The role of orthographic input in the distributional and lexical learning of non-native speech sounds

by

Abdulaziz Alarifi

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics
University of Alberta

Examining committee:

Benjamin V. Tucker, Supervisor
Rachel Hayes-Harb, Supervisory Committee
Juhani Järvikivi, Supervisory Committee
Elena Nicoladis, Examiner
Christine E. Shea, External examiner

Abstract

This dissertation explores the interaction between orthographic information and auditory forms in the distributional and lexical learning of consonant length contrast by monolingual English speakers. Two potentially important orthographic variables were examined: Orthographic compatibility (whether the orthographic information supports or contradicts the distributional or lexical information) and orthographic familiarity (whether the native and target languages share the same orthography).

In the first experiment, 10 groups of learners were trained on either a unimodal or bimodal distribution of two length continua. Out of the 10 groups, 8 groups were additionally exposed to orthographic cues that varied in their compatibility with the distributional information (compatible vs. incompatible) and familiarity with the orthography of learners' native language (Roman vs. Arabic). Following training, all participants performed an AX discrimination task to examine their perception of the length contrast. The results revealed that, in general, the availability of either familiar or unfamiliar orthographic input that signaled the existence of a single length category significantly lowered learners' discrimination of the length contrast regardless of the auditory distribution. Further, the exposure to orthographic input that supported two-category length distinction enhanced the discrimination of length contrast irrespective of the distribution. However, the most significant improvement occurred when both distributional information and familiar orthographic input were compatible.

In the second experiment, the same orthographic variables were tested in the lexical learning of 12 pseudo-words containing either a singleton or geminate along with their pictured-meanings. The results revealed that presenting learners with compatible Roman orthography (where singleton and geminate consonants had distinct spellings) significantly improved their lexical encoding and subsequent retrieval of words containing the length contrast. For the compatible Arabic orthography group, the improvement was only significant for those who had multiple training cycles, indicating that unfamiliar orthography may require more exposure before it can be learned. Finally, no significant differences were found for learners who were presented with incompatible Roman or Arabic orthography (where both singletons and geminates had the same spellings) compared to those who received no orthography.

Together, these findings indicate that orthographic input, regardless of its level of compatibility or familiarity, influences the acquisition of non-native speech sounds. However, systematic individual variations were present, suggesting that learners exhibit differential preference towards learning from written cues. Overall, the results provide original contributions to the body of work on the interaction between orthography and L2 phonology and offer theoretical implications for the models of L2 phonology.

Preface

This dissertation is an original work by Abdulaziz Alarifi. The research project, of which this dissertation is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name "The Role of Orthography in Second Language Phonology", No. Pro00058839, November 9, 2015.

This dissertation is dedicated to my wife Eman and my son Sultan

Acknowledgements

Throughout my time at the University of Alberta, I have had the great fortune to work with many wonderful people. First and foremost, I would like to express my deepest appreciation to my supervisor, Dr. Benjamin V. Tucker, whose unwavering support has profoundly helped this dissertation come to light. Ben has been a constant source of inspiration, a great mentor and always a pleasure to work with. I also wish to express my sincere gratitude to Dr. Rachel Hayes-Harb, who graciously agreed to serve on my supervisory committee. Rachel's extensive knowledge and shrewd advice were essential in helping me tackle the complexity of the topic. My gratitude is extended to Dr. Juhani Järvikivi for his incisive feedback that both deepened and refined my analyses. It would be very remiss of me not to acknowledge the invaluable contribution of my previous supervisor Dr. Anne-Michelle Tessier whose insights and stimulating discussions helped stoke my interest in the interaction between orthography and second language phonology.

This research would not have been possible without the multitude of individuals who have agreed to participate in my experiments, to them I say thank you. I also owe a great deal to the lab assistants at the Alberta Phonetics Lab, in particular Kira Shelton and Melina Sinclair, for spending countless of hours assisting me in running the experiments. I also wish to recognize my fellow lab members, past and present, for their fantastic support throughout the years. I especially would like to thank Filip Nenadić and Mathew Kelly for always offering to lend a helping hand (or brain), and Annika Nijveld and Ryan Podlubny for giving me a much-needed boost just days before my doctoral defense.

My acknowledgements would not be complete without thanking my family. I am deeply grateful to my mother, Latifah, for her countless prayers and encouraging words that kept me going over the years; my late father Saad, who always believed in my ability to succeed; and my late brother Waleed, whose passion for learning languages ignited my interest in linguistics.

Finally, I would like to extend my sincere thanks to King Saud University for funding my graduate studies, and the Saudi Cultural Bureau in Canada for facilitating my stay there.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Learning non-native speech sounds is a challenging task for adult learners especially if the target language has a phonemic contrast that is not available in their native language. Researchers have extensively examined the phonological and phonetic factors that make a particular sound difficult to acquire for adult learners. A well-established finding in second language phonology research states that the degree of the perceived phonetic similarities between the target sound and its corresponding native category can either hinder or facilitate the perception and production of non-native segments. Based on this notion, several models and frameworks have been posited to explain precisely the roles that a learner's native language play in her or his acquisition of L2 phonology (Hancin-Bhatt, 1994; Flege, 1995; Best, 1995; Brown, 2000; Best & Tyler, 2007). An equally important question, which has rarely been given the same attention until recently, is how the orthographic representations of auditory input affect L2 learners' ability to acquire non-native phonemes. This dissertation aims to contribute to the growing body of work on the complex relationship between orthography and novel speech sounds in second language phonology.

Perhaps one of the most fundamental differences between first and second language acquisition is the nature of the linguistic input that is available for infants and adults learners. Research on infants' phonological development has shown that sound categories are mostly acquired during the first year of life (Werker & Tees, 1984a). It has been suggested that this perceptual development is achieved through distributional learning (Maye, Werker, & Gerken, 2002; Maye, Weiss, & Aslin, 2008; Yoshida, Pons, Maye, & Werker, 2010; Cristià, McGuire, Seidl, & Francis, 2011). Although adults possess similar abilities (Maye & Gerken, 2000, 2001; Hayes-Harb, 2007; Perfors & Dunbar, 2010; Ong, Burnham, Escudero, & Stevens, 2017), albeit less effective (Wanrooij, Boersma, & Van Zuijen, 2014), their exposure is not limited to the relative frequency of the auditory stimuli available in speech. Unlike infants, the linguistic experience of literate adult learners is shaped by many factors including the target language input, which typically encompasses various data types that can be utilized by learners to acquire the linguistic system of the target language.

In addition to the frequency distributions of sound categories, learners have access to the written representations of these categories. Previous L2 research has largely overlooked the ways in which orthography influences the acquisition of nonnative speech sounds. Until recently, the impact of written input on the perception and production of L2 sound categories has been implied (e.g., Flege, 1988; Best and Tyler, 2007), but not experimentally teased apart. Still, recent contributions have provided valuable insights into the relationship between orthographic representations and nonnative speech sounds in perception (Escudero & Wanrooij, 2010; Simon, Chambless, & Kickhöfel Alves, 2010; Pytlyk, 2011; Mok, Lee, Li, & Xu, 2018), production (Rafat, 2015; Bassetti & Atkinson, 2015; Young-Scholten & Langer, 2015; Hayes-Harb, Brown, & Smith, 2018; Han & Kim, 2017; Bassetti, 2017; Bassetti, Sokolović-Perović, Mairano, & Cerni, 2018), word learning (Hayes-Harb, Nicol, & Barker, 2010; Simon et al., 2010; Showalter & Hayes-Harb, 2013; Escudero, Simon, & Mulak, 2014; Showalter & Hayes-Harb, 2015; Hayes-Harb & Hacking, 2015; Mathieu, 2016; Hayes-Harb & Cheng, 2016; Simonchyk & Darcy, 2018; Showalter, 2018) and L2 spoken word recognition (Veivo & Järvikivi, 2013; Veivo, Järvikivi, Porretta, & Hyönä, 2016; Qu, Cui, & Damian, 2018; Veivo, Porretta, Hyönä, & Järvikivi, 2018).

No previous study though has offered a systematic assessment of the impact of orthographic input in the perceptual and lexical learning of nonnative speech sounds. In particular, the way orthographic information interacts with the acquisition of non-native speech sounds at the perceptual as well as recognition levels has not been fully understood.

This dissertation examines the effects of orthographic cues on the acquisition of the length contrast by native English speakers. Although consonant length is not contrastive in English as opposed to Arabic, native speakers manipulate the duration of consonants phonetically. Perceptually, native English speakers exhibit sensitivity to durational cues, identifying, for example, /tɔpɪk/ with a /p/ closure duration shorter than 150 ms as *topic*, and as *top pick* if the duration is longer than 250 ms (Pickett & Decker, 1960). In production, native English speakers tend to produce a phoneme spelled with more than one letter longer than when same phoneme spelled with a single letter (Brewer, 2008). Thus, the availability of this feature in English at the phonetic level is expected to facilitate learning the contrast phonemically. Previous studies have shown that English learners can improve their discrimination of the singleton/geminate contrast with language experience (Hayes, 2001; Kato & Tajima, 2002; Hayes-Harb, 2005; Hayes-Harb & Masuda, 2008); or, more relevant to the current dissertation, after brief training (Hirata, 2004; Tajima, Kato, Rothwell, Akahane-Yamada, & Munhall, 2008; Hirata, Whitehurst, & Cullings, 2007; Motohashi-Saigo & Hardison, 2009; Pajak & Levy, 2011; Porretta & Tucker, 2015). In the training studies, the learning consisted of auditory exposure only or auditory exposure coupled with explicit instruction, visual presentation of the waveforms of the contrast or instant feedback. Based on these findings, this particular contrast offers a unique opportunity to evaluate the effects of orthographic cues following a brief exposure.

The remainder of this chapter is organized as follows. Section (1.2) briefly reviews some of the models and theories of L2 phonology and provides an overview of the research on distributional learning as a learning mechanism utilized by adult learners in the acquisition of non-native speech sounds ; section (1.3) offers a detailed review of studies that have explored the relationship between orthography and L2 phonology; and finally, section (1.4) concludes this chapter with the objectives and research questions of this dissertation.

1.2 L2 phonological acquisition

1.2.1 Models of L2 phonology

One of the earliest models that attempted to explore the interaction between L1 and L2 phonology is Contrastive Analysis (Lado, 1957), which claims that L2 phonemes which are similar to the learner's native language are easier to acquire than dissimilar phonemes. However, several studies have also pointed out that L2 learners have no difficulty in acquiring segments that are unavailable in their native language.

Following this simple phonemic comparison between L1 and L2 phonological systems, researchers have sought to develop more complex models that can tackle the various types of learning difficulties often encountered by adult L2 learners. Among these is the Speech Learning Model (SLM) (Flege, 1995). One of the assumptions this model makes is that L2 learners exploit the same learning mechanisms and processes, such as forming phonetic categories in the long-term memory, that are available during the acquisition of their native language. Crucially, the SLM hypothesizes that sounds in L1 and L2 are perceptually related to each other at the surface level rather than at the abstract level. Data on the production and perception of English by Japanese learners has shown that the allophonic variants of a phoneme are less prone to errors than other variants in different positions (Sheldon & Strange, 1982). For instance, adult Japanese learners of English performed better in discriminating /ɹ/ and /l/ in word-final positions than in word-initial positions, possibly due to durational cues, phonotactic constraints as well as coarticulatory effects. For example, the phonemes /ɹ/ and /l/ are consistently longer word-finally, thus providing listeners with more acoustic information to perceive the contrast (Lively, Logan, & Pisoni, 1993). Regarding the relative ease or difficulty in learning a non-native category, the SLM classifies L2 sounds in terms of their similarities and dissimilarities to L1 segments. Identical and different sounds are more easily perceived and produced, while similar sounds are predicted to be more difficult to learn.

Another model that stresses the importance of surface representation in L2 was posited by Best (1995). According to this model, naive listeners (those without prior experience in the target language) perceive non-native sounds in terms of their similarities and dissimilarities to their native language by means of articulatory gestures. That is, listeners assimilate a non-native phonetic segment into a native phoneme based on the articulatory similarities between the two. The model did not originally intend to explore the L2 learners' perception of speech sounds; instead, the focus was placed on the interaction between listeners' L1 phonology and non-native sounds of unfamiliar languages. Noting that the Perceptual Assimilation Model (PAM) is often reported in the L2 phonology literature, Best and Tyler (2007) extended the model to account for the perceptual difficulties for listeners who are actively learning a second language. According to PAM-L2, L2 learners base their perception of non-native phonetic segments on the degree of similarities between the articulatory properties of these L2 segments and the L1 phonological categories. Unlike PAM, however, PAM-L2 allows for perceptual assimilation to occur at the phonological level and thus differs from SLM, which considers phonetic properties as the only level of investigation. PAM-L2 hypothesizes that the degree of difficulty L2 learners have in discriminating a given non-native contrast is determined by the type of perceptual assimilation process: L2 learners are able to perceive the contrast if it is assimilated into two distinct L1 categories. However, difficulty will arise in perceiving the L2 contrast if it is assimilated into the same L1 category.

1.2.2 Distributional learning

One way adult learners acquire a non-native contrast is through minimal pair comparison. Learners recognize that a sound is important in the target language once they realize that replacing it with another one in the same position alters the meaning. Their attention then is focused on learning the phonetic properties that differentiate these sounds. Once the contrasting properties are discerned, new categories are formed. However, as the SLM model postulates, L2 adult learners retain the same learning mechanisms used when learning the sound system of their L1 (Flege, 1995). Experimental research has shown that adults—like infants—can form new phonetic categories by extracting the statistical regularities of the L2 sounds structures that are available in the input. In training studies, adult learners exhibited sensitivity to the statistical information in the auditory input and learned sound categories based on their distribution. For instance, Maye and Gerken (2000) exposed adult English speakers to a sound continuum ranging from the voiced /d/ (as in **d**ay) to an unaspirated /t/ (as in s**t**ay) in 8 equal steps. While English listeners can perceive the contrast phonetically, they are not able to discriminate /d/ and an unaspirated /t/ on a phonemic level, as both sounds belong to the same phonemic category /d/. The distribution was either unimodal, where learners heard tokens on the center of the continuum more often, or bimodal, where learners heard tokens near both ends of the continuum more frequently (See Figure 1.1).

Following training, the learners performed an AX discrimination task where pairs of stimuli extracted from endpoints of the continuum were presented and they had to decide whether they were the same or different. If the bimodal group relied on statistical distributions to establish phonemic categories, they would judge tokens one and eight as different, while the unimodal group would only establish a single category and thus perceive both endpoints as similar. The results indicated that participants who were exposed to a bimodal distribution indeed performed significantly better in discriminating tokens along the /d/-/t/ continuum than learners

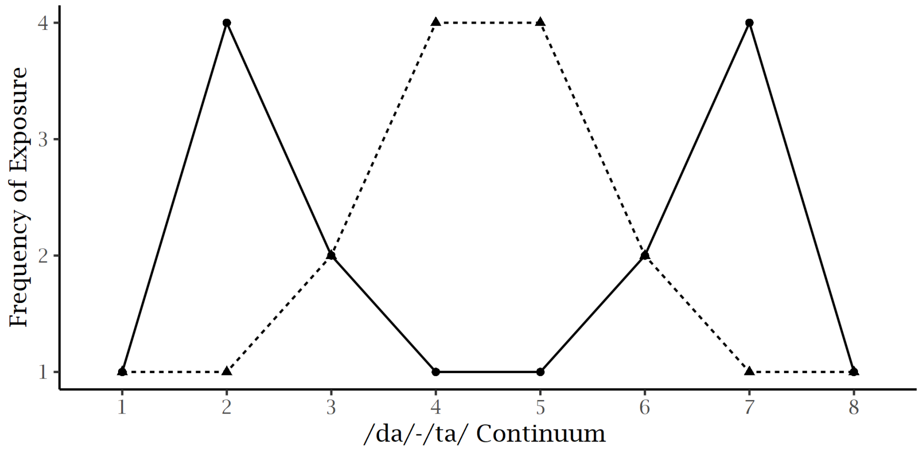


FIGURE 1.1: An illustration of the distributional learning participants were exposed to (Solid line = bimodal distribution. Dotted line = unimodal distribution).

in the unimodal group. The researchers concluded that phoneme acquisition could be achieved through distributional learning.

This learning mechanism was further tested on the acquisition of phonologically conditioned allophones. Peperkamp, Pettinato, and Dupoux (2003) exposed adult native French listeners to a continuum stretching from [ʁ] to [χ], where the end points were the allophones of the voiced uvular fricative /ʁ/. In French, the voiceless uvular fricative is the result of voice assimilation with adjacent voiceless consonants. In order to determine whether contextual factors affect distribution-based learning, the researchers presented a group of listeners with a bimodal distribution of [ʁ]-[χ] continuum followed by CV syllable chosen at random, while another group was exposed to the same type of distribution with the addition of the contextual factors that condition the allophones, i.e. a voiced obstruent follows tokens containing [ʁ] while a voiceless one follows tokens containing [χ]. The follow-up AX discrimination task revealed that listeners who were exposed to the target allophones in their rightful contexts made more errors in discriminating the pair than the other group, which were not provided with same contexts. Peperkamp et al. (2003) concluded that distribution-based learning is sensitive to the contextual cues that condition phonetic segments. As the results have shown, allophones are discriminated better when presented out of context; however, the ability to perceive the allophonic contrast diminishes when they are embedded in their phonological contexts. The researchers attributed this discrepancy in discriminating seemingly distinct segments to the listeners' different levels of processing when perceiving phonetic segments. Listeners rely on their auditory and acoustic knowledge to process allophones in isolation, while their phonological representation comes into play if context is introduced.

Relevant to the current study is the question of whether adult learners of a second language also utilize distribution-based learning. Shea and Curtin (2005) addressed this issue by training native Spanish and English speakers to perceive the Arabic pharyngealized consonant+low vowel sequence. In Arabic, the pharyngealized consonants /tˤ sˤ dˤ ðˤ/ alter the quality of adjacent vowels by lowering their F2 values. The resulting allophone of the Arabic /æ/ in pharyngealized context has phonemic status in English, but it does not have a direct equivalent in Spanish. Thus, an English listener may map this allophone into his or her native phoneme /ɑ/. In contrast, a Spanish listener will probably assimilate both Arabic [æ] and [ɑ] into the Spanish /ɑ/. Shea and Curtin explored the effect of L1 and contextual cues that condition allophonic alternation on distribution-based learning. The researchers hypothesized that native English speakers may be sensitive to the vowel quality in the input to acquire pharyngealized/non-pharyngealized vowel distinction because they are both phonemic in English. As for Spanish listeners who lack this contrast, the types of consonants that condition vowel distinction may trigger the acquisition.

To test these hypotheses, Shea and Curtin (2005) created two types of continua from natural speech; in one, the consonant varied along a Voice Onset Time (VOT) continuum and the vowel remained constant; in the other, the vowel varied along an F2 continuum with the consonant held steady. Both groups of listeners were exposed to unimodal and bimodal distributions of the two continua. An ABX discrimination task on the pharyngealized/plain consonant+ low vowel sequence was given before and after training in order to examine the effects of statistical distribution-based learning on the acquisition of the non-native contrast. The participants heard three stimuli in a row and had to decide which of the first two (A or B) was similar to the third one (X). The results indicated an overall significant increase in the accuracy of discriminating the Arabic contrast in both English and Spanish listeners. Spanish listeners who were exposed to a bimodal distribution of the low vowel continuum achieved the greatest improvement. With the caveat that these results were achieved under this particular experimental setup, the researchers argued that adult learners of a second language could utilize distribution-based learning to create non-native categories.

Similarly, Pajak and Levy (2011), Pajak (2012) trained 48 monolingual English listeners on the novel singleton/geminate contrast using sonorants and obstruents that varied along the length continuum. Participants were exposed to either continua containing sonorants or obstruents. Each continuum consisted of 8 tokens separated equally by 15 ms with the end points measured at 100 ms for short consonants and at 205 ms for long consonants. The participants were divided into unimodal and bimodal training groups and subsequently were given an AX discrimination task. The aim was to see whether participants would be able to infer categories based on the type of distribution they were exposed to and whether they could transfer their newly acquired knowledge to other class of consonants. The testing materials contained pairs with a combination of tokens 1 and 8: (1-8, 8-1) for different pairs or (1-1, 8-8) for same pairs. The results showed that participants exposed to sonorants were able to infer two categories from the bimodal distribution and one category from on the unimodal distribution. Additionally, they were able to utilize their training on sonorants to perceive the length contrast in obstruents. Conversely, participants in the bimodal group who were trained on obstruents failed to learn the contrast and eventually replicated their inability to perceive the length contrast when tested on sonorants. Pajak and Levy (2011) attributed this discrepancy to the fact that obstruents, especially fricatives, are generally longer than sonorants in English and thus participants in the bimodal group might have perceived the short and long tokens as members of the same category.

Distributional learning has also been shown to enhance the lexical encoding of novel speech sounds. Perfors and Dunbar (2010) trained 61 participants on two tasks: Phonetic and word learning. One group was exposed to a bimodal distribution of a novel voicing continuum ranging from /g/ to unaspirated /k/; the other group was presented with a bimodal distribution of a familiar contrast ranging from /d/ to the aspirated /$\mathrm{t^h}$/. Following the phonetic training, both groups were asked to learn multiple pseudo-words containing the same novel contrast with their pictured-meanings. Following the learning phase, participant were tested on their ability to associate each pictured-meaning with its corresponding auditory form. During the auditory-picture matching test, participants saw an image and heard two words differing in whether and word-initial sound was /g/ or the unaspirated /k/, and were asked to indicate which auditory form corresponded with the image. The results showed that those who received distributional training of the same contrast established more accurate lexical representations than those who were trained on the /d/-/$\mathrm{t^h}$/ contrast. Hayes-Harb (2007), however, found no effects for distributional information in improving the lexical encoding of nonnative speech sounds.

