, as it was predicted that /ɯ/ would shift primarily on the frontness-backness scale, not vowel height. There is little to no linguistic motivation to map /ɯ/ to a [-low] phoneme over /i/ or /u/ for speakers who do not speak Turkish as their L1, nor is there any orthographically derived motivation to produce a lower vowel. Regarding Turkish F2, on the other hand, sociolinguistic factors played a much larger role. When speakers produced /ɯ/ tokens while reading, the more fronted pronunciations support the idea that orthographic <ı>~/ɯ/ would give rise to an articulatory alignment with <i>~/i/. Fronting is also associated with a lower Turkish language level, and as speakers' self-identified proficiency increases, so does their backing of the /ɯ/ phoneme. F3 also patterns with F2 with respect to language level. As language level increases, speakers have less rounding of /ɯ/. It is worth noting, however, that this effect of F3 and language level is quite subtle, despite its statistical significance. Each increase in language level results in only a .018 bark increase in F3. To place this back into a more interpretable Hz scale, .018 bark roughly corresponds to 8 Hz at the 3000 Hz range[3].

Lastly, for speaker E, specifically, their /i/-backing during the Turkish reading task may also be in part due to orthographic effects. There is a possibility that this speaker, trying to produce more native-like speech, has hyper-extended for the backing of [+high, -round] vowels to /i/ because of the orthographic alignment between /i/~<i> and /ɯ/~<ı>. In a weaker version of this hypothesis, the speaker may simply be doing /i/-backing because they misread the grapheme as <i>. A much stronger claim would be that the graphemic alignment is actually leading to a new archiphoneme for this speaker of [+high, -round]. Further research is needed to discuss why this backing occurs more concretely.

Korean, much like Turkish, has an /ɯ/ phoneme which fits somewhere between /i/ and /u/. Nevertheless, Korean /ɯ/ is much closer to /u/ on the F2 axis than Turkish, meaning that Korean /ɯ/ is backer. Pure comparison between Turkish and Korean formant values is not sufficient to definitively say that orthography alone would result in more fronted Turkish pronunciations of /ɯ/ than Korean. The strongest support for orthographic effects cross-linguistically, on the other hand, can be seen in Korean's dramatically different patterns of sociolinguistic factors as predictors of /ɯ/ formant values when compared to Turkish. For F1, the picture task resulted in higher pronunciations of Korean /ɯ/. Upon first glance, this vowel raising in non-read contexts seems to run counter to the idea that the grapheme <으> would reinforce a [-low] feature. However, orthographic effects could still be potentially playing a role in the examined F1 differences for two reasons. One potential effect is that Korean /ɯ/ is becoming less phonetically distinct from /u/ in non-read contexts, as the visual differences between <으> and <우> encourage the /ɯ/ and /u/ phonemes to be phonetically distinct from one another. This is to say, two grapho-phonemes may be visually and featurally similar to one another, but their subtle visual differences serve as an impediment to a complete merger. The second possible explanation for spoken contexts contributing to vowel raising relies on the actual shape of the graphemes <우>~/u/, <으>~/ɯ/, <오>~/o/. The high back rounded vowel, /u/, corresponds to a high horizontal bar, and the mid back rounded vowel corresponds to a low horizontal bar. Perhaps then, when early learners produce /ɯ/, the interpretation of the plain horizontal bar shape is as an intermediary vowel height between /u/ and /o/. By this explanation, there is still a

[3] This approximate conversion between 8 Hz and .018 bark is 7 *(arcsinh(3008/650) - arcsinh(3000/650)) = 0.018 .

visual alignment effect between /u/ and /ɯ/, but the difference between the graphemes corresponds to differing vowel heights.

The F2 of /ɯ/ in Korean also changes in the opposite direction from Turkish as the language level increases. While Turkish F2 decreases by .059 bark for every unit of language-level increase, Korean F2 increases by .059 for every unit of language-level increase. This is to say, less-proficient Korean speakers start by having very backed F2 values, more similar to the pre-existing /u/ and /o/ phonemes in their L1 phonology, and begin to front slightly more as proficiency increases, aligning with the vowel space described in Lee (1999).