Although the relationship between perception and production is often complex, there is evidence that perceptual learning based on statistical distribution can transfer to production. Baese-Berk (2010) examined the relationship between perception and production using the statistical learning paradigm. In the first experiment, two groups of participants were exposed to either a bimodal or unimodal distribution of multiple voicing continua ranging from the voiced /d/ to the unaspirated /t/ in 8 equal steps preceding three different vowel environments. The subjects were tested on their production of the learned contrast using a repetition task. The participants were presented each time with any of tokens 1, 3, 6 and 8 from the three continua and asked to reproduce the token as accurately as possible. The results showed that the participants in the bimodal group improved slightly by producing the target contrast with a larger distinction between the end points of the continuum after the second day of training. Furthermore, there was a correlation between accuracy in perception and improved production. Baese-Berk concluded that implicit statistical learning based on distributional information could filter through production.

The findings of these studies suggest that adult learners, aided only by limited exposure to the statistical information available in the input, could acquire non-native contrasts with notable success. Clearly, L2 learners possess the ability to overcome quickly their native language biases when implicitly learning a second language, albeit under experimental conditions. However, the following question remains: How does orthographic information affect learning non-native phonemes? In the next section, an review of the research on the relationship between orthographic representations and L2 phonology is presented.

1.3 Orthography and second language phonology

Interest in the interaction between orthographic information and auditory forms has increased recently, especially given the limited explanations offered by the traditional models and theories of phonological acquisition about the intricate nature of language learning. Among the models discussed, only PAM-L2 considers orthography as a potential influence in L2 phoneme acquisition. Best and Tyler (2007) attributed the tendency of French learners to produce the English [ɹ] as their native voiceless uvular fricative [ʁ] to the fact that both phonemes are represented by the same letter in both languages. The lack of experimental scrutiny of the potential role of orthography in second language learning has prompted some researchers to explore the relationship between written representations and auditory input in L2 phonology. Overall, previous studies have found that the orthographic input may have a positive, negative or no effect at all on learning non-native speech sounds based on factors such as L1 orthography and the level of familiarity with L2 writing system. The following subsections will discuss the nature of the effects of L1 orthography and the degree of familiarity with L2 orthography on the acquisition of L2 phonemes.

1.3.1 The influence of L1 orthographic representations on L2 phonological acquisition

In order for orthography to inhibit the acquisition of non-native phonemes, the orthographic forms of the learners' native phonemes have to be incompatible with the auditory input of the target language. The evidence for the interference of L1 orthography was provided by a longitudinal study of native English learners of German (Young-Scholten & Langer, 2015). The researchers tracked the phonological development of three Americans learning German as a second language over a one-year period. The linguistic item of interest was the word-initial /z/, which is orthographically represented by the letter <s> in German. The analysis of both spontaneously produced and naturally elicited data showed that the American learners' realization of the phoneme /z/ mostly corresponded with their representation of the letter <s> word-initially in their native English, that is [s]. This finding suggests that orthographic forms may curb the influence of the auditory input on the the learners' ability to acquire second language phonemes. This is particularly interesting because the target phoneme has an equivalent in the English phonemic inventory; nonetheless, the study concluded that the mismatch in the orthographic representations between German and English caused learning difficulties. Earlier, Young-Scholten (2002) also found that English-speaking learners of German produced the word-final devoiced obstruents with a voicing quality, possibly due the of the influence of the orthographic-auditory correspondences of their native language.

Inconsistency in the spelling norms of L2 phonemes can also lead to learning difficulties. Evidence pointing to the hindering effect of the orthographic input was found in a study conducted on 18 participants, most of them were native English-speaking learners of Chinese (Bassetti, 2006). These learners were exposed to the pinyin orthographic forms, a somewhat phonologically transparent Chinese writing system that uses Roman characters, for an average of eight months. The researcher examined the learners' phonological acquisition of three Chinese rimes [uei], [iou] and [uən] using phoneme counting and phoneme segmentation tasks. These rimes are spelled without the central vowel after an onset consonant (e.g., [kuən] is spelled <kun>). The results showed that the number of phonemes counted and segmented in the consistently spelled syllables were significantly higher than the number of phonemes counted and segmented in the syllables that lacked the central vowel in the pinyin spelling. The researcher concluded that difficulties in acquiring L2 sounds are not only caused by the learners' L1 phonology but rather by a combination of L1 bias and the orthographic input of L2.

Bassetti and Atkinson (2015) explored how the irregularities of L2 orthographic forms affect the production of L2 phonemes among advanced Italian-speaking learners of English. In a reading task, 85% of the experienced learners pronounced an added phoneme in words containing silent letters, while 56% produced these letters in a repetition task where both the spellings of the words and their pronunciations were presented. In a second experiment, the Bassetti and Aktinson investigated whether vowels represented by double letters would be produced with longer duration than vowels represented by a single letter (e.g., <moon> vs. <june>). The results confirmed the negative influence of orthography. The learners' production of vowels spelled with double letters were longer than vowels spelled with a single letter. Moreover, Bassetti and Atkinson (2015) examined the production of the allophonic variants of the morpheme <ed>, which marks the past tense and the past participles ([d], [əd] and [t]). The learners were given the written base form of 21 regular verbs and were asked to produce their past tense and past participle. The researchers found that L2 learners produced /d, t/-ending verbs with an added vowel, while their realization of verbs ending with /t/ mostly contained /d/ or /əd/. These results indicate that learners could be negatively influenced by the orthographic forms of L2 sounds.

The hindering role of the orthographic input in learning new phonemes has been also attested in training studies. Hayes-Harb et al. (2010) examined native English speakers' ability to learn new words based on the following information: auditory input, pictures indicating meanings and orthographic input. Three groups of learners were given the same auditory input matched with corresponding images, but they differed with regard to orthography. Crucially, the first group was presented with spellings consistent with English norms (e.g., <kamad> for [kɑməd]), while the second group received spellings that were inconsistent with English (e.g., <faza> for [faʃə]). The third group was not exposed to the orthographic forms of the auditory input. All three groups were given the same picture-matching task where an auditory form of a word was accompanied by a picture. The learners were asked to decide whether the picture matched the auditory token or not based on the training they received. The results showed that the learners who were presented with incompatible orthography were less successful in matching auditory tokens with their corresponding pictures, which indicates that inconsistency between L1 and L2 grapheme-phoneme correspondences could negatively impact learners' ability to lexically encode new words.

However, an eye-tracking experiment found that the availability of orthography improved learners' ability to acquire a non-native contrast. Escudero, Hayes-Harb, and Mitterer (2008) trained 50 highly proficient Dutch learners of English to learn the problematic English /æ-ɛ/ contrast using non-words. The participants were divided into two groups: an auditory-only group and an auditory-orthography group. During training, the participants were presented with an auditory form along with its corresponding image as well as the written representation of the non-word only in the auditory-orthography group (e.g., <gebbet> for /gɛbət/ and <gabble> for /gæbəl/). In Dutch, the graphemes <e> and <a> correspond to the phonemes /ɛ/ and /ɑ/, respectively. Following the training, the participants engaged in an eye-tracking task where an auditory form was played and two images were presented; they had to fixate on the correct image associated with the auditory form. The results revealed that participants who were exposed to only the auditory forms looked at both images associated with words containing /æ/ and /ɛ/ when they heard a word with either /æ/ or /ɛ/. In contrast, the auditory-orthography group fixated only on images associated with words containing /ɛ/ when they heard a word with /ɛ/, but when they heard tokens with /æ/, these participants looked at images that were associated with words containing /æ/ or /ɛ/. The researchers concluded that L2 learners, aided by their L1 phoneme-grapheme correspondences, utilize the available orthographic representations of the auditory input in order to acquire the target contrast.

Other studies have reported mixed results regarding the nature of the influence that L1 orthography exerts on the acquisition of L2 phonemes. Escudero and Wanrooij (2010) examined the effects of L1 orthographic input on the perception of six Dutch vowels /a ɑ i ɪ y ʏ/ by Spanish-speaking naive listeners. The Spanish vowel inventory has only five monophthongs /a i e o u/. Therefore, it was predicted that Spanish learners would have difficulty perceiving the following Dutch vowel contrasts: /i-y/, /ɪ-ʏ/, /i-ɪ/, /y-ʏ/ and particularly /a-ɑ/ because they would be most likely assimilated to their single native categories /a/, /i/ and /u/, respectively. In terms of orthography, Spanish has a transparent writing system as opposed to Dutch orthography in which a grapheme can represent more than one phoneme. Based on these discrepancies, both orthographic and phonological, the researchers hypothesized that the Spanish learners' L1 writing system would interfere with their ability to process L2 non-native contrasts. To test this hypothesis, an auditory XAB discrimination task was performed first on the perception of the following five vowel contrasts: /a-ɑ/, /i-ɪ/, /i-y/, /ɪ-ʏ/ and /y-ʏ/. The results of the auditory task confirmed that the Spanish-speaking listeners were not able to accurately discriminate the Dutch novel contrasts, particularly /a-ɑ/ and /ɪ-ʏ/. To see whether their ability would improve or worsen with the presence of orthography, a second task was conducted in which the listeners were given an XAB test along with the Dutch spellings of the auditory stimuli. Here, the participants heard a vowel and had to decide if it was similar to the second or third auditory tokens accompanied by their orthographic representations in Dutch. The results showed that the orthography had differential effects: the listeners displayed significant improvement in their perception of /a-ɑ/, whereas they performed worse in discriminating /i-ɪ/ and /y-ʏ/, although the differences were not significant. The researchers attributed the improvement in the perception of the /a-ɑ/ contrast to the alignment of both the acoustic and the orthographic information. Acoustically, /ɑ/ is longer than /a/ in Dutch and this durational difference is marked orthographically with double letters (<a> for /a/ and <aa> for /ɑ/). For the remaining contrasts, access to orthography neither helped nor hindered the ability of the Spanish listeners to perceive these contrasts.

Escudero et al. (2014) extended their examination of the influence of orthography on the acquisition of Dutch vowels by looking at both monolingual and experienced Spanish-speaking listeners. In this study, the participants were trained to learn novel words containing perceptually easy and difficult vowel contrasts that had either similar or different orthographic representations in both Dutch and Spanish. For instance, the Dutch vowels /ɪ-ʏ/ are perceptually similar to the Spanish /i-u/; moreover, both pairs have identical orthographic forms in Dutch and Spanish: <i> for /ɪ-i/ and <u> for /ʏ-u/. Conversely, the Dutch /ɑ-a/ contrast is perceptually difficult because it has one equivalent in Spanish /a/, and it is orthographically represented by two graphemes (<aa> and <a> as opposed to one in Spanish <a>). The training involved one group of the participants listening to non-words containing these vowels and simultaneously seeing their corresponding pictures, while the other group was exposed, in addition to auditory forms and pictures, to the written forms of these items. To test their acquisition of these contrasts, the participants heard an auditory form and were asked to choose the picture that corresponded with the non-word. The results showed that the performance of the participants correlated with the L1 grapheme-phoneme correspondences; both monolingual listeners and learners performed worse in identifying novel words that featured incompatible orthographic forms. Conversely, the participants who were exposed to

Dutch non-words with orthographic forms that were compatible with their L1 orthography performed better than those in the auditory group, with the experienced learners achieving more accuracy than the naive listeners.

In a follow-up study, Escudero (2015) evaluated the performance of Australian English listeners in learning novel minimal pairs containing the same Dutch vowels and compared it with the Spanish listeners' performance in (Escudero et al., 2014). The grapheme-phoneme correspondences in English are not as transparent as in Spanish. For instance, the Dutch /ɑ-a/ are orthographically represented in Dutch by <a> and <aa>, respectively, while their closest equivalents in Australian English /ɔ-ɐː/ are spelled with <o> and <ar> in English. The participants underwent the same learning phase where they heard a non-word containing one of the Dutch vowels while seeing its corresponding picture. In addition to the auditory forms and their corresponding pictures, the participants in the orthography group were also exposed to the Dutch spellings of these words. The results of the follow-up test showed that access to orthography only improved the learning of two contrasts which were already perceived with intermediate to high accuracy in the auditory group, indicating that orthography reinforces learning when the contrast is relatively easy to perceive. For the remaining contrasts, the availability of orthographic information neither helped nor hindered their acquisition.

Overall, these studies demonstrate that orthography can have a positive effect on phonological acquisition and/or word learning when the target contrast is already easy to discriminate and the orthographic input is compatible with the learner's L1.

1.3.2 Unfamiliar orthography and L2 phonemes

The level of familiarity with the orthography of the target language can also determine the learning outcome. Mathieu (2014) explored the impact of orthographic familiarity on the lexical encoding of non-native phonemes. In the first set of experiments, native English speakers were trained to perceive the Arabic voiceless uvular and pharyngeal fricatives /χ/- /ħ/ and were assigned to five groups. The first group was exposed only to the auditory tokens along with pictures indicating their meanings, while the other groups were additionally presented with orthographic labels varying in their degrees of resemblance to Roman orthography. After the training, all participants underwent a phonological test in which they were asked if the picture presented matched the auditory token played. The results demonstrated that generally, the participants exposed to Arabic orthographic forms performed significantly worse than the participants who only received auditory input. Mathieu (2014) attributed the poor performance of the participants in the orthography group to the combined effect of the complexity of the contrast and the unfamiliarity of the script.

The same result was obtained when using orthographic labels that were more familiar to native English speakers' writing system; participants who were exposed only to the auditory tokens performed significantly better than the participants who were presented with the Cyrillic spellings and of these tokens. In order to increase the level of familiarity, a phonologically transparent Roman orthography with added superscripts (a bar and a dot to mark the phonological contrast) was introduced to a third group, but the learners who were trained only on auditory tokens still performed significantly better. The only two conditions in which the inhibitory effect of orthography was neutralized occurred when the two phonemes were contrasted orthographically by diacritics (e.g., <x̃al> vs. < ẋal>), and when a hybrid system consisting of a word-initial Cyrillic letters followed by Roman letters was used. In both conditions, the performances of the participants who were exposed only to auditory input were not significantly better than those who also were exposed to orthographic information.

The second set of experiments was designed to test monolingual English speakers' ability in learning novel words containing the Japanese consonant length contrast. Overall, orthography exerted no influence when completely unfamiliar (Hiragana symbols) and familiar (Roman letters) scripts were presented. These findings were in contrast to the results on the acquisition of the Arabic /χ/ and /ħ/ where Arabic and Roman alphabets negatively affected learning the contrast (Mathieu, 2014).

Similarly, Showalter and Hayes-Harb (2015) trained monolingual English speakers to learn Arabic non-words containing the velar-uvular stop contrast /k/ and /q/. The participants were assigned to two groups: a group that heard the auditory forms and simultaneously saw their Arabic spellings and pictures indicating their meanings; the second group only lacked the orthographic information. Following the training, the participants were tested on their acquisition of the contrast by deciding whether the words they heard matched the pictures they saw. The results showed that both groups performed similarly, which suggests that Arabic orthography neither improved nor hampered the learners' ability to learn the /k/-/q/ contrast. The follow-up experiment attempted to familiarize the English learners with Arabic orthography by instructing them on the direction of the script (right-to-left as opposed to left-to-right) and the number of letters in each word, but the results did not reveal any effect of that instruction on their ability to learn the contrast. Another subsequent experiment substituted the Arabic script with Roman alphabet (e.g., /kubu/ is spelled <kubu> while /qubu/ is spelled <qubu>). Because of the L1 grapheme-phoneme correspondences (both /k/ and /q/ are orthographically represented by the letter <k>), the participants who were exposed to Roman spellings performed significantly worse in identifying the Arabic velar-uvular contrast than the control group and the Arabic script group.

These findings suggest that while unfamiliar orthography can generally hinder learning L2 phonemes, the effect of the two sources of information on learning is far more complex. Indeed, there is even evidence that unfamiliar writing system can improve learning. Showalter and Hayes-Harb (2013) examined the ability of monolingual English speakers to learn eight Mandarin non-words containing contrastive lexical tones. The participants were assigned to the "No Tone Marks" and "Tone Marks" learning groups. Both groups heard an auditory form containing a lexical tone along with its corresponding picture as well as the written form. The "No Tone Marks" group was presented with a pinyin spelling that lacked diacritics, whereas the "Tone Marks" group saw a spelling that contained a diacritic that marked the tone type (e.g., <gi> vs. <gì> for a high-falling tone). In the first experiment, the learners were asked to decide whether the auditory form they heard matched the correct image based on the training they had. The results revealed that the participants in the "Tone Marks" group performed significantly better than the participants in the "No Tone Marks" group, which indicates that unfamiliar orthography can help learning non-native speech sounds. The second experiment was intended to determine if improved learning was driven by the availability of diacritics. Instead of pictured meanings, participants were presented with an auditory form along with its spelling and were asked whether they matched or not. Although their accuracy in matching the diacritics with their corresponding auditory tones was not significantly high, the participants in the "Tone Marks" group outperformed their counterparts in the "No Tone Marks" group. The researchers concluded that partially unfamiliar orthography may help L2 learners to acquire non-native speech sounds.

In sum, our understanding of the role of orthography in language learning has grown because of the recent interest in the relationship between written forms and L2 phonology. The previous research has identified a number of cases in which orthography may influence the acquisition of non-native phonemes. First, in languages that share similar writing systems, the incompatibility between L1 and L2 grapheme-phoneme correspondences may impede learning. Second, the inconsistencies in the spelling norms in L2 orthography may also lead to a defective learning of L2 phonemes. Third, the exposure to completely foreign orthography can hinder learners' ability to acquire non-native sounds. However, research also has shown that the type of contrast being learned may determine whether orthography is influential or not. For example, learning two contrastive L2 phonemes that lack counterparts in L1 may prove to be too difficult for orthography to have any effect. Finally, orthography can improve learning if L1 and L2 grapheme-phoneme correspondences are compatible, L2 spellings are consistent and if the target language employs a partially familiar writing system, such as the Roman alphabet with diacritics.

1.4 Research objectives and questions

Second language learners are exposed to multiple and diverse cues—both linguistic and non-linguistic—during the early stages of L2 phonological development. Chief among them, particularly for literate learners in instructional settings, is the written representations of speech sounds. Although recent research has advanced our understanding of the role of orthography in L2 phonological acquisition, there are still many areas in L2 orthography-phonology interface that need to be explored. The main objective of this dissertation is to understand how orthographic cues interact with the perceptual and lexical learning of non-native speech sounds. In particular, this research intends to probe the effects of orthographic compatibility and familiarity on the acquisition of the consonant length contrast via both statistical and lexical learning. Orthographic compatibility is determined by whether the written cues support or contradict the auditory statistical information in distributional learning; and the auditory lexical information in word learning. Orthographic familiarity is defined by whether L1 and L2 share the same orthographic system. Three main questions will be addressed in this dissertation:

1. Does the exposure to orthographic cues facilitate the distributional and lexical learning of the consonant length contrast?

2. If the orthographic information contradicts other cues, does orthography override the effect of statistical and lexical learning in a similar way?

3. Does the level of familiarity with L2 orthography determine the learning outcome? If the target language employs an unfamiliar alphabet with diacritics, does the availability of diacritics neutralize or even reverse the negative influence of unfamiliar orthography in L2 phonemes?

1.5 Outline of the dissertation

Chapter (2) presents three experiments that tested native English speakers' ability to learn the consonant length distinction via distributional learning. Chapter (3) describes two experiments designed to examine the impact of orthographic compatibility and familiarity on the distributional learning of consonant length contrast. The first experiment (3.2) tested monolingual English speakers' perception of the length contrast following the exposure to either a unimodal or bimodal distribution paired with compatible or incompatible Roman spellings. The second experiment (3.3) implemented the same training procedure with Arabic orthography. Chapter (4) examines the effects of the same variables—orthographic compatibility and familiarity—on the lexical encoding of the length contrast. Two word learning experiments were designed; Experiment One (4.2) presented 12 novel words and their pictured-meanings paired with either compatible Roman orthography, incompatible Roman orthography or no orthography at all; the second experiment (4.3) presented the same words with Arabic orthography instead. Finally, Chapter (5) summarizes the main findings in this dissertation and revisit the research questions, offering interpretation and future research directions.

Chapter 2

The Distributional Learning of an L2 Consonant Length Contrast

2.1 Introduction

The purpose of this chapter is to uncover the role of orthography in perceptual learning. The research is driven by the observation that naive listeners are able to infer non-native sound categories based on the type of distribution they are exposed to, and the difficulty L2 learners persistently face when learning non-native phonemes indicates that other factors, such as orthography, may have a contributing role in L2 phonological acquisition. In order to examine this relationship, there is a need to first establish the role of statistical learning in the acquisition of the consonant length contrast. This chapter reports three experiments demonstrating whether exposure to the distributions of sound categories in the input influences the perceptual learning of the length contrast.

2.2 Experiment One

The first experiment attempts to replicate the study conducted by Pajak and Levy (2011) on the statistical learning of the consonant length contrast by naive English learners. Their findings showed that learners in the bimodal group were significantly more accurate in their perception of the contrast than the learners who were exposed to a unimodal distribution. The design of the speech materials and the training and testing procedures closely follows their study. Based on their findings, it is predicted that a similar pattern will emerge; learners exposed to a bimodal distribution will exhibit more sensitivity to the length contrast than learners who have unimodal exposure .

2.2.1 Methods

2.2.1.1 Participants

Twenty four native English speakers (20 females and 4 males) with an average age of 20.8 (SD= 1.4, range= 18-22) were recruited from the University of Alberta Department of Linguistics participant pool to take part in this experiment. The participants received course credits for their participation. A screening process was implemented to ensure that the participants had not been exposed to any language that utilizes the consonant length distinction phonemically. Four of the participants were monolingual English speakers; the remaining participants spoke, along with their native language, one or multiple languages with varying degrees of proficiency. These languages included French, German, Mandarin, Cantonese, Tagalog and Spanish. None of the participants reported any hearing or speech impairments.

2.2.1.2 Training materials

A length continuum was created for each of the consonant /m n l tˤ/ using naturally produced tokens. The items were recorded by a 30 year old male Arabic speaker from the Al-Ahsa region in Saudi Arabia. The recording was made in a sound-attenuated booth at Alberta Phonetics Laboratory, using an Alesis ML9600 Masterlink CD recorder with a head-mounted microphone (Countryman E6). The sampling rate was set at 44.1 kHz (16 bit, mono). The materials were presented as a series of slides on a computer screen and each slide contained one word written in Arabic. The words were in a CVCV structure and the target consonant occurred word-medially. The singleton and geminate consonants were orthographically differentiated by the addition of a diacritic, referred to as *shaddah* (pronounced as /ʃaddah/ and spelled in Arabic as <شدَّة >). The *shaddah*, which is placed above the consonant as in <تّ>, is used to disambiguate words that are spelled identically but have different meanings (e.g., < علم> /ʕalam/ "flag" vs. < علّم> /ʕallam/ "he taught").

The training materials were recorded three times. In order to create a length continuum for each consonant, the most naturally produced words, as determined by a native Arabic speaker, from the recordings were chosen and segmented using PRAAT (Boersma & Weenink, 2018). The criteria adopted in segmenting the target sounds generally followed those of Turk, Nakai, and Sugahara (2006). In the case of stop closure duration, the boundary was placed where the amplitude sharply drops and the formants of the preceding vowel fully disappear. The offset of the stop closure was marked by the presence of a burst release in the waveform (See Figure 2.1). For the duration of nasals, the boundaries were marked by the abrupt decline in amplitude between the closure onset and release. For the lateral /l/, the change in amplitude was also adopted in determining the boundaries of the closure. An example of a segmented stimulus is shown in Figure 2.2.

Following the segmentation process, a PRAAT script was used to create a duration continuum of eight equal steps differing by 15 ms for each consonant. The endpoints of each continuum were set at 100 ms for a singleton and 205 ms for a geminate. Items 1-4 were created form a naturally produced singleton, while items 5-8 were extracted from a naturally produced geminate. Filler items were recorded using different short consonants. The tokens had the same structure as the target materials and consisted of the following consonants /k f s d t ʃ b g /. Finally, the items were normalized for 78 dB amplitude.

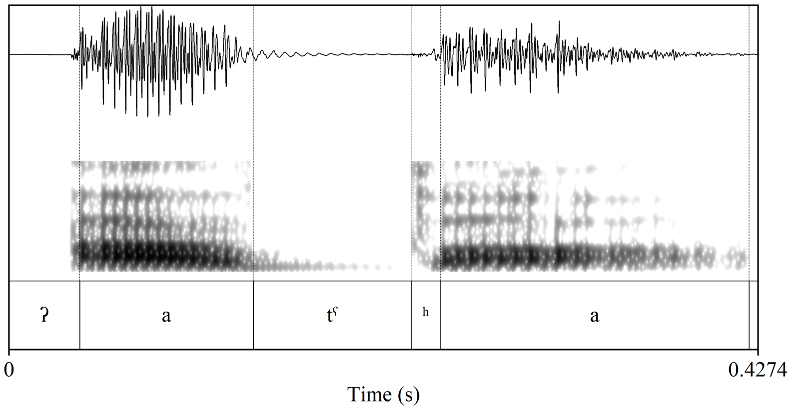


FIGURE 2.1: A waveform and spectrogram for the auditory form /ʔatˤa/. The boundaries for the pharyngealized stop were placed between the offset of preceding vowel (marked by the sharp drop in amplitude) and the onset of the burst release.

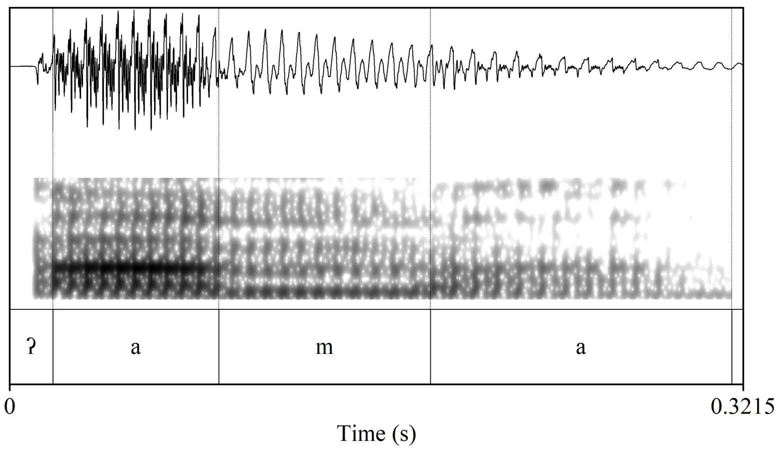


FIGURE 2.2: A waveform and spectrogram for the auditory form /ʔama/. The boundaries for the nasal were placed according to the abrupt spectral changes relative to the preceding and the following vowels.

2.2.1.3 Training phase

The participants were assigned to the following learning conditions: Unimodal and Bimodal. In the Bimodal group, the participants were presented with tokens 2 and 7 four times more often than any token (see Figure 2.3), indicating that the language they heard had a length contrast. Conversely, the subjects in the Unimodal group heard tokens 4 and 5 four times more often than any token, illustrating that the language they were exposed to did not use segmental length contrastively, as shown in Figure 2.3. Both groups were presented with the same filler tokens. The participants went through four blocks of training; each block contained 64 target tokens and 32 fillers, bringing the total number of items to 384. The participants controlled the presentation of items in order to focus on the task. Each token played when a button was pushed after a short delay of 1 second.

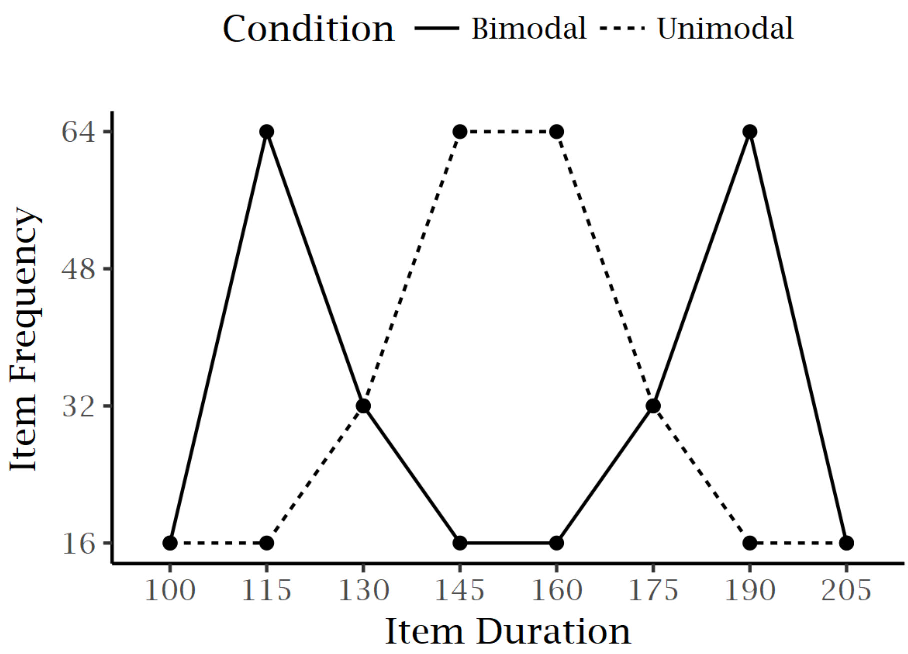


FIGURE 2.3: An illustration of the frequency of exposure to the duration continuum for the Unimodal and Bimodal groups.

The training was administered using E-Prime 2.0 (Psychology Software Tools, 2012). The participants were seated in front of a computer screen inside a sound attenuated booth and were fitted with over-the-ear headphones (MB QUART QP-805 HS). The materials were played in a random order and the training lasted for approximately 10 minutes.

2.2.1.4 Testing phase

Following the training phase, all participants performed an AX discrimination task, which was used to assess their perceptual phonological acquisition based on the type of training they had received. The participants listened to a pair of auditory tokens separated by 700 ms inter-stimulus interval, and had to decide whether the items they heard were the same or different. The pairs consisted of different combinations of tokens 1 and 8: (1-8, 8-1) for different pairs or (1-1, 8-8) for identical pairs.The same training materials were used to test the learners' performance on the consonant length contrast. The participants also were tested on the fillers items that consisted of identical pairs (e.g., /aka/ vs./aka/) and different pairs (e.g., /aka/ vs. /aga/). The testing phase consisted of four blocks, and each block contained 64 randomly presented trials. Half of the trials were target pairs and the other half were fillers. The duration of the testing phase was approximately 15 minutes. Prior to the training phase, the participants practiced the AX task for 5 minutes.

2.2.2 Data visualization and statistical analysis

The data in this dissertation were visualized using raincloud plots (Allen, Poggiali, Whitaker, Marshall, & Kievit, 2018); a combination of traditional plots that conveys the key statistical information in a transparent and informative way. The plot includes probability density; medians and means with confidence interval; and raw data from each individual subject (see Figure 2.4).

The statistical analysis in this dissertation was conducted using mixed-effects logistic regression models (Baayen, Davidson, & Bates, 2008; Morrison, 2007), which was performed in the R software (version 3.5.1) environment (R Core Team, 2016) using the*lme4* package version 1.1-17 (Bates, Mächler, Bolker, & Walker, 2015). Unlike ANOVA, mixed-effects logistic regression models require no transformation of the categorical dependent variable (see Jaeger (2008) for an explanation of the problems with using ANOVA for categorical data analysis). In addition, both categorical and continuous independent variables can be analyzed in the same model, and participants and items can be included as random effects.

The coefficients of the logistic regression models were converted to odds ratios to provide a more intuitive interpretation of the model output. Odds ratios compare the odds that an event will happen under two different conditions. Odds are the probability that an event will occur divided by the probability that the same event will not occur. For instance, in an AX discrimination task the odds of accuracy is the proportion of correct responses divided by the proportion of incorrect responses. The odds ratio compares the odds of accuracy for one group versus another by dividing the two odds. If the odds ratio is greater than 1 then the exposure is associated with higher odds of accuracy, and if the odds ratio is lower than 1 then the exposure is associated with lower odds of accuracy. An odds ratio that equals 1 indicates no effect of the exposure on the odds of accuracy.

2.2.3 Results

The accuracy on filler items is presented in Figure 2.4 for both the Unimodal and the Bimodal groups. Overall, both groups performed near ceiling on filler items, particularly on identical pairs. On filler Different pairs, the participants in the Unimodal group were slightly less accurate than their cognates in the Bimodal group; however, the difference was not statistically significant.

For target items, the proportion of correct responses on the consonant length contrast was calculated for each participant in both the Unimodal and Bimodal group and then presented in Figure 2.5. On Same pairs, the performance of both groups was similar; all participants achieved near-perfect accuracy scores. On Different pairs, however, participants in the Unimodal and Bimodal groups had differential success in perceiving length contrast. A closer inspection of the plots reveals that the accuracy of the participants in the two groups varies greatly, with the majority of the accuracy scores lower than 50%, particularly for the Bimodal group.

To examine whether statistical learning is associated with the accuracy of the perception of the length contrast, the responses to the discrimination test on Different pairs were modeled using mixed-effects logistic regression. The dependent variable

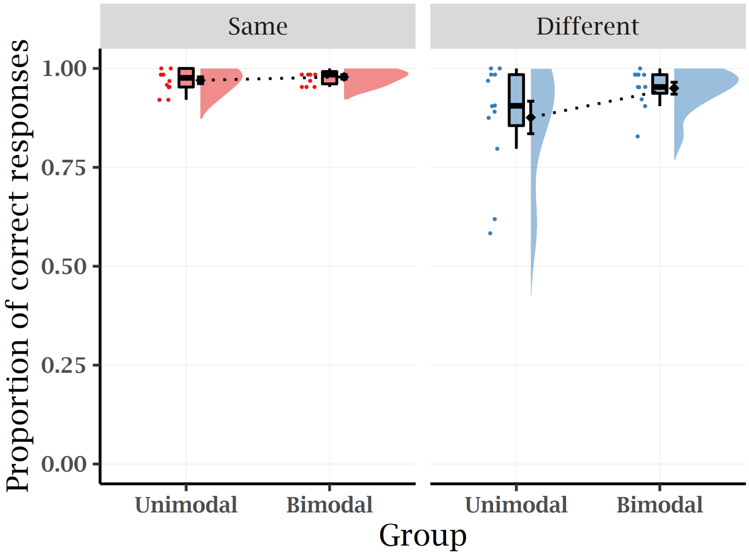


FIGURE 2.4: A combined plot representing the accuracy of both the Unimodal and the Bimodal groups on filler items. The plot contains probability distribution (split-half violin); box-plots and line-plots visualizing the means with standard errors for each group. Data from each individual are represented as points.

was the responses to the AX discrimination task on target Different pairs (Same vs. Different). The fixed-effect factor was GROUP (Unimodal vs. Bimodal). SUBJECT and ITEM served as random-effect factors. Of the 1472 trials, 26 none-responses were eliminated from the model (1.8% of the data).

According to the model, GROUP was not significantly associated with the degree of accuracy in the perception of the consonant length contrast. The performance of the Bimodal group was not significantly different from the Unimodal group (β= -1.18, *SE*= 0.95, *p*=0.21).

2.2.4 Discussion

The present experiment aimed to replicate the findings obtained in Pajak and Levy (2011). The experimental design followed closely the statistical learning paradigm implemented in Maye and Gerken (2000). The results of this experiment failed to show any significant effects for statistical learning in the acquisition of the consonant length contrast.

The unsuccessful replication of the learning effect observed in Pajak and Levy

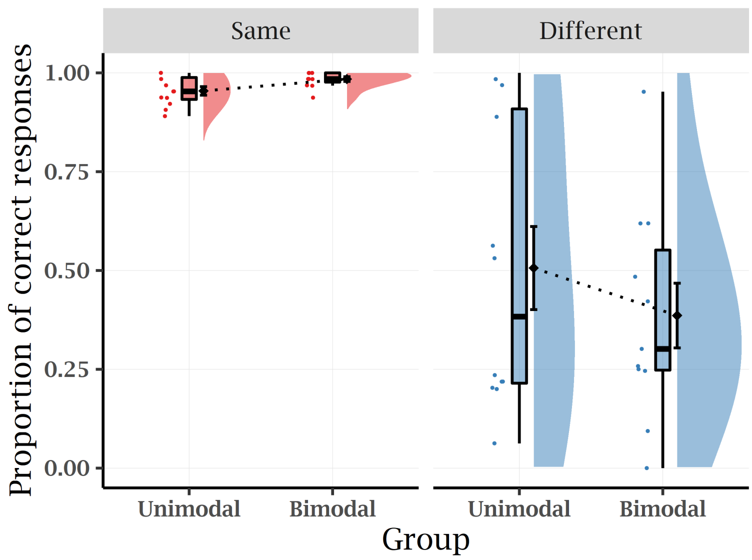


FIGURE 2.5: A combined plot representing the accuracy of both the Unimodal and the Bimodal groups on target items. The plot contains probability distribution (split-half violin); box-plots and line-plots visualizing the means with standard errors for each group. Data from each individual are represented as points.

(2011) may be attributed to two methodological differences: Population size and population sampling. The current experiment used a relatively smaller sample size compared to the original study which was 48 participants. However, there were no indications in the current data that the performances of the Unimodal and the Bimodal groups were in the same direction as the original study. On the contrary, the summary statistics suggest that the accuracy of the Bimodal group was lower than the Unimodal group. The other potential reason for this discrepancy was the population sampling. The current study included both monolingual and bilingual native English speakers; the participants in Pajak and Levy (2011) were only monolinguals. A separate analysis conducted to compare the performances of monolinguals and bilinguals also failed to detect any difference between the two groups of learners.

One possible reason why this experiment did not detect any evidence for learning on the basis of the distributional information was the level of processing involved in the AX discrimination task. In general, a short ISI allows listeners to utilize an auditory level of processing to make judgment about the stimuli they hear, while a longer ISI forces listeners to rely on their established phonemic representations to perceive any incoming stimuli (Werker & Logan, 1985). The following experiment addresses this issue.

2.3 Experiment Two

Previous studies have demonstrated that the length of the inter-stimulus interval (ISI) in discrimination tasks may influence the perception of sound contrasts (Pisoni, 1973; Crowder, 1982; Werker & Tees, 1984b; Cowan & Morse, 1986; Van Hessen & Schouten, 1992). Werker and Logan (1985) have provided evidence for the existence of three modes of perception in discrimination tasks. In a series of experiments, they have demonstrated that ISI length activates different levels of processing; a short ISI (250 ms) enhances the perception of auditory differences, while a medium length ISI (500 ms) facilitates access to phonetic differences. A longer ISI (1500 ms), however, limits access to phonemic distinctions.

This line of research may explain the lack of difference between the Unimodal and the Bimodal groups in their perception of the length contrast. Specifically, the learners in the Unimodal group, who were predicted to perceive short and long consonants as members of the same phonemic category, were equally sensitive to the contrast. The relatively short ISI (700 ms) may have increased learners' sensitivity to the phonetic variants of their single phonemic category.

Therefore, extending the interval between stimuli may force learners to rely on their phonemic representations in order to process the contrast. The argument put forward states that as the interval between stimuli increases, the auditory memory of the first stimulus fades very rapidly. This in turn forces listeners to recruit their pre-established phonemic categories in order to categorize the first stimulus and compare it with the second one.

This experiment implements the same design as the previous one, except for ISI duration. It is predicted that with an increased ISI, learners in the Unimodal group will perceive the length contrast as belonging to the same phonemic category, while learners in the Bimodal group will exhibit more sensitivity to the contrast.

2.3.1 Methods

2.3.1.1 Participants

Thirty two native English speakers with an average age of 20 (*SD*= 1.7, range= 18-23) from the University of Alberta Department of Linguistics participants pool took part in this experiment for course credit. The participants were comprised of 18 monolinguals and 14 bilinguals (other languages spoken included French, Mandarin, Spanish, Filipino, Ukrainian and Cantonese). None of the bilingual participants spoke a language that has a phonemic consonant length contrast. None of the participants reported hearing or speech impairments.

2.3.1.2 Training materials

The same training stimuli described in Section 2.2.1.2 were used for this experiment.

2.3.1.3 Training phase

The same training procedure was implemented in this experiment. The participants were assigned to the same experimental groups; 16 for the Unimodal group and 16 for the Bimodal group.

2.3.1.4 Testing phase

The participants were tested on the same items described in Section 2.2.1.4. The only difference was the ISI which was was set at 1300 ms, as opposed to the 700 ms used in the previous experiment.

2.3.2 Results

The performance on filler items for each group is shown in Figure 2.6. The majority of learners in both groups achieved near-perfect scores on same filler pairs, while the performance of both groups on different pairs exhibited some variability. No significant differences between the two groups were found.

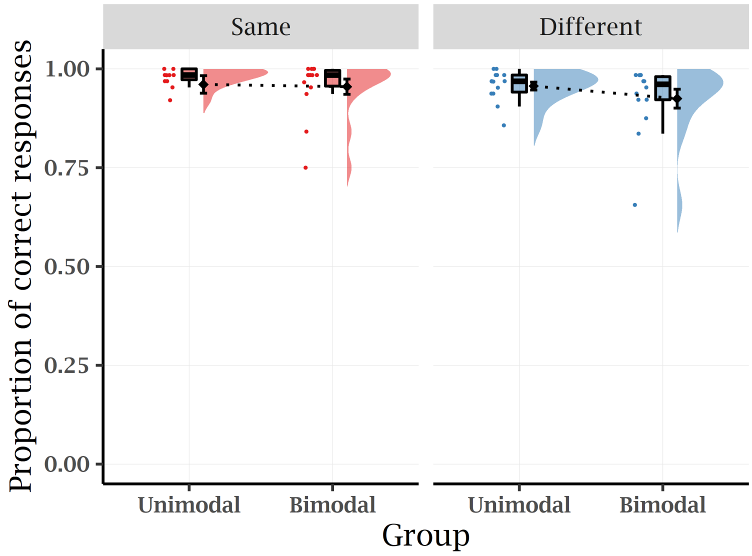


FIGURE 2.6: A combined plot representing the accuracy of both the Unimodal and the Bimodal groups on filler items. The plot contains probability distribution (split-half violin); box-plots and line-plots visualizing the means with standard errors for each group. Data from each individual are represented as points.

Figure 2.7 displays the distributions of the accuracy scores for both the Unimodal and the Bimodal groups on Same and Different target pairs. The accuracy of the two groups on Same items was similar; the learners in both groups judged most of the items to be the same. The performance on Different target pairs also shows a similar pattern; both groups had accuracy scores that ranged from low to high as evident from the spread of the distributions. However, a greater proportion of the accuracy scores for the Bimodal group clusters around the lower end of the distribution.