Lastly, for the Korean F3 model, it is quite notable that there is no baseline significant difference in F3 between /u/ and /ɯ/, but that /ɯ/ F3 increases as language level increases. This means that /ɯ/ is actually mapped more closely to the rounded vowels /u/ and /o/. Only as proficiency increases, the vowel becomes closer to the prototypical unrounded pronunciation.

The different patterns in vowel shifting between Turkish and Korean suggest that the different kinds of alphabetic systems result in distinct phonological interpretations. As predicted in the hypothesis, there is an alignment between <ı> and <i>, where <i> pulls <ı> away from the back of the vowel space. Given this phenomenon, we can posit that the vertical line shape shared by both of these letters may correspond to a kind of orthographic cue which contributes to a [+front] feature. Also in line with the hypothesis, Korean /ɯ/~<으> seems to align with the high vowel /u/~<우> and potentially with /o/~<오>, where <으> is analyzed by readers as having an intermediary vowel height between /u/ and /o/. Because the experiment was designed primarily to focus on F2 and F3 differences to compare /i/, /u/, and /ɯ/, more research is needed to see whether /ɯ/~<으> has a significantly lower F1 value than speakers' /o/ phoneme. This idea of alignment with the [+back, +round] series is supported by /ɯ/ becoming more rounded in read contexts and for speakers with lower language levels. Given this pattern in Korean, we can posit that a horizontal line is an orthographic cue which contributes to a speaker's production of a vowel as [+back, +round], and that the height of this horizontal line is interpreted as a correspondence to vowel height. Ultimately, the results support the existence of orthographic cues that have the potential to bias the interpretation of a non-L1 phoneme, much like how acoustic or visual cues can influence the perceived phonological features associated with a specific phone.

As learners begin to associate particular graphemes with specific phones, these correspondences may function in a similar "mutually reinforcing" acoustic cue-based process, as speakers rely on to make phonological inferences, as described by Kingston et al. (2008). The vowel variation in the present study expands on Kingston et al.'s (2008) work by positing that an "Intermediate Perceptual Property" connects not only acoustic cues to one another, but also to graphemic or other visual correspondences through learned mutual reinforcement. As speakers begin to connect a variety of multi-modal linguistic inputs, the characteristics of these inputs become mutually reinforcing and develop a phonological feature. Figure 4 demonstrates one possible model for combining multimodal cues through learned associations into a single phonological feature.

The strengthening of orthographic cues can be imagined as an increasing weighted coefficient, improving the likelihood that a word is spoken aloud with a feature associated with the particular grapheme shape. As orthographic cues lessen, speakers need to rely more on stored exemplars of acoustic and visual cues to reproduce a given phonological segment. This explanation aligns with the present data, as the picture-task seemed to elicit pronunciations from learners which mirrored the prototypical pronunciations of Turkish or Korean /ɯ/.

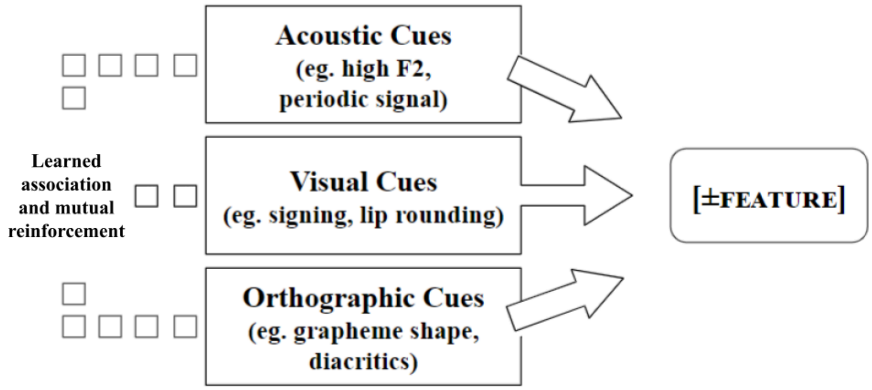

Figure 4. A generalized model for acoustic, visual, and orthographic cueing, all of which contributes to a featural output.

Lastly, the lack of statistical difference between textbook and non-textbook words demonstrates that word-level processing versus grapheme-level processing does not significantly affect how much orthographic cues shift speaker pronunciations.