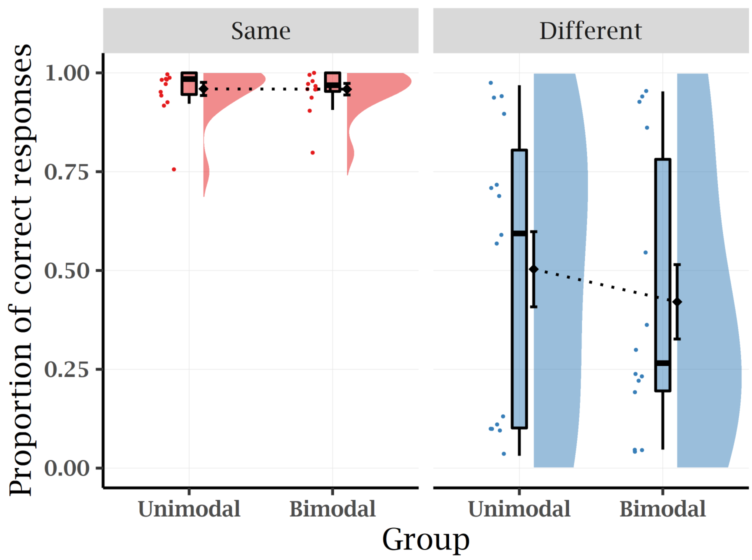


FIGURE 2.7: A combined plot representing the accuracy of both the Unimodal and the Bimodal groups on target items. The plot contains probability distribution (split-half violin); box-plots and line-plots visualizing the means with standard errors for each group. Data from each individual are represented as points.

A mixed-effects logistic regression analysis was conducted to determine whether statistical learning had any effect on the perception of the consonant length contrast. The dependent variable was the responses to the AX task on Different target pairs (Same vs. Different). The independent variable was GROUP (Unimodal vs. Bimodal). SUBJECT and ITEM served as random factors. Trials with responses were removed from the model (26 out of 1856). The results of the model did not reveal any significant relationship between statistical learning and the accuracy on the discrimination task (β= -0.39 , *SE*= 0.86, *p*=0.65).

2.3.3 Discussion

This experiment sought to test learners' sensitivity to the length contrast with a longer ISI after receiving either a unimodal or bimodal distribution of multiple length continua. It was predicted that with an increased ISI learners would recruit a higher level of processing. In particular, the Unimodal group was expected to perceive the endpoints of the length continuum as phonetic variants of the same phonemic category and, thus, show significantly less sensitivity to the contrast.

The findings of this experiment did not differ from the first one which showed that both the Unimodal and the Bimodal groups performed similarly on the discrimination of consonant length contrast. The distributional properties of the auditory input had no significant effect on the learning outcome despite lengthening the ISI to elicit phonemic processing.

This persistent lack of effect in both experiments shifts the attention to another possibility: The durational difference between short and long consonants in the previous two experiments may have been too difficult for learners to perceive and subsequently acquire. The following experiment aims to alleviate this difficulty by altering the experimental design in order to improve learning.

2.4 Experiment Three

The prior two experiments failed to provide evidence that learners could acquire sound segments by tracking their distributional properties. A possible cause of failure is the length ratio used in the training and the testing materials. In these experiments learners were trained on length continua that ranged from 100 ms to 210 ms. Data from several studies suggest that naive English listeners, and to a lesser extent learners of languages that employ durational contrast phonemically, experience perceptual difficulties when processing length contrast (Hayes-Harb, 2005; Tajima et al., 2008; Motohashi-Saigo and Hardison, 2009; Hisagi and Strange, 2011; Tsukada, Cox, and Hajek, 2014; Okuno, 2013; Porretta and Tucker, 2015).

In Hayes-Harb (2005), monolingual English listeners were asked to identify whether the token they heard contained a geminate or a singleton consonant. The tokens were extracted from length continua with endpoints at 70 ms and 310 ms. In the identification task, naive English listeners' perception of the length contrast was continuous in nature, in contrast to the categorical perception observed by native Japanese listeners and to a lesser extent by relatively experienced English learners of Japanese. A /kk/ with a closure duration of 250 ms was identified as a geminate around 50% of the time by naive English listeners, 80% by experienced English-speaking learners of Japanese and 100% of the time by Japanese listeners.

In addition, Porretta and Tucker (2015) tested naive English listeners' ability to discriminate Finnish non-words containing long and short consonants. Half the participants were instructed on how the duration of the consonant is manipulated phonemically in Finnish. A length continuum consisting of 10 equal steps was created for the word medial consonant in each word. The results of the AX discrimination task showed that while the accuracy of native Finnish control in discriminating singleton/geminate consonants with a ratio of 2.5 was above 80%, naive English listeners' accuracy was around 20%. Crucially, the participants who received instruction about the contrast in Finnish judged the singleton/geminate consonants to be different at an accuracy of around 60%, significantly higher than participants who received no such instruction.

Following these studies and the results of the previous experiments, the current experiment aims to provide training and testing materials that are neither too difficult to learn nor too easy to perceive. Therefore, the ratio between the endpoints of the length continua has been expanded from 1:2 to approximately 1:3 (details are provided in the Methods section). Further, to gain a better understanding of the role of distributional learning in the acquisition of length contrast, it is important to determine first the baseline performance of naive English listeners. This way the effects of training can be assessed by comparing the performance of learners who are exposed to either a unimodal or bimodal distribution to the performance of those who have not received any training.

Finally, in order to evaluate the robustness of distributional learning, the current experiment probes learners' ability to generalize learning to novel segments that share the same contrasting property. To this end, learners in this experiment are additionally tested on untrained singleton vs. geminate contrasts.

To sum up, this experiment attempts to answer the following questions: (I) How well can monolingual English listeners perceive the consonant length contrast, (II) do the enhanced length continua help learners track the distributional properties of the input and acquire sound segments accordingly, and finally, (III) are learners able generalize learning to other consonants that share the same phonological feature?

2.4.1 Methods

2.4.1.1 Participants

Participants in this experiment were 81 monolingual English speakers with no L2 experience beyond high school French. Seventy-five participants were recruited from the University of Alberta Department of Linguistics participant pool and received course credit, while the remaining participants (six) were recruited from the Edmonton community and were paid $10 for their participation. The participants were 28 male and 53 female speakers and had a mean age of 21 (*SD*= 5.4, range= 18-37). None had reported any hearing or speech impairments. The participants were randomly assigned to three groups; Control, Unimodal and Bimodal.

2.4.1.2 Training materials

The training stimuli used in the current experiment were recorded by the same Arabic speaker in Experiment One. However, instead of 4 length continua, the participants in this experiment were only trained on two created from non-words that contained either /m/ or /n/. These two continua were extracted from naturally produced singleton consonants in a CVCV structure, where the target consonant was embedded between two vowels. Crucially, each continuum consisted of 10 equal steps differing by 20 ms with the endpoints set at 80 ms for a singleton and 260 ms for a geminate. The singleton to geminate ratio was approximately 1:3.2.

Unlike in the previous two experiments, the filler items in the current experiment were carefully selected to match the complexity and the difficulty of the target items. This approach made it less likely that the participants were aware of the main purpose of the experiment. In both of the previous experiments, the filler items included mostly consonants that exist in the learners' L1 phonemic inventory. Here, the filler items consisted of largely Arabic consonants that do not have equivalents in English. The learners were exposed to the pharyngealized consonants /tˤ sˤ ðˤ/, the voiceless uvular stop /q/, the voiceless uvular fricative /χ/ and the voiced uvular fricative /ʁ/. The filler items were in the same structure CVCV. Both the target and the filler items were normalized for amplitude at 78 dB.

2.4.1.3 Training phase

Forty of the participants in this experiment were randomly assigned to two learning conditions; Unimodal and Bimodal (20 participants per group). Prior to commencing the training phase, the participants were told they would hear multiple words in an unidentified language and their task was to listen carefully to these words in order to learn them. Additionally, a practice test of the AX discrimination task was administered before the training phase. On completion of the practice test, which lasted for 5 minutes, the exposure phase started. For the Unimodal group (see Figure 2.8), items on the middle of the length continua (5 and 6) were heard four times more than items near the endpoints of each continuum (2 and 9). This type of exposure favored learning a single category. On the other hand, the Bimodal group was presented with items near the endpoints of the continua (2 and 9) four times more than items on the middle of each continuum (5 and 6), which in turn favored learning two categories that differed in duration. Critically, as can be seen in Figure 2.8, the frequency of exposure to items 1 and 10 was the same for both groups.

The participants went through 4 blocks of training. In each block, participants heard a total of 100 tokens (72 targets and 28 fillers), bringing the number of items to 400. The items were presented in a random order and the training lasted for approximately 12 minutes. The training was conducted in soundproof booths where participants sat in front of a computer screen and fitted with over-ear headphones

(MB QUART QP-805 HS). The experiment was presented using E-Prime 2.0 (Psychology Software Tools, 2012).

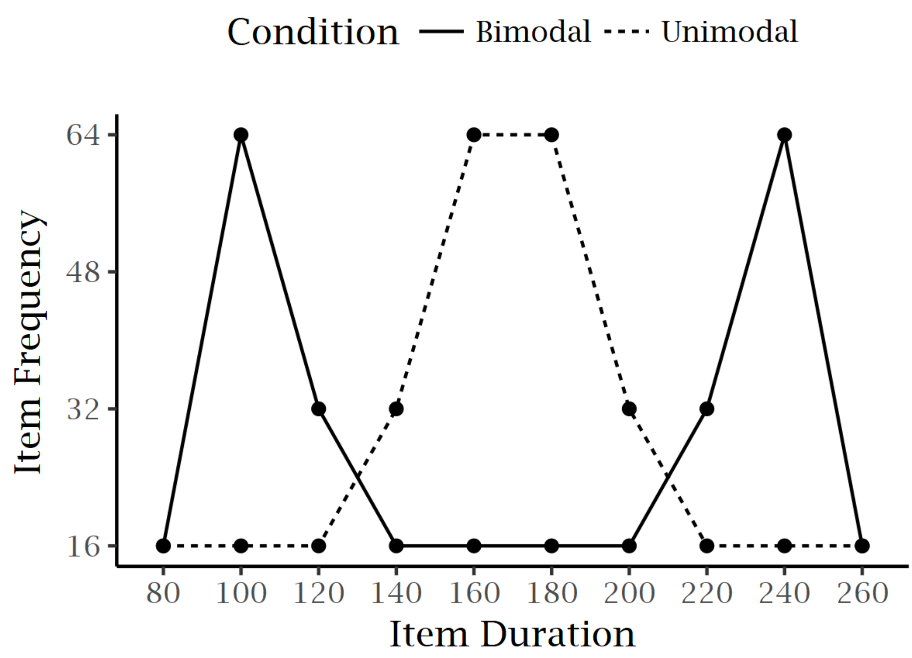


FIGURE 2.8: An illustration of the frequency of exposure to the duration continuum for the Unimodal and Bimodal groups.

2.4.1.4 Testing phase

Once the training phase was completed, each participant performed an AX discrimination task. The testing phase consisted of three blocks: Trained, Untrained and Untrained Natural. In the Trained block, the participants were tested on their perception of the same tokens heard in training, namely pairs from the endpoints of the /m/ and /n/ continua. The second block examined the participants' ability to generalize learning to untrained pairs extracted from length continua similar to the trained items. The consonants used in the Untrained block were the lateral approximant /l/ and the voiceless alveolar fricative /s/. The third testing block examined the participants' ability to generalize learning to untrained naturally produced length contrasts, which included the alveolar stop /t/ and the voiced alveolar fricative /z/ (See Table 2.1).

There were 64 trials in each block; 32 target pairs and 32 filler pairs. For target pairs, combinations of items 1 and 10 from each continuum were presented in the Trained and Untrained blocks (1-10, 10-1, 1-1, 10-10). In the Untrained Natural block, the participants were presented with pairs that consisted of short-long, long-short, short-short and long-long consonants. There were 16 same and 16 different target pairs in each block.

The filler pairs were composed of the same items used in training. The participants were tested on contrasts that differed either in voicing, place or manner of articulation. These included nonnative Arabic pairs such as the velar/uvular and the pharyngealized/plain contrasts. Similar to Experiment Two, the ISI was set at 1300 ms and the duration between each trial was 3 seconds. Participants in the Control group performed the same task but without any prior training. The testing phase lasted for approximately 5 minutes.

TABLE 2.1: A summary table of the testing blocks for target different pairs and their durations

Testing Block	**Singleton**	**Geminate**
Trained	/ama/ (80 ms) /ana/ (80 ms)	/amma/ (260 ms) /anna/ (260 ms)
Untrained	/ala/ (80 ms) /asa/ (80 ms)	/alla/ (260 ms) /assa/ (260 ms)
Untrained Natural	/ata/ (108 ms) /aza/ (82 ms)	/atta/ (272 ms) /azza/ (187 ms)

2.4.2 Results

The performance of each group on filler items is shown in Figure 2.9. Overall, all groups performed similarly on filler items. The accuracy scores were not at ceiling since some of the fillers were composed of nonnative contrasts that were difficult for participants to rapidly learn with limited exposure. No significant differences were found between the three groups on filler items.

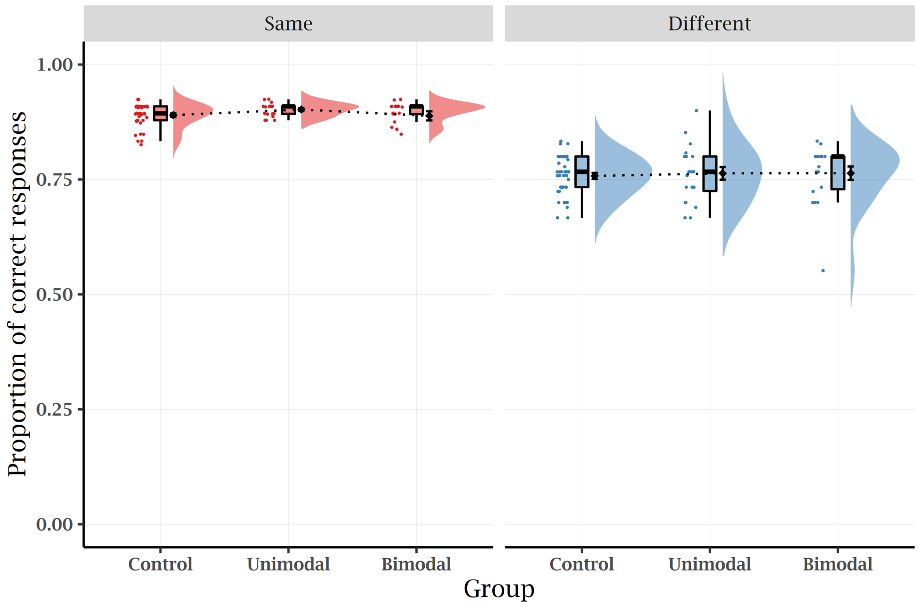


FIGURE 2.9: A combined plot representing the accuracy of the Control, Unimodal and the Bimodal groups on filler items. The plot contains probability distribution (split-half violin); box-plots and line-plots visualizing the means with standard errors for each group. Data from each individual are represented as points.

For target items, the proportions of correct responses on Same and Different pairs were calculated for each participant and then plotted in Figure 2.10. As shown in the figure, all groups had almost identical accuracy scores on Same target pairs. Of greater interest is the performance on Different target pairs. First, the accuracy scores on the length contrast for the Control group seem to display a bimodal distribution, with the majority of the participants (approximately 61%) scoring below chance. In comparison, although the accuracy scores for Unimodal and Bimodal groups exhibit variability, the majority of the learners in both groups achieved accuracy percentages above chance, with 35% of the learners in the Unimodal group and only 15% in the Bimodal group having accuracy scores below chance. In terms of means, it is apparent that both the Unimodal and Bimodal groups had higher mean accuracy values (*M*= 0.67, *SD*= 0.47 and *M*= 0.75, *SD*= 0.43, respectively) than the Control group (*M*= 0.48, *SD*= 0.50).

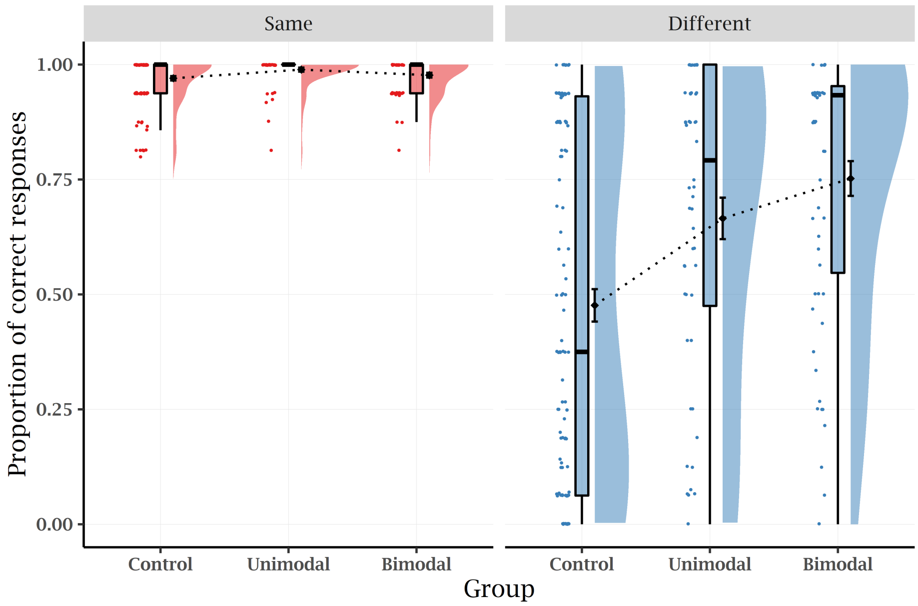


FIGURE 2.10: A combined plot representing the accuracy of the Control, the Unimodal and Bimodal groups on target items. The plot contains probability distribution (split-half violin); box-plots and line-plots visualizing the means with standard errors for each group. Data from each individual are represented as points.

To examine whether statistical learning influences the perception of consonant length contrast, the accuracy on Different target pairs was analyzed by fitting two separate mixed-effects logistic regression models. The first model was conducted to compare the performance of the two experimental groups (Unimodal and Bimodal) to the Control group. The dependent variable was the response to the AX discrimination task on Different target pairs (Same vs. Different) and the fixed predictors were GROUP (Control vs. Unimodal and Bimodal) and TEST TYPE (Trained vs. Untrained and Untrained Natural), with SUBJECT and ITEM as random intercepts and by-subject random slope for TEST TYPE. A summary of the model is presented in Table 2.2.

According to the model, although the Unimodal group had an increased odds ratio (OR 3.77, 95% CI: 0.92-15.46) of perceiving length contrast as 'Different' compared to the Control group, the difference was not significant. However, the difference between the Bimodal group and the Control group was significant. The odds of learners in the Bimodal group responding 'Different' to the AX task were 5.2 times that of the Control group (95% CI: 1.20-22.31). As for TEST TYPE, the performance

TABLE 2.2: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on target different items. The reference level for GROUP is Control, and for TEST TYPE is Trained.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	0.152	0.621	0.246	0.806
Unimodal	1.328	0.720	1.845	0.065
Bimodal	1.649	0.743	2.220	**0.026**
Untrained	-0.208	0.609	-0.341	0.733
Untrained Natural	0.070	0.621	0.112	0.911

on the Untrained and Untrained Natural pairs was not significantly different from the Trained pairs. The predicted probabilities obtained from the model are shown in Figure 2.11.

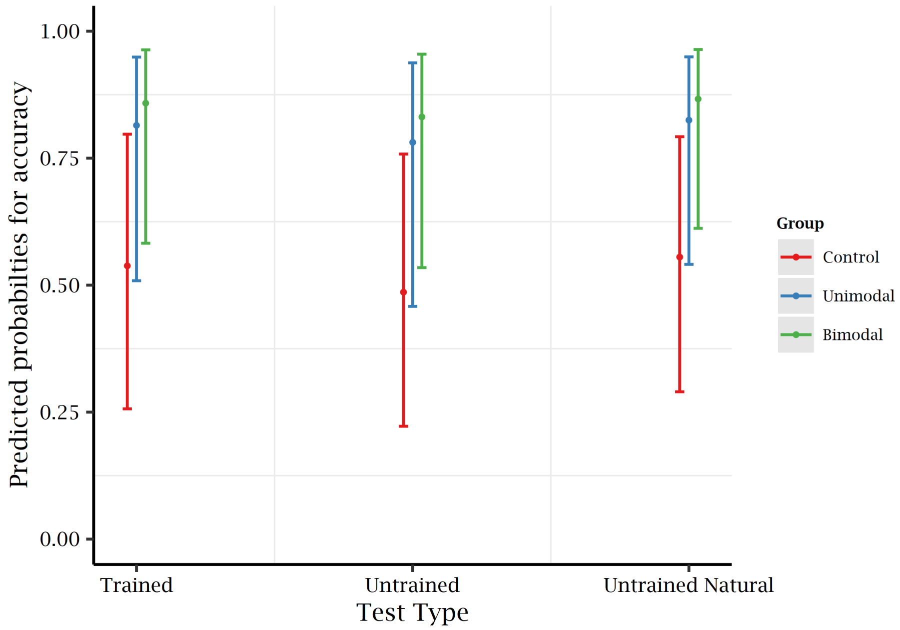


FIGURE 2.11: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast. The probabilities are presented with 95% confidence interval.

Having established that the Bimodal group was significantly better at perceiving the length contrast than the Control group, it remains unclear whether their accuracy was significantly higher than the Unimodal group. For this purpose, a second logistic mixed effects model was constructed to determine if the distributional information was associated with the performance on the length contrast. The fixed predictors were GROUP (Unimodal vs Bimodal) and TEST TYPE (Trained vs Untrained and Untrained Natural). SUBJECT and ITEM served as random effects, with by-subject slope for TEST TYPE. Table 2.3 summarizes the model.

TABLE 2.3: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on target different items. The reference level for GROUP is Unimodal, and for TEST TYPE is Trained.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	2.149	0.731	2.941	0.003
Bimodal	0.166	0.740	0.225	0.822
Untrained	-0.583	0.605	-0.964	0.335
Untrained Natural	-0.734	0.611	-1.201	0.230

The result of the analysis yielded no significant relationship between the type of distribution learners were exposed to and the accuracy scores on the AX discrimination task. Although the odds for the learners in the the Bimodal group to perceive the length contrast as 'Different' was 1.2 times that of the Unimodal group (95% CI: 0.28-5.04), the difference was not significant. Further, the difference in accuracy between Trained, Untrained and Untrained Natural pairs was not significant for learners in both groups, indicating that learners were able to generalize the newly acquired knowledge to novel segmental length contrasts.

2.4.3 Discussion

This modified experiment set out with the aim of assessing the effects of distributional learning on the acquisition of sounds categories from enhanced training stimuli. To this end, two length continua ranging from 80 ms to 260 ms were created to train monolingual English speakers on the nonnative length contrast. Follow-up AX discrimination task was administered to test learners' perceptual sensitivity to the trained stimuli as well as to novel segmental length contrasts. In addition, naive English listeners were tested on their perception of the length contrast in order to offer a better measurement of the impact of distributional learning on phonological acquisition.

With respect to the first research question, the findings of current experiment suggest that naive English listeners have differential ability in perceiving the singleton/geminate contrast. Although some listeners were able to perceive the contrast with high accuracy, the majority of naive English listeners displayed chance level discrimination. The same pattern of perception has been found in both manipulated and naturally produced consonant length contrasts.

The central question of this experiment pertains to the impact of the enhanced length continua on learners' ability to acquire the relevant sound categories compared to the Control group. The results showed that unlike the Unimodal group, learners who were exposed to a Bimodal distribution had a significantly higher sensitivity to the contrast than the Control group. The experiment, however, failed to obtain a significant difference between the Unimodal and the Bimodal groups in their discrimination of the length contrast. The performance of the Unimodal group was neither statistically better than the Control group, nor was it significantly worse than the Bimodal group. This pattern of discrimination has been also attested in novel segmental contrasts that varied in durational differences. Both the Unimodal and the Bimodal groups seem to generalize learning to the untrained synthetic and natural stimuli.

This lack of difference in performance between the Unimodal and Bimodal groups is contrary to previous studies which have suggested that adult learners can acquire non-native sound categories by tracking their distributions in the input (Maye and Gerken, 2000, 2001; Gulian, Escudero, and Boersma, 2007; Hayes-Harb, 2007; Perfors and Dunbar, 2010; Pajak and Levy, 2011; Ong et al., 2017). In these studies, learners who were exposed to a unimodal distribution showed either a significantly worse discrimination than learners trained on a bimodal distribution or no difference between their pre-training and post-training discrimination scores.

Other studies reported improvements in discriminating nonnative contrasts by comparing pretest and posttest scores of participants who were trained on a bimodal distribution and a control group who only had music exposure (Wanrooij and

Boersma, 2013; Escudero and Williams, 2014). In some studies, significant improvements in discrimination were only found when learners were trained on bimodal distribution with exaggerated stimuli (Escudero, Benders, and Wanrooij, 2011; Wanrooij, Escudero, and Raijmakers, 2013).

Perhaps the most comparable findings come from a study that trained Mandarin listeners on the nonnative Thai lexical tones (Ong, Burnham, and Escudero, 2015). In their experiment, 36 Mandarin participants were assigned to either a unimodal or a bimodal training and their pretest and posttest discrimination scores were compared. The results showed that only the bimodal group experienced significant improvement in discrimination over their pretest scores. However, the posttest scores of the unimodal and the bimodal groups were not significantly different. Likewise, the current study found that while the performance of the Unimodal and Bimodal groups were not statistically different, only the Bimodal group showed significantly better sensitivity to the length contrast than the Control group.

Still, the indistinguishable performances of the Unimodal and the Bimodal groups question the effectiveness of distributional learning in this experiment. Some researchers attributed such findings to the degree of attention sustained by learners during training (Terry, Ong, & Escudero, 2015). To examine this, Ong et al. (2015) assigned Australian English learners of lexical tones to a unimodal or bimodal group with or without a cover ask. The cover task entailed indicating on a response sheet the sequence in which randomly interspersed beeping sounds occurred during the training phase. Their results showed that only when learners performed the cover task to maintain attention during the exposure phase the bimodal group exhibited significantly superior sensitivity to the contrast than the unimodal group. While this may be true for learning lexical tones, maintaining attention during distributional training seems to be ineffective for Australian English learners of the nonnative Dutch /ɑ/-/aː/ vowel contrast (Ong, Terry, & Escudero, 2016).

In sum, the results of this experiment and the reported studies on distributional learning highlight the complex nature of this learning mechanism. Indeed, many factors including the target contrast, attentional listening and learners' L1 can influence the robustness of implicit learning. The next step is to examine whether exposing learners to orthographic information can impact the detection of the underlying distributional properties of the input.

2.5 Conclusion

The present chapter aimed to establish whether naive English listeners could acquire length contrasts implicitly by tracking their distributional properties. Three experiments were conducted. Experiment One attempted to reproduce the findings in Pajak and Levy (2011) without success. Experiment Two conducted the same experiment but with an increased ISI in order to elicit phonemic processing. Again, learners who had either unimodal or bimodal exposure displayed similar sensitivity to the segmental length contrast. Experiment Three hypothesized that the failure of the previous two experiments to induce distributional learning was possibly due to the difficulty of the learned contrast. Thus, an expanded length continua were used to train participants distributionally. Although learners in the Bimodal group were significantly more likely to perceive the length contrast as different compared to the Control group, their performance was comparable to that of of the Unimodal group. All in all, these findings offer the perfect conditions to examine how the orthographic input interacts with the seemingly unstable nature of distributional learning in adults phonological acquisition. Chapter 3 presents two experiments that investigate the effects of orthographic compatibility and familiarity in the distributional learning of the consonant length contrast.

Chapter 3

Orthographic input and distributional learning

3.1 Introduction

The previous chapter investigated the role of distributional learning in the acquisition of nonnative speech sounds. Three experiments were conducted to test whether the exposure to unimodal or bimodal distributions of length continua could help learners uncover and subsequently acquire the underlying structure of the input. Experiments 1 and 2 did not show any effect for distributional learning; learners who were exposed to the bimodal distribution failed to perceive the contrast more accurately than learners who had the unimodal exposure. Experiment Three revealed that learners who were exposed to the bimodal distribution had a significantly higher sensitivity to the length contrast than listeners who had no prior exposure. The results yielded no significant difference between the performance of the Unimodal and Bimodal groups. Overall, the findings of these experiments have failed to provide conclusive evidence that distributional exposure induces category formation for the consonant length distinction.

In light of these findings, the present chapter examines how access to orthographic information interacts with the distributional learning of consonant length contrast. Two experiments were designed to answer the following questions:

1. Does the exposure to written input facilitate the distributional learning of the consonant length contrast?

2. If the auditory and orthographic stimuli convey conflicting information, will learners rely on orthography to learn the length contrast?

3. Will similar patterns emerge if learners are exposed to unfamiliar orthography?

3.2 Experiment One

The purpose of this experiment was to examine whether the exposure to a familiar orthography—one that shares learners' native orthographic system—enhances or hinders the perceptual learning of the consonant length contrast. To this end, the following four experimental groups were designed: Unimodal Same Roman Spelling (USRS), Bimodal Distinct Roman Spelling (BDRS), Unimodal Distinct Roman Spelling (UDRS) and Bimodal Same Roman Spelling (BSRS). The main difference between these groups was whether the orthographic cues were compatible or incompatible with the distributional information. Orthographic compatibility was determined by whether the orthographic information supported or contradicted the auditory distribution. For example, in a unimodal distribution with compatible orthography, the auditory forms of each continuum had the same spellings (e.g., <ama>); while in a bimodal distribution the spellings of the auditory forms reflected the existence of a length contrast represented by two orthographic graphemes (e.g., <ama> versus <amma>). Conversely, in groups where the orthographic cues contradicted the distributional information, the auditory forms had distinct spellings (two different orthographic representations) in a unimodal distribution and the same spellings in a bimodal distribution.

It was predicted that learners who were exposed to a unimodal distribution paired with compatible Roman orthography would learn to perceive short and long consonants as variants of the same category, while learners who were trained on a bimodal distribution along with compatible orthography would learn to perceive short and long consonants as two distinct categories. Conversely, incompatible orthography was expected to lead learners trained on a unimodal distribution to perceive the length contrast more accurately than learners trained with unimodal exposure only, whereas pairing a bimodal distribution with incompatible orthographic input was predicted to interfere with the perception of the length contrast compared to bimodal training only.

3.2.1 Methods

3.2.1.1 Participants

The participants were 80 monolingual English speakers, 78 of whom were recruited from the University of Alberta Department of Linguistics participant pool and received course credit. Two of the participants were recruited from the University of Alberta community and were paid $10 for taking part in the study. The mean age of the participants was 20.6 (*SD*= 4.2, range= 18-43), with 57 female and 23 male speakers. None of the participants reported a significant exposure to other languages beyond high school French.

3.2.1.2 Training materials

The auditory training stimuli used in this experiment were the same as in Experiment 3 in Chapter 2. Two length continua of 10 equal steps created from /m/ and /n/ were used to train participants on the consonant length contrast. The orthographic stimuli were presented using Roman script (more details are provided in the following section). The filler items of consonants that do not have an equivalent in English were romanized using the ALA-LC (American Library Association-

Library of Congress) romanization standards for Arabic. For instance, the pharyngealized alveolar stop /tˤ/ was written as <ṭ>. A list of these auditory stimuli and their written representations is provided in Table 3.1.

TABLE 3.1: A list of the nonnative auditory stimuli and their Romanized written representations.

Auditory Stimuli	Written Stimuli	Auditory Stimuli	Written Stimuli
/atˤa/	aṭa	/asˤa/	aṣa
/aðˤa/	aḍa	/aʁa/	agha
/aqa/	aqa	/aχa/	akha

3.2.1.3 Training phase

Half of the participants were assigned to two groups (20 participants per group) in which the orthographic input supported the distributional information. For the Unimodal Same Roman Spelling (USRS) group, the participants were exposed to a unimodal distribution of two length continua, in which tokens in the middle of the continuum were repeated four times more than any token in the distribution. Simultaneously, each token was paired with an orthographic representation where the target consonant was spelled with a single letter. For the Bimodal Distinct Roman Spelling group (BDRS), learners heard tokens around the endpoints of each continuum (2 and 9) more than any tokens in the distribution. The target consonant in tokens 1 to 5 was spelled with a single letter, while tokens 6 to 10 were shown with double letters (e.g., /ama/ with a consonant duration of 100 ms was spelled as <ama>, while consonant duration of 200 ms was spelled as <amma>). An illustration of the training design for these two compatible groups is given in Figure 3.1.

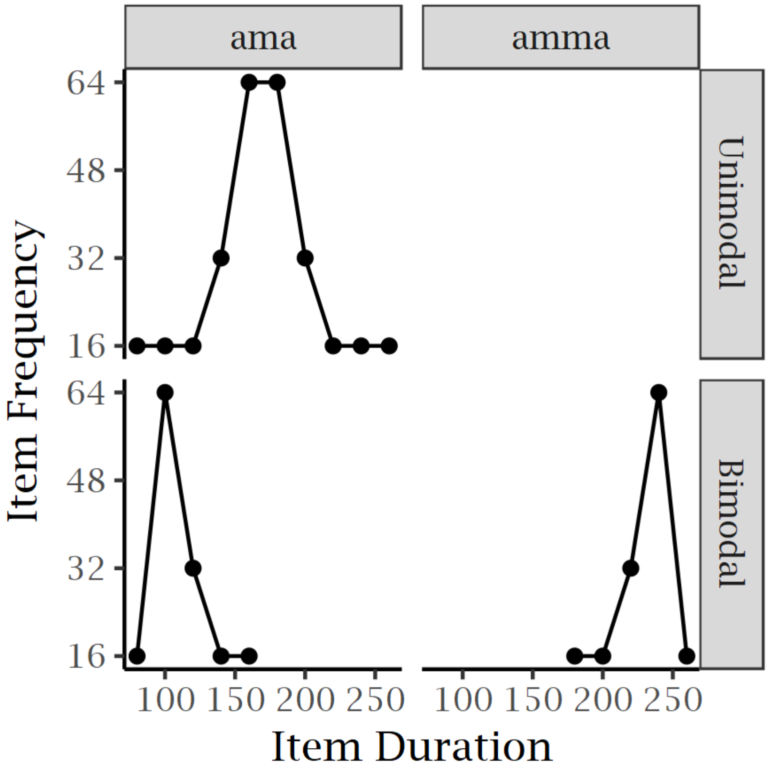


FIGURE 3.1: An illustration of the **compatible** orthographic pairing with the unimodal and bimodal distributions of the /ama/ duration continuum. All the auditory tokens in the unimodal distribution were presented with the same spellings <ama>; however, in the bimodal distribution, the first 5 tokens were spelled as <ama>, while the remaining tokens were spelled as <amma>.

Conversely, the other learners were assigned to two experimental groups in which the auditory and the orthographic stimuli conveyed conflicting information. For the Unimodal Distinct Roman Spelling group (UDRS), while learners were exposed to the unimodal distribution that favored learning a single category, the target consonants in tokens 1 to 5 were spelled with a single letter, whereas the consonants in tokens 6 to 10 were shown with double letters. Likewise, the auditory stimuli for the Bimodal Same Roman Spelling group (BSRS) were paired with incompatible orthographic labels. Learners in this group heard items 2 and 9 more frequently than any token, while simultaneously saw their written representations in which the target consonant spelled with a single letter (see Figure 3.2). Table 3.2 summarizes the auditory and orthographic stimuli provided to each group in this experiment.

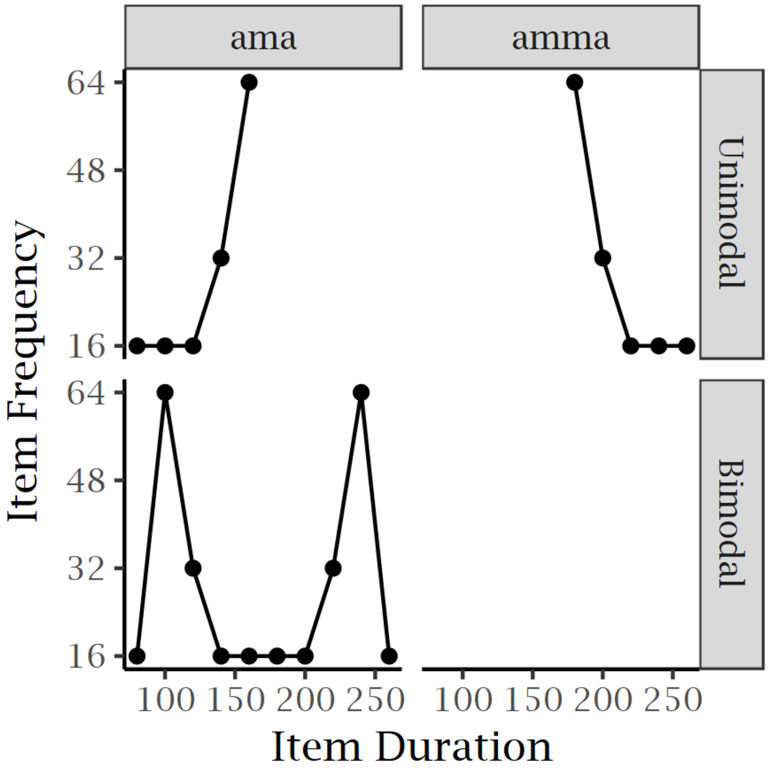


FIGURE 3.2: An illustration of the **incompatible** orthographic pairing with the unimodal and bimodal distributions of the /ama/ duration continuum. In the unimodal distribution, the first 5 tokens were spelled as <ama>, while the remaining tokens were spelled as <amma>. All the auditory tokens in the bimodal distribution were presented with the same spellings <ama>.

All groups were presented with the same filler items (see Section (2.4.1.2) of the previous chapter for more details) and their orthographic representations. Participants went through 4 blocks of training, with each block containing 72 target tokens and 28 fillers, bringing the total number of items to 400 (an example of a trial is shown in Figure 3.3). The training was conducted in a sound attenuated booth at the Alberta Phonetics Laboratory. The participants were seated in front of a computer screen and were fitted with over-the-ear headphones (MB QUART QP-805 HS). The materials were played in a random order and the training lasted for approximately 12 minutes. The training was administered using E-Prime 2.0 (Psychology Software Tools, 2012).

TABLE 3.2: A summary table of the learning groups examined in this experiment.

Group	Distribution	Orthographic stimuli
Control	None	None
Unimodal	Unimodal	None
Bimodal	Bimodal	None
USRS	Unimodal	Same Roman spelling (compatible)
BDRS	Bimodal	Distinct Roman spelling (compatible)
UDRS	Unimodal	Distinct Roman spelling (incompatible)
BSRS	Bimodal	Same Roman spelling (incompatible)

anna

FIGURE 3.3: An example of a trial presented to the participants during the learning phase. Participants heard an auditory token from the /ana/ length continuum and simultaneously saw its spelling.

3.2.1.4 Testing phase

This experiment replicated the testing procedures and materials described in Experiment Three from the previous chapter (see Section 2.4.1.4). Following the training phase, all participants performed three blocks of testing. In the first block, participants were tested on the items they heard during training, while in the second block the testing materials were extracted from the consonant length continua. In these blocks, participants listened to pairs of tokens and had to decide whether the items they heard were the same or different. The pairs consisted of different combinations of items 1 and 10: (1-10, 10-1) for different pairs or (1-1, 10-10) for identical pairs.

In the third block, participants listened to pairs of naturally produced singleton and geminate consonants. The purpose of these additional testing blocks was to examine whether learners were able to generalize what they learned to different nonnative contrasts that share the same length feature (see Table 2.1 for more details). Further, participants were tested on the fillers items. Each block contained 64 pairs, 32 target items and 32 fillers. The inter-stimulus interval was set at 1300 milliseconds in order to elicit phonological processing instead of a phonetic or auditory processing (Werker & Logan, 1985). A practice test was conducted prior to the training phase.

3.2.2 Results

The accuracy scores of the four learning groups on the length contrast are presented in Figure 3.4 (responses to filler items and Target Same pairs are not shown[1]). All four groups exhibited performances that were heavily influenced by the type of orthography that was paired with each distribution. For groups with compatible orthography, the mean proportion of correct responses for learners exposed to the unimodal distribution paired with the same Roman spelling (USRS) (*M*= 0.31, *SD*= 0.46) was lower than learners who had the bimodal distribution with distinct Roman spellings (BDRS) (*M*= 0.89, *SD*= 0.31). Conversely, the mean proportion of correct responses for learners who were exposed to the unimodal distribution paired with distinct Roman spellings (UDRS) was higher (*M*= 0.87, *SD*= 0.34) than those who heard the bimodal distribution that was paired with the same spelling (BSRS) (*M*= 0.42, *SD*= 0.49).

Interestingly, while the accuracy scores of the Bimodal (BDRS) and Unimodal (UDRS) Distinct Roman Spelling groups were notably similar, the discrimination scores for learners who received the same orthographic input, i.e. the spellings indicated no length distinction, displayed more variation, particularly those in the Bimodal Same Roman Spelling (BSRS).

[1]No significant differences were found between the four groups on these items.

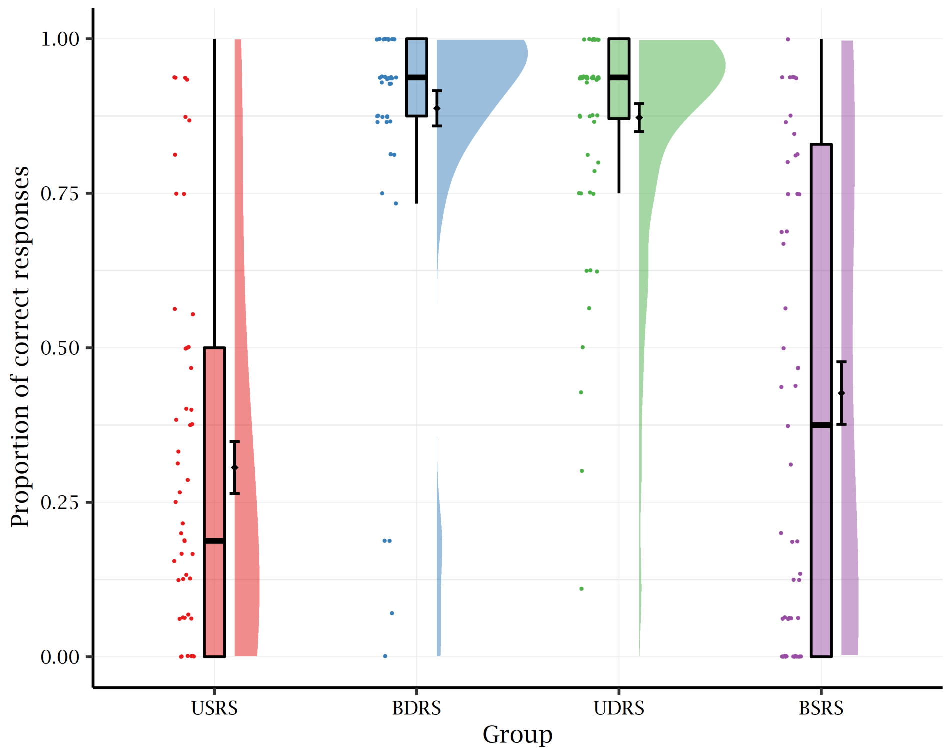


FIGURE 3.4: A combined plot representing the accuracy of the USRS (Unimodal Same Roman Spelling), BDRS (Bimodal Distinct Roman Spelling), UDRS (Unimodal Distinct Roman Spelling) and BSRS (Bimodal Same Roman Spelling) groups on Different pairs. The plot contains probability distributions (split-half violin); box-plots and means with standard errors for each group. Data from each individual are represented as points.