6. Concluding Remarks. Orthography has the potential to affect the pronunciation of speakers and to change the features of language learners' phonologies. Isolating an orthographic variable can be tested for by examining phonetic phenomena (eg. formant frequency) and measuring its interaction between language proficiency and read versus non-read speech. Moreover, this phonetic consequence likely extends to these grapho-phonemes having altered phonological features. These phonological differences can be modeled by positing that features are compositions of multimodal linguistic cues, extending previous analyses by Lisker (1986), Kingston (2008), and Pfiffner (2021) to also include orthographic cues.

Nevertheless, more research needs to be conducted to more concretely determine the extent to which orthography affects one's phonological system. While the present data suggests that orthographic cueing plays a role in speaker phonology, more work needs to be done to determine how strong these effects are and how they vary with each language's unique phonology. Within Turkish, the grapho-phonemes /œ/~<ö> and /y/~<ü> are also potential candidates for the examination of orthographic effects, also being modifications of the preexisting Latin script. Extending beyond Turkish, there also exists an /ɯ/~<ı> grapho-phonemic correspondence in Crimean Tatar, a related Turkic language. One special characteristic of Crimean Tatar is that it utilizes multiple alphabets: The Latin-Turkic alphabet containing (/ɯ/~<ı>) and the Cyrillic Alphabet (/ɯ/~<ы>). For languages like Crimean Tatar, which utilize multiple alphabets for a single language, orthographic effects can be tested for by eliciting data with two identical word-lists, only differing in the type of script used. Additionally, because the effect of a textbook word versus a non-textbook word was an insignificant factor in the present study, this factor may not be strongly correlated with WLP or GLP. To more concretely tell whether words are produced with one of the two processing types, eye-tracking software could potentially solve this issue. With an eye-tracking study, an increased amount of time concentrating on individual graphemes (GL processing) may correspond with orthographic cues becoming more heavily weighted. Ultimately, future research is needed to better understand the orthography-phonetics interface, such as the continued examination of reading processing as a predictor of orthographic cue-weightedness. Each writing system, diacritical difference, and varied linguistic profiles of speakers have the potential to result in highly varied effects on pronunciation and must be expanded on in future research.

The demonstrated importance of orthographic cueing is relevant to applied fields such as language acquisition and language revitalization. Pedagogical frameworks and orthographic planning need to be adjusted with the aforementioned considerations in mind. In language acquisition, orthography can accentuate specific acoustic or visual cues, leading learners to develop altered pronunciations when reading aloud. In developing curriculum for language acquisition, it may be necessary to provide explicit instruction about pronunciation so students have a metalinguistic awareness about the articulatory targets associated with a certain phone. In language revitalization, these findings are relevant to the creation or modification of an orthographic system. These findings highlight a challenge for linguistic communities that make an effort to be as linguistically conservative as possible when reclaiming their language, where the use of orthographies derived from Latin may inadvertently accelerate sound change. Moreover, the development of featural orthographies, like the development of Hangeul for Korean, must be carefully considered so that grapheme shapes intuitively represent phonological features. The visual difference between <우> and <으> seems to be one such example where Hangeul's intuitiveness falters, as the round/unround pair is reinterpreted as a height difference by speakers of lower Korean proficiency. This means that both Latin-derived (or derivations of other pre-existing writing systems) and newly developed orthographies all have the potential to be affected by orthographic cueing.

References

Arık, Eylem. 2015. An experimental study of Turkish vowel harmony. Poznań Studies in Contemporary Linguistics 51(3). 359–374. https://doi.org/10.1515/psicl-2015-0014

Birdsong, David, Libby M. Gertken, & Mark Amengual. 2012. *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism*. Austin, TX: COERLL, University of Texas at Austin. https://sites.la.utexas.edu/bilingual/

Bisani, Maximilian & Hermann Ney. 2008. Joint-sequence models for grapheme-to-phoneme conversion. *Speech Communication* 50(5). 434–451. https://doi.org/10.1016/j.specom.2008.01.002

Boersma, Paul & David Weenink. 2024. *Praat: doing phonetics by computer* (version 6.4.11). [Computer program]. http://www.praat.org/

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

Cho, Young-mee O., Hyo Sang Lee, Carol Schulz, Ho-min Sohn & Sung-Ock Sohn. 2000. *Integrated Korean: Beginning 1*. Honolulu: University of Hawai‘i Press.