To assess the relationship between distributional learning and orthographic information in the perception of length contrast, multiple logistic regression models were fitted. The dependent variable in these models was the response to the AX discrimination task (Same vs. Different). The independent variables of interest were GROUP, a between-subject fixed effect, and TEST TYPE, a within-subject fixed effect. SUBJECT and ITEM served as random effects. By-subject random slopes were considered for TEST TYPE and always resulted in a better model fit. Adding by-item random slopes for GROUP did not improve the models. The result section is organized as follows. Section 3.2.2.1 analyzed data from the compatible orthography groups, where multiple logistic regression models were fitted to compare the accuracy scores of learners in the compatible orthography groups to those in the Control, Unimodal and Bimodal groups. Section 3.2.2.2 replicated the same analyses with data from the incompatible orthography groups. Finally, 3.2.2.3 presented two separate models that compared the accuracy scores of learners in the compatible and incompatible orthography groups. The justification for each model is provided below.

3.2.2.1 Compatible distributional and orthographic information

In this section, three models were fitted. The first model compared the accuracy scores of learners in the two compatible orthography groups (Unimodal Same Roman Spelling and Bimodal Distinct Roman Spelling) to those who received no training (Control); the second model compared the accuracy scores of learners who were exposed to the unimodal distribution paired with the same Roman spelling (USRS) to learners who received the unimodal training only (Unimodal); and finally, the third model compared the accuracy scores of learners who had the bimodal training with distinct Roman spelling (BDRS) to those who had the bimodal training only (Bimodal). A detailed explanation for each model is provided below.

The purpose of the first model was to examine whether the additional exposure to compatible Roman spellings had an effect on the perception of the length contrast in comparison to the Control group. The mixed-effects logistic regression model contained GROUP (Control vs. Unimodal Same Roman Spelling and Bimodal Distinct Roman Spelling), TEST TYPE (Trained vs. Untrained and Untrained Natural) and their interaction as fixed effects. Additionally, the model contained random intercepts for SUBJECT and ITEM, with by-subject random slopes for TEST TYPE. The model is summarized in Table 3.3.

As shown in the table, the results of the model reveal that learners who had the unimodal exposure along with compatible orthography (USRS) were significantly less likely to respond 'Different' to the Trained stimuli compared to the Control group. That is, the Control group had 7.44 times the odds of perceiving the length contrast as different (95% CI: 1.18-47) compared to the USRS group. No significant differences were found between the two groups on both Untrained and Untrained Natural pairs. Those who were exposed to the bimodal distribution paired

TABLE 3.3: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on target different items. The reference level for GROUP is 'Control', and the reference level for TEST TYPE is 'Trained'.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
Intercept	-0.388	0.626	-0.619	0.536
USRS	-2.033	0.889	-2.285	**0.022**
BDRS	5.076	0.958	5.297	**0.000**
Untrained	0.032	0.681	0.047	0.963
Untrained Natural	0.647	0.683	0.947	0.344
USRS:Untrained	0.769	0.747	1.030	0.303
BDRS:Untrained	-1.070	0.867	-1.235	0.217
USRS:Untrained Natural	1.003	0.753	1.332	0.183
BDRS:Untrained Natural	-2.682	0.849	-3.158	**0.002**

with compatible orthography (BDRS), on the contrary, were significantly more likely to respond 'Different' to the Trained pairs. In comparison to the Control group, training learners with the bimodal distribution alongside distinct orthographic labels significantly increased the odds of perceiving the length contrast as different (OR=160.2, 95% CI: 24.48-1048). However, there was a significant interaction between GROUP and TEST TYPE. The degree by which the BDRS group performed on the Trained pairs compared to the Control group significantly decreased when perceiving the Untrained Natural stimuli. In that case, the odds ratio dropped to 10.97 but remained statistically significant (95% CI: 2.87-41.96). Figure 3.5 displays the predicted probabilities for the model described above.

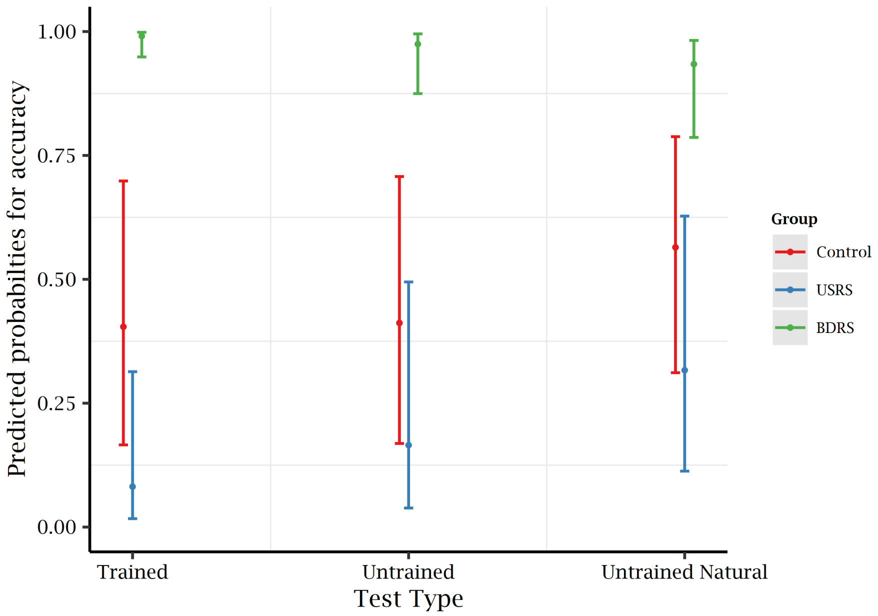


FIGURE 3.5: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.3 . The probabilities are presented with 95% confidence interval.

The results so far confirm that exposing native English listeners to auditory distributions alongside compatible orthographic labels significantly influences their perception of the length contrast compared to those without prior exposure. Still, it is unclear whether combining orthographic labels with auditory items had any advantage over distributional exposure only. Thus, a second mixed-effects logistic regression model investigated data from learners who were trained with a unimodal distribution with or without compatible orthographic input. The model predictors were GROUP (Unimodal vs. Unimodal Same Roman Spelling), TEST TYPE (Trained vs. Untrained and Untrained Natural) and their interaction. The random predictors were SUBJECT and ITEM, with by-subject random slopes for TEST TYPE. A summary of the model is provided in Table 3.4.

The results indicate that training learners on the unimodal distribution alongside compatible orthographic information significantly lowered the likelihood of responding 'Different' to the Trained stimuli compared to training learners with a unimodal distribution only. The odds of discriminating the length contrast for the

TABLE 3.4: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy for participants trained on the unimodal distribution with or without orthography. The reference level for GROUP is 'Unimodal', and the reference level for TEST TYPE is 'Trained'.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	2.247	0.943	2.382	0.017
USRS	-4.765	1.255	-3.796	**0.000**
Untrained	-1.069	0.791	-1.351	0.177
Untrained Natural	-0.941	0.785	-1.199	0.230
USRS:Untrained	2.043	0.905	2.256	**0.024**
USRS:Untrained Natural	2.765	0.889	3.109	**0.002**

Unimodal group were 117.37 times higher than the USRS group (95% CI: 10.02-1373). However, the model reveals a significant interaction between GROUP and TEST TYPE. In comparison to their performances on the Trained stimuli, the differences in accuracy between the two groups were significantly reduced when perceiving the Untrained and the Untrained Natural stimuli. The odds ratio decreased to 15.22 (95% CI: 2.28-101.62) for the Untrained pairs, and to 7.39 (95% CI: 1.99-27.4) for the Untrained Natural stimuli, while remaining statistically significant in both cases. Figure 3.6 shows the predicted probabilities of the model.

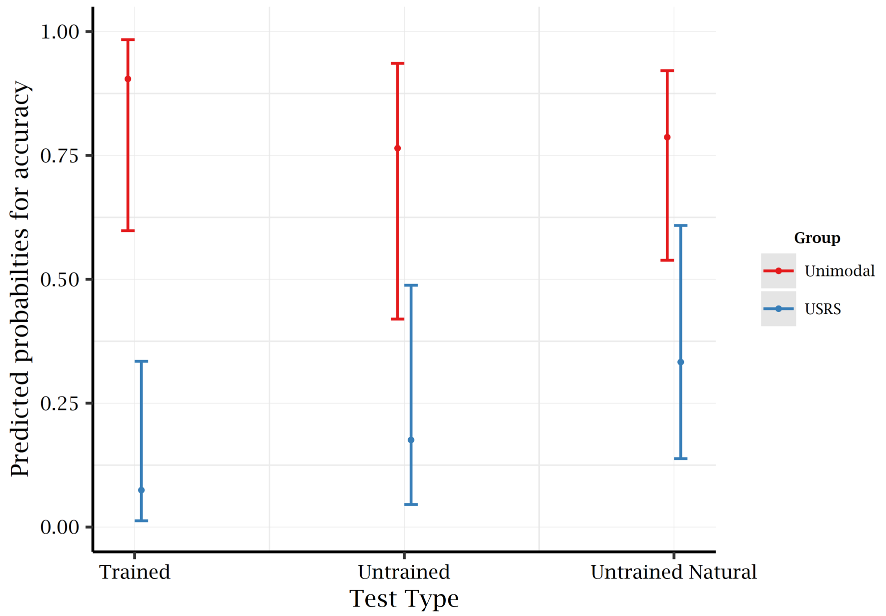


FIGURE 3.6: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.4. The probabilities are presented with 95% confidence interval.

The third model investigated whether the exposure to the bimodal distribution paired with compatible Roman orthography improved the discrimination of the length contrast compared to training learners with the bimodal distribution only. The model included response data from two groups of learners who either had the bimodal training only or the bimodal training with distinct Roman spelling. The fixed predictors were GROUP (Bimodal vs. Bimodal Distinct Roman Spelling), TEST TYPE (Trained vs. Untrained and Untrained Natural) and their interaction. The random predictors were SUBJECT and ITEM, with by-subject random slopes for TEST TYPE. Table 3.5 summarizes the model.

The results reveal that training learners with the bimodal distribution paired with compatible orthography significantly increased the likelihood of responding 'Different' to the Trained stimuli compared to exposing learners to the bimodal distribution only. The odds of discriminating the length contrast for the Bimodal Distinct Roman Spelling (BDRS) group was 8 times greater than the Bimodal group (95% CI: 1.67-38.43). No significant differences were found between the two groups

TABLE 3.5: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy for participants trained on a bimodal distribution with or without orthography. The reference level for GROUP is 'Bimodal', and the reference level for TEST TYPE is 'Trained'.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.237	0.619	3.612	0.000
BDRS	2.080	0.800	2.599	**0.009**
Untrained	-0.121	0.739	-0.164	0.870
Untrained Natural	-0.554	0.739	-0.750	0.453
BDRS:Untrained	-1.004	0.876	-1.146	0.252
BDRS:Untrained Natural	-1.234	0.875	-1.411	0.158

on Untrained and Untrained Natural pairs. The predicted probabilities for the both groups on the three testing blocks are shown in Figure 3.7.

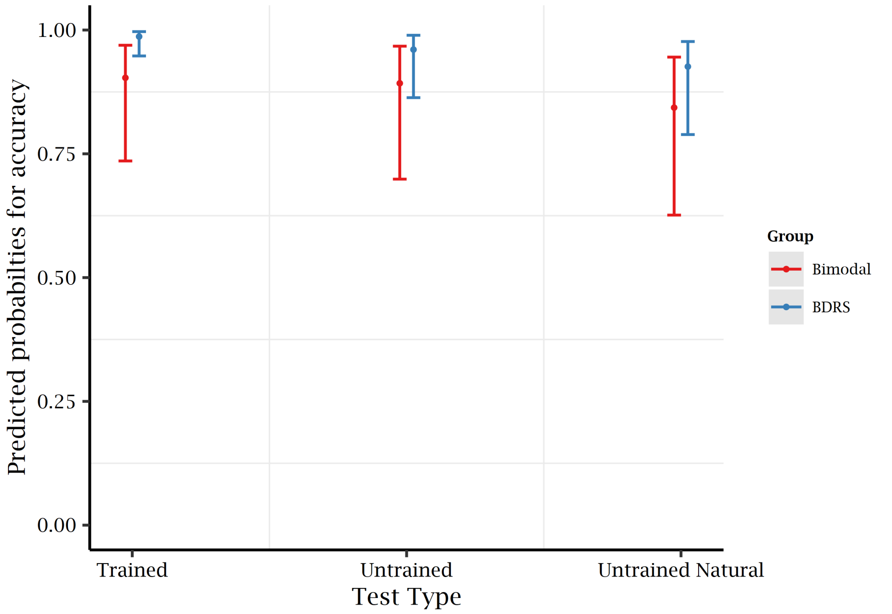


FIGURE 3.7: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.5. The probabilities are presented with 95% confidence interval.

3.2.2.2 Incompatible distributional and orthographic information

The previous section demonstrates that the exposure to compatible distributional and orthographic information could support a single length category perception when a single orthographic representation is paired with a unimodal distribution. Likewise, when two distinct orthographic representations are paired with a bimodal distribution, a two category length distinction could be inferred. The present section examines if the learning patterns differ when the orthographic input does not correspond with the distributional information. To this end, three mixed-effects logistic regression models were conducted. These analyses explored whether the conflicting orthographic information would alter learners' perception of the length contrast. The first model was fitted to compare the accuracy scores of the Unimodal Distinct Roman Spelling and Bimodal Same Roman Spelling groups to that of the Control group. The fixed predictors were GROUP, TEST TYPE and their interaction. SUBJECT and ITEM served as random effects, with by-subject random slopes for TEST TYPE. Table 3.6 summarizes the model.

TABLE 3.6: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy for participants trained on either the unimodal or bimodal distribution paired with mismatching orthography. The reference level for GROUP is 'Control', and the reference level for TEST TYPE is 'Trained'.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-0.364	0.568	-0.642	0.521
UDRS	3.590	0.826	4.345	**0.000**
BSRS	-0.544	0.820	-0.663	0.507
Untrained	0.045	0.543	0.082	0.934
Untrained Natural	0.617	0.553	1.117	0.264
UDRS:Untrained	0.302	0.581	0.520	0.603
BSRS:Untrained	-0.268	0.564	-0.475	0.635
UDRS:Untrained Natural	-1.524	0.592	-2.575	**0.010**
BSRS:Untrained Natural	-0.009	0.579	-0.016	0.987

The results of the model indicate that learners in the Unimodal Distinct Roman Spelling (UDRS) group were significantly more likely to respond 'Different' to the Trained stimuli than the Control group. The odds of perceiving the length contrast as different for learners trained on the unimodal distribution alongside distinct orthographic labels was 36.23 times greater than the Control group (95% CI: 7.27-182.97). The same pattern was observed when comparing the performance of the two groups on Untrained stimuli. However, as evidenced by the the significant interaction, the difference in accuracy between the Control and UDRS groups was significantly lower on the Untrained Natural stimuli compared to Trained pairs. The odds ratio of discriminating the Untrained Natural pairs dropped to 7.89 (95% CI: 2.33-26.73), but still remained statistically significant. For learners trained with the bimodal distribution alongside incompatible orthographic information (BSRS), no significant difference was evident in their accuracy scores compared to the Control group. The predicted probabilities obtained from the model are shown in Figure 3.8.

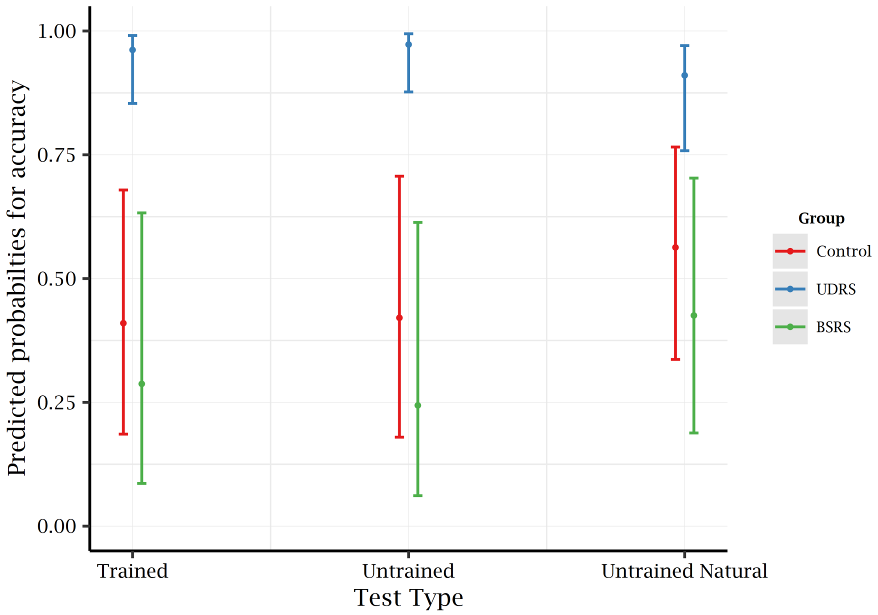


FIGURE 3.8: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.6. The probabilities are presented with 95% confidence interval.

The second model compared the performance of learners who had the unimodal training only to those who were trained on the unimodal distribution with incompatible orthographic exposure. The purpose of this analysis was to examine whether combining the unimodal training with two distinct orthographic labels could significantly improve the perception of the length contrast in comparison to the unimodal training only. The fixed predictors were GROUP (Unimodal vs. Unimodal Distinct Roman Spelling) and TEST TYPE (Trained vs. Untrained and Untrained Natural). The random effects were SUBJECT and ITEM, with by-subject slopes for TEST TYPE. A summary of the model is provided in 3.7.

TABLE 3.7: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy for participants trained on either the unimodal distribution only or the unimodal distribution with incompatible orthography. The reference level for GROUP is 'Unimodal', and the reference level for TEST TYPE is 'Trained'.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	1.710	0.617	2.772	0.006
UDRS	1.431	0.794	1.803	0.071
Untrained	-0.587	0.529	-1.110	0.267
Untrained Natural	-0.501	0.502	-0.997	0.319
UDRS:Untrained	0.733	0.605	1.213	0.225
UDRS:Untrained Natural	-0.555	0.530	-1.046	0.295

The model reveals that both groups had similar accuracy scores on Trained and Untrained Natural items. However, when setting the reference level for TEST TYPE to Untrained, the difference between the two groups was statistically significant (β= 2.16, *SE*= 0.83, *p*=0.009). The odds of discriminating the length contrast for the UDRS group was 8.71 times that of the Unimodal group (95% CI: 1.72-44.18). Figure 3.9 illustrates this difference in performance between the two groups on the there testing types.

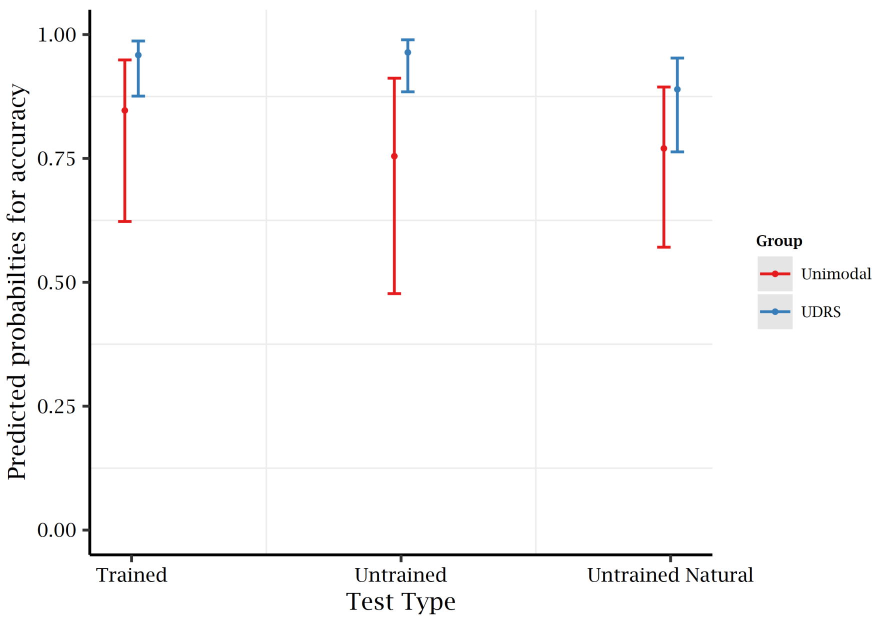


FIGURE 3.9: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.7. The probabilities are presented with 95% confidence interval.

A similar analysis was conducted to examine whether learners who were exposed to the bimodal distribution paired with the same spelling for each token (incompatible) perceived the length contrast less successfully than those who were exposed to a bimodal distribution only. The fixed predictors of the model were GROUP (Bimodal vs. Bimodal Same Roman Spelling) and TEST TYPE (Trained vs. Untrained and Untrained Natural). The random effects were SUBJECT and ITEM, with by-subject slopes for TEST TYPE. A summary of the model is provided in 3.8.

TABLE 3.8: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy for participants trained on either the bimodal distribution only or the bimodal distribution with incompatible orthography. The reference level for GROUP is 'Bimodal', and the reference level for TEST TYPE is 'Trained'.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	2.418	0.810	2.985	0.003
BSRS	-3.355	1.040	-3.227	**0.001**
Untrained	-0.077	0.663	-0.117	0.907
Untrained Natural	-0.844	0.680	-1.241	0.214
BSRS:Untrained	-0.107	0.692	-0.155	0.877
BSRS:Untrained Natural	1.504	0.703	2.140	**0.032**

The result of the analysis indicates that the Bimodal Same Roman Spelling group (BSRS) had significantly lower accuracy scores on Trained pairs compared to the Bimodal group. The odds of perceiving the length contrast as different for the Bimodal group was 28.65 times greater than the BSRS group (95% CI: 3.73-219.79). The model reveals, however, a significant interaction between GROUP and TEST TYPE. On Untrained Natural stimuli, the difference between the two groups significantly decreased compared to Trained pairs. The odds ratio dropped to 6.36 (95% CI: 1.56-25.89), while remaining statistically significant. The predicted probabilities obtained from the model are shown in Figure 3.10.

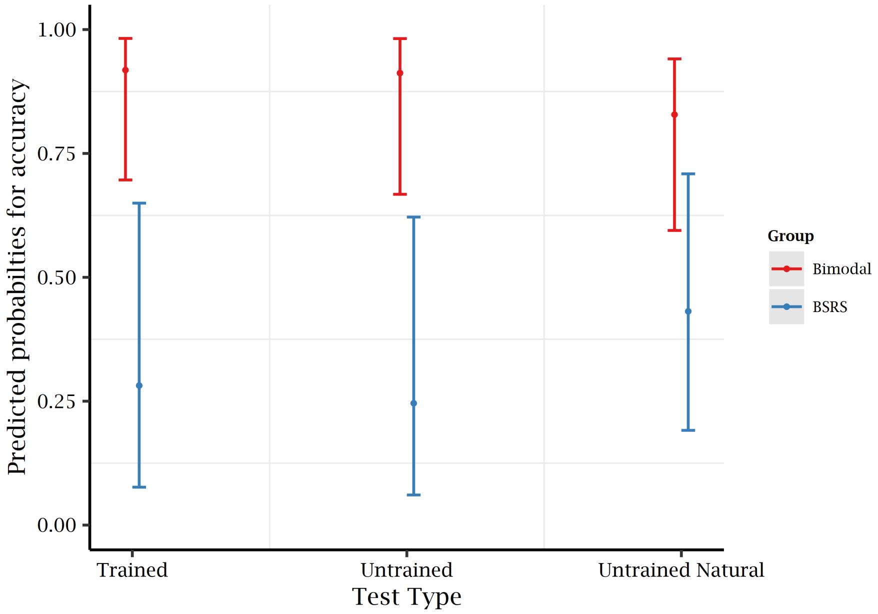


FIGURE 3.10: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.8. The probabilities are presented with 95% confidence interval.

In summary, the results so far demonstrate that exposing learners to either compatible or incompatible orthographic information influences their perception of the length contrast. Both the Bimodal Distinct and Unimodal Distinct orthography groups performed significantly better on the AX discrimination task than the Control, Unimodal and Bimodal groups. In the same way, the Unimodal Same and Bimodal Same Roman spelling groups performed significantly worse than the Unimodal and Bimodal groups and similar to the Control group. The following section explores whether the distributional information could still contribute to the learning outcome in the presence of orthographic representations.

3.2.2.3 Compatible versus incompatible distributional and orthographic information

Thus far, there is clear evidence that orthographic information — regardless of its type — had an effect on the perceptual learning of the consonant length contrast.

The purpose of the following analysis was to examine whether the type of distribution could exert any influence on the discrimination scores. Two mixed-effects logistic regression models were conducted. The first model compared the accuracy scores of learners who either had the unimodal distribution paired with compatible orthography or a bimodal distribution that was paired with incompatible orthographic labels. Both groups were exposed to the same orthographic stimuli; all tokens on each continuum had the same spelling, but the auditory distribution was different (unimodal vs. bimodal). The model fixed effects were GROUP (Unimodal Same Roman Spelling vs. Bimodal Same Roman Spelling), TEST TYPE (Trained vs. Untrained and Untrained Natural) and their interaction. The random effects were SUBJECT and ITEM, with by-subject slopes for TEST TYPE. Table 3.9 summarizes the results of the model.

TABLE 3.9: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy for participants trained on either a unimodal distribution with compatible or a bimodal distribution with incompatible orthography. The reference level for GROUP is 'USRS' (Unimodal Same Roman Spelling), and the reference level for TEST TYPE is 'Trained'.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-2.548	0.943	-2.704	0.007
BSRS	1.489	1.210	1.231	0.218
Untrained	0.828	0.809	1.024	0.306
Untrained Natural	1.864	0.793	2.351	**0.019**
BSRS:Untrained	-1.063	0.890	-1.195	0.232
BSRS:Untrained Natural	-1.068	0.854	-1.250	0.211

The model showed that although learners in the Bimodal Same Roman Spelling group (BSRS) were more likely to perceive the length contrast as different than the Unimodal Same Roman Spelling group (USRS), the difference was not significant. The only significant difference revealed by the model was the performance of the USRS group on Untrained Natural pairs compared to Trained stimuli. The odds of accuracy on the Untrained Natural pairs were 6.45 times greater than the Trained pairs (95% CI: 1.36-30.51). The predicted probabilities of the model are visualized in Figure 3.11.

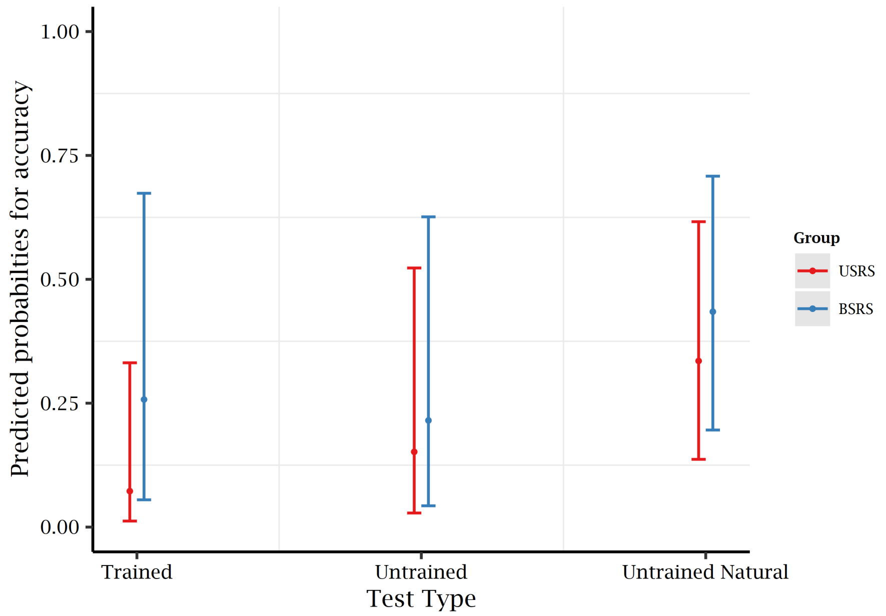


FIGURE 3.11: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.9. The probabilities are presented with 95% confidence interval.

Similarly, the second model compared the response accuracy for learners who either had the bimodal training with compatible orthography or a unimodal training alongside incompatible orthographic input. Here, the two groups had the same orthographic exposure (two distinct orthographic labels for the auditory tokens on each continuum) but received two different auditory distributions (unimodal vs. bimodal). The model fixed effects were GROUP (Bimodal Distinct Roman Spelling vs. Unimodal Distinct Roman Spelling), TEST TYPE (Trained vs. Untrained and Untrained Natural) and their interaction. The random effects were SUBJECT and ITEM, with by-subject slopes for TEST TYPE. A summary of the model is provided in Table 3.10.

According to the model, training learners with the bimodal distribution paired with distinct spelling significantly increased the likelihood of responding 'Different' to the Trained stimuli compared to training them with the unimodal distribution paired with incompatible orthography. The odds of discriminating the contrast for the Bimodal Distinct Roman Spelling group (BDRS) was 2.60 times greater than the

TABLE 3.10: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy for participants trained on either the unimodal distribution with compatible or the bimodal distribution with incompatible orthography. The reference level for GROUP is 'UDRS' (Unimodal Distinct Roman Spelling), and the reference level for TEST TYPE is 'Trained'.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	2.662	0.359	7.425	0.000
BDRS	0.955	0.485	1.970	**0.049**
Untrained	0.561	0.614	0.914	0.361
Untrained Natural	-0.543	0.540	-1.006	0.314
BDRS:Untrained	-1.078	0.804	-1.341	0.180
BDRS:Untrained Natural	-0.753	0.723	-1.041	0.298

Unimodal Distinct Roman Spelling group (UDRS) (CI 95%: 1.01-6.72). For the Untrained and Untrained Natural stimuli, no significant differences were found. Figure 3.12 shows the predicted probabilities obtained from the model.

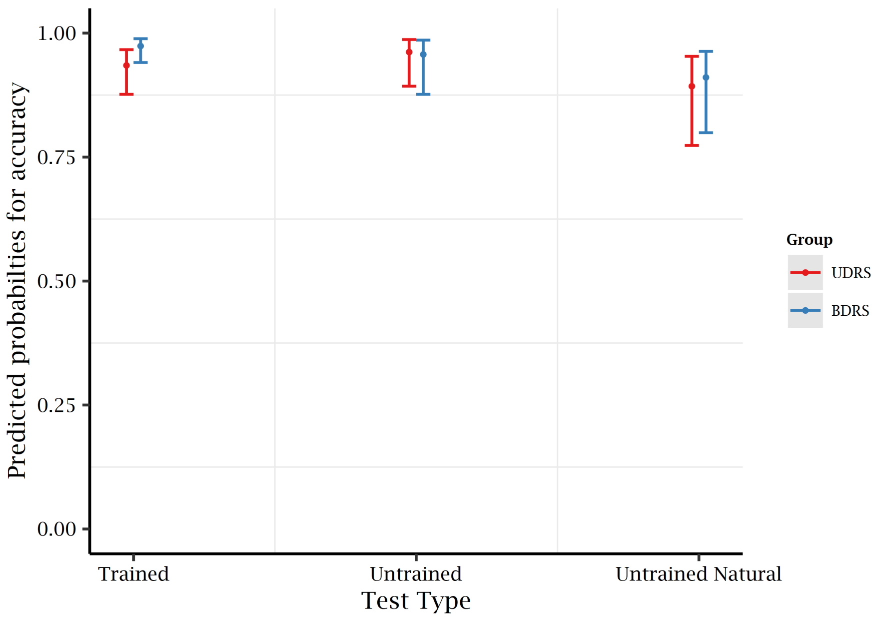


FIGURE 3.12: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.10. The probabilities are presented with 95% confidence interval.

In sum, the results of the two models revealed that in the presence of orthographic stimuli, distributional information seems to only influence the accuracy of the Bimodal Distinct Roman Spelling group on the Trained pairs.

3.2.3 Discussion

The present experiment sought to explore the nature of the interaction between auditory and familiar written stimuli in the perceptual learning of consonant length by native English speakers. Specifically, compatible and incompatible orthographic information were paired with unimodal and bimodal distributions of two length continua to examine whether orthography could facilitate or offset the distributional learning of the length contrast.

Firstly, it was predicted that exposing learners to a unimodal auditory distribution alongside compatible orthographic labels, i.e. tokens in each continuum had the same spelling, would greatly reduce learners' ability to perceive the length contrast. The findings of this experiment confirmed this prediction. Learners who had the unimodal training paired with the same written representations demonstrated significantly lower accuracy in discriminating the length contrast than both learners who had no training and those who had the unimodal exposure only, thereby successfully inferring one length category. In the same way, it was predicted that learners who were presented with a bimodal distribution alongside two distinct orthographic representations for short and long consonants would exhibit greater perceptual sensitivity to the contrast. The findings showed that learners with such exposure outperformed both the Control and the Bimodal groups. In fact, integrating compatible orthographic input with bimodally distributed length continua resulted in a near perfect discrimination of the consonant length contrast.

Secondly, it was predicted that contradictory orthographic and auditory input would increase the sensitivity to the length contrast for learners who had the unimodal training and decrease it for learners trained on the bimodal distribution. The findings suggest that presenting learners with the unimodal distribution represented by two distinct orthographic labels significantly enhanced the perception of the contrast compared to both the Control the Unimodal groups. In addition, exposing learners with the bimodal distribution represented orthographically by a single label neutralized the effects of the auditory distribution as evident by the significantly lower perception of learners who received such training compared to those who had the bimodal exposure only. However, the effect of the bimodal distribution was not completely offset by the incompatible orthographic labels. This was evident by the lack of significant difference in accuracy between learners with such exposure and the Control group. The only group of learners who exhibited significantly less accuracy in discriminating the length contrast compared to the Control group were those who received the unimodal training paired with the same orthographic labels for short and long consonants.

Given the evidence that incompatible orthography significantly alters the perception of the length contrast, a follow-up analysis compared the performances of compatible and incompatible orthography groups. The findings suggest that, for the most part, written input determines the learning outcome regardless of the auditory distribution it represents. The only exception was the performance of learners in the Bimodal Distinct Roman Spelling group who significantly outperformed the Unimodal Distinct Roman Spelling group on Trained items. In this instance, the bimodal distribution appears to aide learners' perception of the length contrast better than the same orthographic input with the unimodal distribution.

Taken together, these findings suggest that access to a familiar orthographic input can play a role in forming L2 categories. Not only do the written representations enhance the distributional learning of such categories, but also override its effects. One might speculate that the robustness of the orthographic cues in guiding learners' perception of the length contrast could be attributed to the status of these representations in learners' L1. Based on the assumption that orthographic information influences the mental representations of words, Brewer (2008) demonstrated that native English speakers indeed produce segments that are represented by multiple letter with significantly longer duration. Thus, it is plausible that the already established association between the number of Roman letters representing a particular phoneme and its durational properties shaped learners' perception of the length contrast in this experiment. One way to explore this possibility is to expose native English speakers to an unfamiliar orthographic system. The following experiment replicated the present experiment with Arabic writing system in place of the familiar Roman alphabet.

3.3 Experiment Two

The second experiment explored whether exposing learners to unfamiliar orthography could have the same impact that was observed in the previous experiment. As previously stated in Chapter 1, several studies have indicated that while learners can generally benefit from familiar written representations of L2 segments, provided that L1 and L2 share the same grapheme-phoneme correspondences (e.g., Escudero et al., 2008; Hayes-Harb et al., 2018; Showalter, 2018), the exposure to novel written input has been shown to inhibit learning. Mathieu (2016) trained native English speakers on six minimal pairs containing the Arabic /ħ/-/χ/ contrast in word-initial positions. Among the experimental groups, two sets of learners heard the words and saw their pictured meanings. Crucially, one group was additionally presented with the Arabic written representations of these words. The auditory-picture matching test revealed that those who were exposed to the Arabic spellings of the minimal pairs were significantly worse at matching the auditory forms with their corresponding meanings than those who lacked orthographic exposure. However, a number of empirical studies suggest that novel orthographic representations exert no adverse or beneficial influence in learning the nonnative phonological forms of new words. In a word learning experiment, Showalter and Hayes-Harb (2015) examined native English speakers' ability to lexically encode the Arabic /k/-/q/ contrast. Two groups were exposed to the auditory forms of six minimal pairs containing the velar-uvular contrast along with their pictured meanings, with one group having to concurrently see the Arabic spellings of these words. The auditory-picture matching test failed to show any significant benefit for the Arabic script in improving the acquisition of the novel phonological contrast.

Relevant to the current experiment, the failure of unfamiliar written stimuli to induce learning has been also attested in learning to perceive nonnative sounds. In Pytlyk (2011), three groups of native English speakers received 4.5 hours of training on L2 Mandarin phonemes. The three groups differed on whether they were simultaneously exposed to Pinyin, Zhuyin or no orthography during training. The Pinyin writing system is familiar to native English speakers since it uses the Roman alphabet, whereas Zhuyin employs Chinese characters. The follow-up perception task revealed no significant differences between the three groups despite the potential negative effects of the incongruency between the grapheme-phoneme correspondences of the Pinyin script and that of learners' L1.

In the present experiment, native English speakers were trained on the consonant length contrast via distributional information and Arabic orthographic input. In line with the previous experiment, learners were assigned to four experimental Groups: Unimodal Same Arabic Spelling (USAS), Bimodal Distinct Arabic Spelling (BDAS), Unimodal Distinct Arabic Spelling (UDAS) and Bimodal Same Arabic Spelling (BSAS). Following the findings of the previous experiment and recent research, it was expected that the exposure to the Arabic script would not have the same profound effect in either facilitating or hindering the perception of the length contrast. Specifically, while a unimodal exposure paired with compatible Arabic input would decrease learners' ability to perceive the length contrast compared to the Unimodal group, their performance was predicted not to differ significantly from the Control group. Additionally, pairing a bimodal distribution with compatible Arabic input was expected to lead to a significantly better discrimination of the contrast compared to the Control group; however, access to unfamiliar orthography was not predicted to cause a significant increase in discrimination compared to the Bimodal group. The same pattern was expected to emerge in the incompatible orthographic groups. The presence of the unfamiliar Arabic input would have less impact in determining the learning outcome for both the Unimodal Distinct and

Bimodal Same Arabic Spelling groups.

3.3.1 Methods

3.3.1.1 Participants

Eighty monolingual English speakers (57 females and 23 males) were recruited from the University of Alberta Department of Linguistics participant pool and received course credit for their participation. The mean age was 20.4 (*SD*= 2.3, range= 18-27). None of the participants had any significant experience with languages that have gemination and reported no prior knowledge of the Arabic script.

3.3.1.2 Training materials

The same auditory stimuli was used to train participants in this experiment. The orthographic stimuli consisted of the Arabic spellings of the auditory tokens. For the target items, the difference between the long and short consonant was marked by the addition of a diacritic referred to as *shaddah*. In Arabic, the *shaddah*, which is placed above the consonant (e.g., <تّ> versus <ت>), is optionally used to mark long consonants. The filler items were also represented orthographically according to the Arabic spelling standards.

3.3.1.3 Training phase

The training phase was similar to one described in the previous experiment. Learners were assigned to four learning groups: Unimodal Same Arabic Spelling (USAS), Bimodal Distinct Arabic Spelling (BDAS), Unimodal Distinct Arabic Spelling (UDAS) and Bimodal Same Arabic Spelling (BSAS). Learners in the USAS group were trained on the unimodal distribution with compatible Arabic orthography. The auditory tokens across each length continuum were represented orthographically with the same spellings. Similarly, learners in the BDAS were presented with the bimodal distribution paired with compatible Arabic orthography; tokens 1-5 on each continuum were spelled without a diacritic indicating the token contained a short consonant (e.g., /ama/ spelled as <اما>), while tokens 6-10 were marked with a diacritic indicating the token contained a geminate consonant (e.g., /amma/ spelled as <امّا>). An illustration of the training is shown in Figure 3.13.

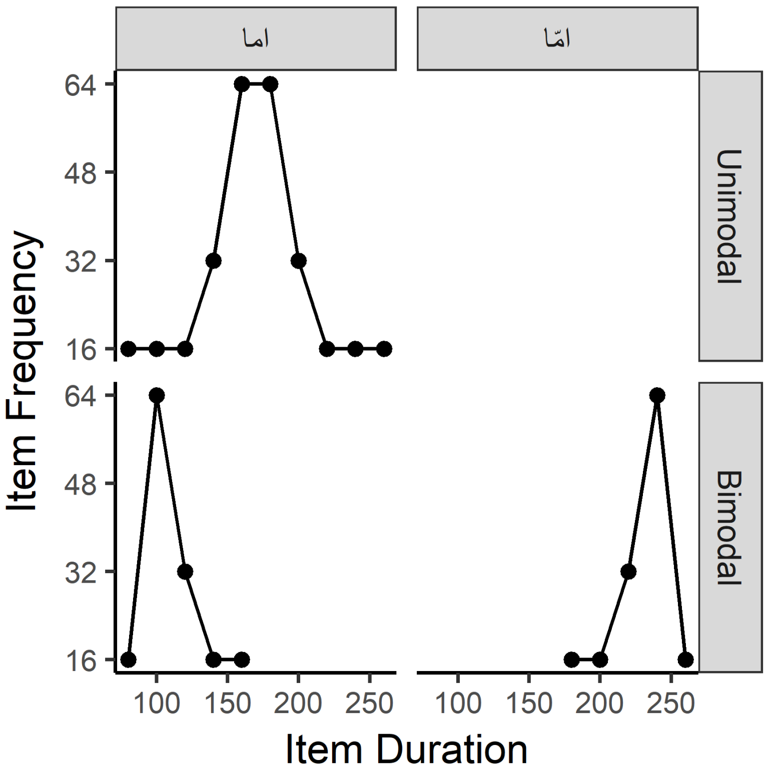


FIGURE 3.13: An illustration of the **compatible** Arabic orthographic pairing with the unimodal and bimodal distributions of the /ama/ duration continuum.

Conversely, learners in the UDAS group were exposed to the unimodal distribution with incompatible Arabic orthography. Tokens 1-5 in each continuum were spelled without a diacritic, while tokens 6-10 were spelled with a diacritic. Likewise, learners in the BSAS group were trained on the bimodal distrbtion with incompatible Arabic orthographic input; all tokens on each length continuum had the same spellings. Figure 3.14 illustrates the auditory and the written exposure for the incompatible orthography groups. Table 3.11 summarizes the auditory and the orthographic exposure for each group.