Elvira-García, Wendy. 2017. vowels. [Praat script]. https://github.com/wendyelviragarcia/vowels/tree/master

Fox, John & Sanford Weisberg. 2019. car. [R package]. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Hanna, Paul R., Jean S. Rudolph & Margaret S. Blouin. 1966. Phoneme-grapheme correspondences as cues to spelling improvement. *Office of Education Cooperative Research Program Report* 97–130. Washington, DC: U.S. Department of Health, Education and Welfare.

Jee, Hana, Monica Tamariz & Richard Shillcock. 2022. Systematicity in language and the fast and slow creation of writing systems: Understanding two types of non-arbitrary relations between orthographic characters and their canonical pronunciation. *Cognition* 226. 105197. https://doi.org/10.1016/j.cognition.2022.105197

de Jonge, Keryn, Olga Maxwell & Helen Zhao. 2022. Learning on the field: L2 Turkish vowel production by L1 American English-speaking NGOs in Turkey. *Languages* 7(4). 252. https://doi.org/10.3390/languages7040252

Kingston, John, Randy L. Diehl, Cecilia J. Kirk & Wendy A. Castleman. 2008. On the internal perceptual structure of distinctive features: The [voice] contrast. *Journal of Phonetics* 36. 28–54. https://doi.org/10.1016/j.wocn.2007.06.002

Lee, Hyun Bok. 1999. Korean. In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*, 120–123. Cambridge: Cambridge University Press.

Lenth, Russel V. 2024. emmeans [R package]. https://github.com/rvlenth/emmeans

Lisker, Leigh. 1986. Voicing in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29(1). 3–11. https://doi.org/10.1177/002383098602900102

McCandliss, Bruce D., Laurent Cohen & Stanislas Dehaene. 2003. The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences* 7(7). 293–299. https://doi.org/10.1016/S1364-6613(03)00134-7

McGurk, Harry & John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264. 746–748. https://doi.org/10.1038/264746a0

Özel, Suzan. 1995. Turkish I: A communicative approach.

Pae, Hye-Kyeong & Min Wang. 2022. The effects of writing systems and scripts on cognition and beyond: An introduction. *Reading and Writing* 35(1). 1315–1321. https://doi.org/10.1007/s11145-022-10289-z

Pfiffner, Alexandra M. 2021. Cue-based features: Modeling change and variation in the voicing contrasts of Minnesotan English, Afrikaans, and Dutch. Washington, DC: Georgetown University dissertation.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee & Paul Hopper (eds.), *Frequency effects and the emergence of linguistic structure*, 137–158. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.45.08pie

Pulgram, Ernst. 1951. Phoneme and grapheme: A parallel. *Word* 7 (1). 15–20. https://doi.org/10.1080/00437956.1951.11659389

R Core Team. 2024. R: A language and environment for statistical computing (version 4.3.1). Vienna: R Foundation for Statistical Computing. https://www.r-project.org/

Sampson, Geoffrey. 1985. *Writing systems: A linguistic introduction*. Stanford, CA: Stanford University Press.

Schroeder, Manfred. R., Atal, Bishnu S. & J. L. Hall. 1979. Objective measure of certain speech signal degradations based on masking properties of human auditory perception. In Björn E. Lindblom & Sven Öhman (eds.). *Frontiers of Speech Communication Research*. Academic Press.

Warner, Natasha, Allard Jongman, Joan Sereno & R. Rachèl Kemps. 2004. Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics* 32(2). 251–276. https://doi.org/10.1016/S0095-4470(03)00032-9

Winskel, Heather. 2022. Script relativity hypothesis: Evidence from reading with different spatial layouts and varied lexical tone. *Reading and Writing* 35(6). 1323–1341. https://doi.org/10.1007/s11145-021-10225-7

Zeileis, Achim & Torsten Hothorn. 2002. Diagnostic checking in regression relationships. *R News* 2(3). 7–10. https://cran.r-project.org/doc/Rnews/