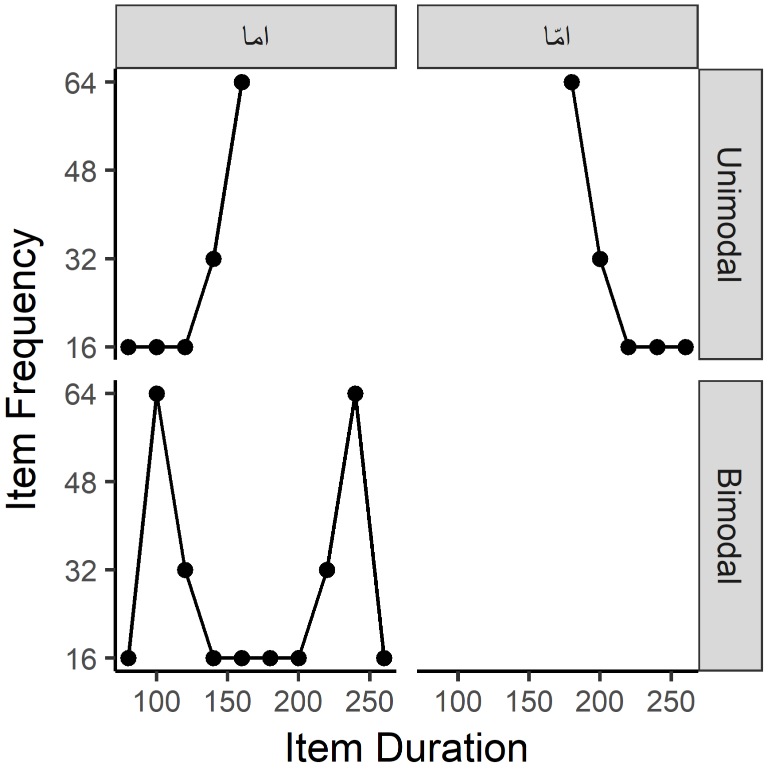


FIGURE 3.14: An illustration of the **incompatible** Arabic orthographic pairing with the unimodal and bimodal distributions of the /ama/ duration continuum.

Prior to the training, learners were told they would hear multiple words and see their spellings in an unfamiliar language. The participants were instructed about direction of writing in Arabic which is right-to-left.

TABLE 3.11: A summary table of the learning groups examined in this experiment. The Control, Unimodal and Bimodal groups are the same ones used in the previous experiment.

Group	Distribution	Orthographic stimuli
Control	None	None
Unimodal	Unimodal distribution	None
Bimodal	Bimodal distribution	None
USAS	Unimodal distribution	Same Arabic spelling (compatible)
BDAS	Bimodal distribution	Distinct Arabic spelling (compatible)
UDAS	Unimodal distribution	Distinct Arabic spelling (incompatible)
BSAS	Bimodal distribution	Uniform Arabic spelling (incompatible)

3.3.1.4 Testing phase

All groups performed the same AX discrimination task described in Section 3.2.1.4.

3.3.2 Results

The accuracy scores of the four groups on the length contrast (Target Different pairs)[2] are shown in Figure 3.15. Clearly, the performances of these groups exhibited variation. Consistent with the previous experiment, the type of orthography that was presented with each auditory distribution seems to determine the degree of accuracy in discriminating the length contrast. Both the Bimodal Distinct Arabic Spelling (BDAS) and Unimodal Distinct Arabic Spelling (UDAS) groups had comparable accuracy means (*M*= 0.82, *SD*= 0.38) and (*M*= 0.81, *SD*= 0.39), respectively. Likewise, the mean accuracy scores for the Unimodal Same Arabic Spelling (USAS) and Bimodal Same Arabic Spelling (BSAS) groups were almost identical (*M*= 0.33, *SD*= 0.46) and (*M*= 0.33, *SD*= 0.39), respectively.

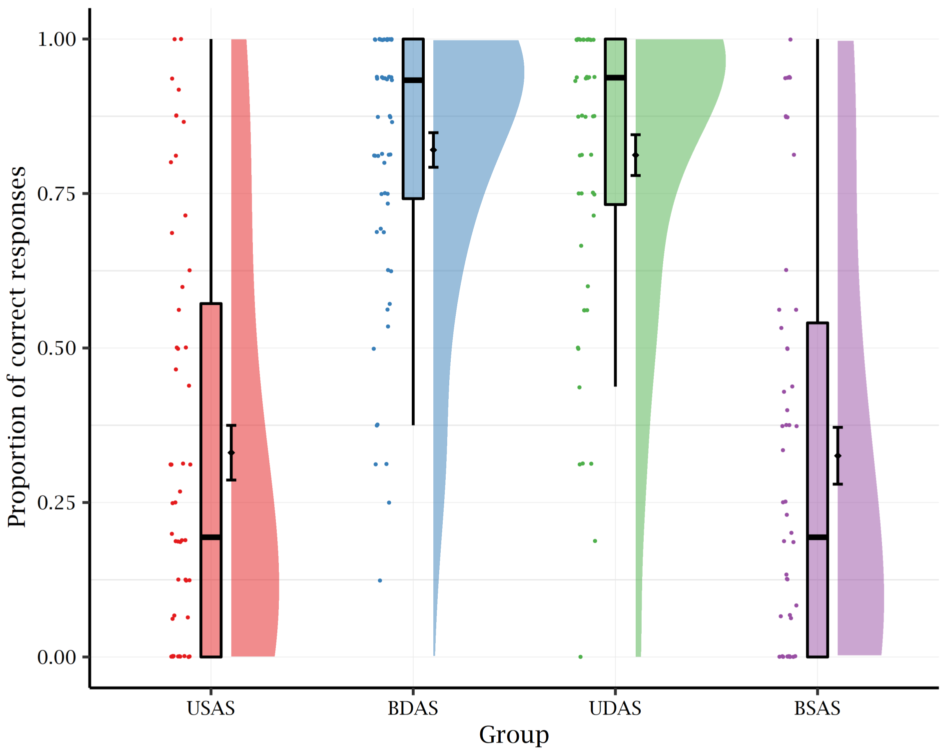


FIGURE 3.15: A combined plot representing the accuracy of the USAS (Unimodal Same Arabic Spelling), BDAS (Bimodal Distinct Arabic Spelling), UDAS (Unimodal Distinct Arabic Spelling) and BSAS (Bimodal Same Arabic Spelling) groups on Different pairs. The plot contains probability distribution (split-half violin); box-plots and means with standard errors for each group. Data from each individual are represented as points.

[2]Analysis on Target Same pairs and filler items did not reveal any significant differences between the four groups.

Responses to the AX discrimination task were analyzed using mixed-effects logistic regression to assess whether the perception of the length contrast was associated with the exposure to Arabic orthographic input. Following the previous experiment, the fixed effects of interest were GROUP, a between-subject factor effect, and TEST TYPE, a within-subject factor. SUBJECT and ITEM served as random effects. By-subject random slopes were included for TEST TYPE and always resulted in a better model fit. Adding by-item random slopes for GROUP failed to improve these models. The result section is divided into three sub-sections. Section 3.3.2.1 presents an analysis of data obtained from the compatible orthography groups in which their performances were compared to groups that received no orthographic cues (Control, Unimodal and Bimodal). Section 3.3.2.2 presents the same comparisons but with the incompatible orthography groups. The third section (3.3.2.3) compared the performances of the compatible orthography groups with those learners who received either the unimodal or bimodal training with incompatible orthography to explore whether distributional information had any effect on the accuracy scores. Finally, section 3.3.2.4 presents a combined analysis of data from all the experimental groups in this chapter.

3.3.2.1 Compatible distributional and orthographic information

The first mixed-effects logistic regression model was conducted to examine whether learners trained with compatible auditory and Arabic orthographic input performed differently from those who received no exposure,i.e., the Control group. The fixed-effects of the model were GROUP (Control vs. Unimodal Same Arabic Spelling and Bimodal Distinct Arabic Spelling), TEST TYPE (Trained vs. Untrained and Untrained Natural) and their interaction term. The random effects were SUBJECT and ITEM, with by-subject random slopes for TEST TYPE. Table 3.12 summarizes the model.

The result of the model indicates that presenting learners with the unimodal distribution paired with matching Arabic orthography significantly decreased the likelihood of responding 'Different' to the Trained pairs compared to the Control group.

TABLE 3.12: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on target different items. The reference level for GROUP is Control, and the reference level for TEST TYPE is Trained.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-0.385	0.625	-0.616	0.538
USAS	-1.940	0.935	-2.075	**0.038**
BDAS	4.053	0.940	4.310	**0.000**
Untrained	0.059	0.595	0.100	0.920
Untrained Natural	0.650	0.608	1.070	0.285
USAS:Untrained	0.853	0.638	1.336	0.182
BDAS:Untrained	-1.138	0.657	-1.730	0.084
USAS:Untrained Natural	1.112	0.677	1.644	0.100
BDAS:Untrained Natural	-1.909	0.697	-2.740	**0.006**

The odds of discriminating the length contrast among the Control group participants were 7 times higher Compared to the Unimodal Same Arabic Spelling group (USAS) (CI 95%: 1.11-43.47). No significant difference between the two groups was evident in the Untrained and Untrained Natural pairs. Conversely, learning from exposure to the bimodal distribution with compatible Arabic script significantly increased the likelihood of responding 'Different' to the Trained pairs compared to the Control group. For the Bimodal Distinct Arabic Spelling group (BDAS), the odds of perceiving the length contrast as different were 57.54 times higher than the Control group (CI 95%: 9.11-363.36). However, as the significant interaction between GROUP and TEST TYPE reveals, the difference in accuracy scores between the BDAS and the Control groups on Untrained Natural pairs significantly decreased compared to Trained stimuli, but the difference remained statistically significant (OR= 8.53, CI 95%: 2.40- 30.23). Figure 3.16 illustrates the predicted probabilities obtained from the model.

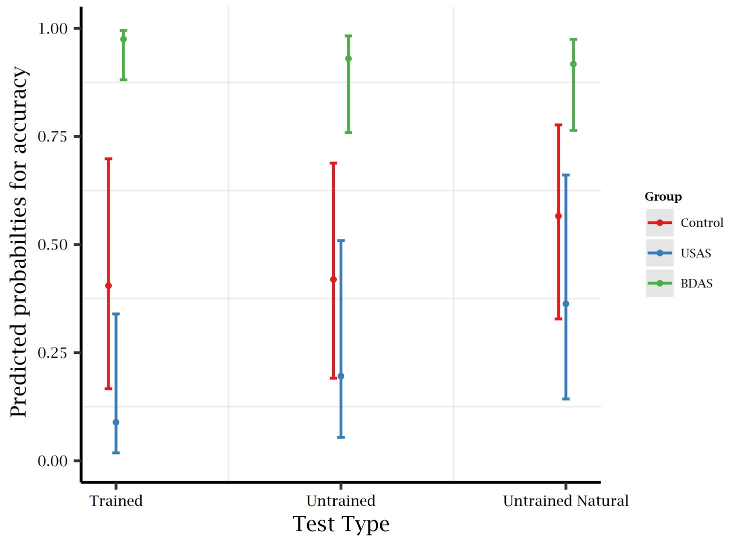


FIGURE 3.16: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.12. The probabilities are presented with 95% confidence interval.

The second model explored whether presenting learners with the unimodal distribution alongside compatible Arabic orthography (same spelling) had an advantage in suppressing a two category perception over unimodal training only. Recall that training learners with the unimodal distribution alone did not lower their discrimination scores compared to the Control group. To this end, a mixed-effects logistic regression analysis was conducted. The fixed predictors were GROUP (Unimodal vs. Unimodal Same Arabic Spelling), TEST TYPE (Trained vs. Untrained and Untrained Natural) and their interaction. SUBJECT and ITEM served as random effects, with by-subject random slopes for TEST TYPE. The model summary is provided in Table 3.13. The results clearly indicate that learners in the Unimodal Same Arabic Spelling group (USAS) were significantly less likely to respond 'Different' to the Trained stimuli than the Unimodal group. The odds of discriminating the length contrast for the Unimodal group were 191.5 times higher compared to those who received unimodal training with compatible Arabic script (CI 95%: 8.9-4121). The model, however, reveals a significant interaction between GROUP and TEST TYPE. On the Untrained and Untrained Natural pairs, the difference between the two groups was significantly reduced compared to their difference on Trained stimuli. The odds ratio dropped to 12.04 (CI 95%: 1.89-76.81) for the Untrained pairs and to 7 (CI 95%: 1.58-31) for the Untrained Natural pairs; however the difference between the two groups on both testing blocks remained statistically significant. Figure 3.13 displays the predicted probabilities of the two learning groups on the three testing blocks.

TABLE 3.13: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on target different items. The reference level for GROUP is Unimodal, and the reference level for TEST TYPE is Trained.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.453	1.097	2.237	0.025
USAS	-5.255	1.564	-3.359	**0.001**
Untrained	-1.317	0.914	-1.441	0.150
Untrained Natural	-1.063	0.932	-1.141	0.254
USAS:Untrained	2.767	1.220	2.268	**0.023**
USAS:Untrained Natural	3.311	1.240	2.669	**0.008**

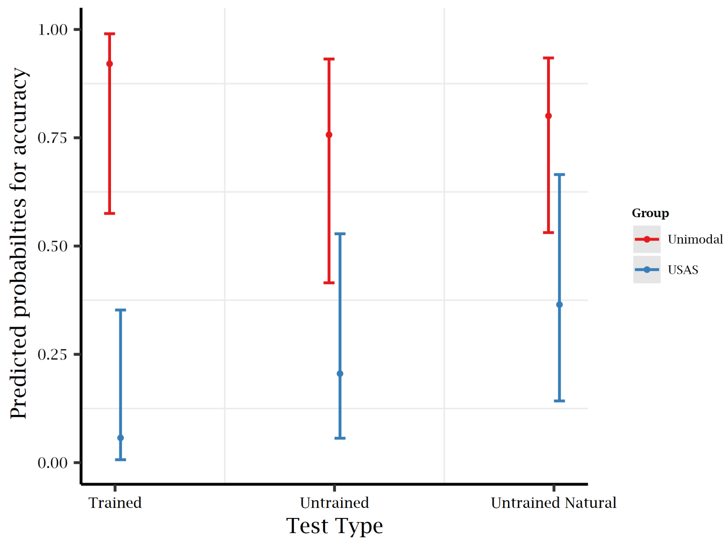


FIGURE 3.17: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.13. The probabilities are presented with 95% confidence interval.

The final model explored whether the exposure to compatible Arabic orthography (distinct spelling) could improve learners' perception of the length contrast over those who were trained with bimodal distribution alone. The fixed effects of the model were GROUP (Bimodal vs. Bimodal Distinct Arabic Spelling), TEST TYPE (Trained vs. Untrained and Untrained Natural). The random effects were SUBJECT and ITEM, with by-subject random slopes for TEST TYPE. The results of the model confirmed that adding compatible Arabic orthography to the bimodal auditory distribution did not significantly increase the likelihood of responding 'Different' to the length contrast compared to bimodal training only. A summary of the model is provided in Table 3.14 and the predicted probabilities of the model are shown in Figure 3.18.

TABLE 3.14: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on target different items. The reference level for GROUP is Bimodal, and the reference level for TEST TYPE is Trained.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.385	0.665	3.586	0.000
BDAS	0.932	0.850	1.096	0.273
Untrained	-0.472	0.579	-0.815	0.415
Untrained Natural	-0.832	0.574	-1.448	0.148
BDAS:Untrained	-0.552	0.576	-0.958	0.338
BDAS:Untrained Natural	-0.317	0.577	-0.550	0.582

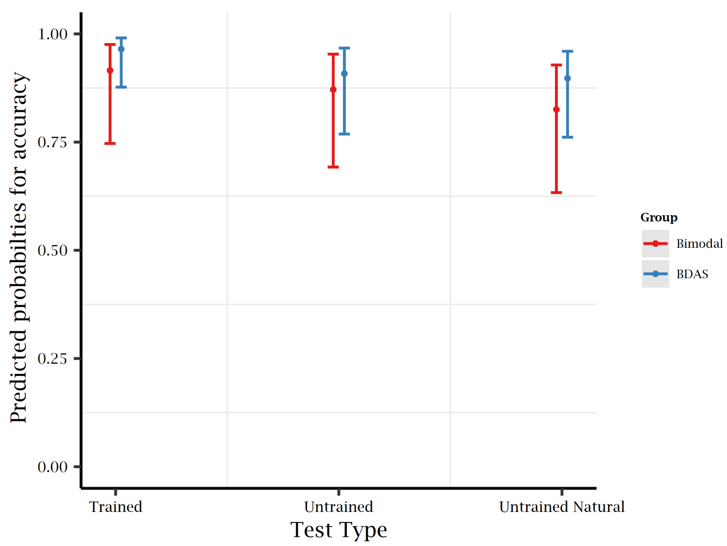


FIGURE 3.18: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.14. The probabilities are presented with 95% confidence interval.

To summarize, the exposure to unimodal distribution with compatible Arabic script significantly lowered learners' accuracy scores compared to unimodal exposure alone. However, the accuracy scores of learners trained with bimodal distribution and compatible Arabic orthographic input failed to significantly improve learners' discrimination scores compared to bimodal training only. The following section investigates the effects of incompatible distributional and orthographic input on learners' discrimination of the length contrast.

3.3.2.2 Incompatible distributional and and orthographic information

The present section examines the interaction between distributional information and incompatible Arabic orthography in the perception of the consonant length distinction. Specifically, the performances of learners in the incompatible Arabic orthography groups (UDAS and BSAS) were compared to the Control, Unimodal and Bimodal groups. The first model analyzed the response accuracy of learners who had the unimodal or bimodal auditory exposure with incompatible orthography in comparison to those who received no training. The model fixed effects were GROUP (Control vs. Unimodal Distinct Arabic Spelling and Bimodal Same Arabic Spelling), TEST TYPE (Trained vs. Untrained and Untrained Natural). SUBJECT and ITEM served as random effects, with by-subject random slopes for TEST TYPE. The results of the model revealed that exposing learners to the unimodal distribution paired with distinct Arabic spellings significantly increased the likelihood of responding 'Different' to the Trained stimuli compared to the Control group. The odds of discriminating the singleton/geminate contrast were 48.45 times higher among learners in the UDAS group compared to the Control (CI 95%: 7.24-324.33). However, a significant interaction between GROUP and TEST TYPE reveals that the difference between the UDAS and the Control groups significantly reduced on Untrained Natural pairs compared to the Trained items; the odds ratio dropped to 10.91 but remained statistically significant (CI 95%: 2.89 -41.22). No significant difference was observed between the BSAS and the Control groups. Table 3.15 summarizes the model and Figure 3.19 illustrates its predicted probabilities.

TABLE 3.15: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on target different items. The reference level for GROUP is Control, and the reference level for TEST TYPE is Trained.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-0.379	0.641	-0.591	0.555
UDAS	3.880	0.970	4.000	**0.000**
BSAS	-1.881	0.966	-1.948	0.051
Untrained	0.027	0.578	0.046	0.963
Untrained Natural	0.626	0.596	1.050	0.294
UDAS:Untrained	0.079	0.624	0.126	0.900
BSAS:Untrained	0.173	0.594	0.291	0.771
UDAS:Untrained Natural	-1.490	0.664	-2.245	**0.025**
BSAS:Untrained Natural	0.898	0.637	1.410	0.159

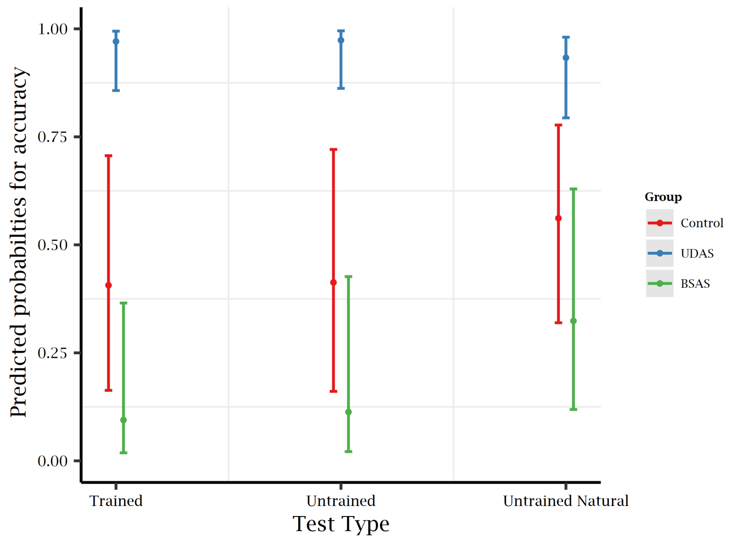


FIGURE 3.19: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.15. The probabilities are presented with 95% confidence interval.

A second analysis modelled the response data from learners who were exposed to the unimodal distribution with or without incompatible Arabic script. The purpose here was to determine if pairing the unimodal distribution with distinct Arabic script would lead to a significantly different discrimination scores compared to the unimodal exposure only. The fixed effects were GROUP (Unimodal vs. Unimodal Distinct Arabic Spelling), TEST TYPE (Trained vs. Untrained and Untrained Natural) and their interaction. SUBJECT and ITEM served as random effects, with by-subject random slopes for TEST TYPE. A summary of the model is shown in Table 3.16.

TABLE 3.16: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on target different items. The reference level for GROUP is Unimodal, and the reference level for TEST TYPE is Trained.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.542	0.653	2.362	0.018
UDAS	1.486	0.864	1.719	0.086
Untrained	-0.398	0.411	-0.968	0.333
Untrained Natural	-0.100	0.412	-0.242	0.809
UDAS:Untrained	0.352	0.347	1.014	0.311
UDAS:Untrained Natural	-0.275	0.344	-0.800	0.423

The model revealed no significant difference between the Unimodal and the UDAS groups on Trained and Untrained Natural items. On Untrained items, however, the difference between the two groups was significant (β= 1.84, SE= 0.86, p=0.03). The odds of discriminating the length contrast for the UDAS were 6.28 times higher compared to the Unimodal group (CI 95%: 1.16-34).

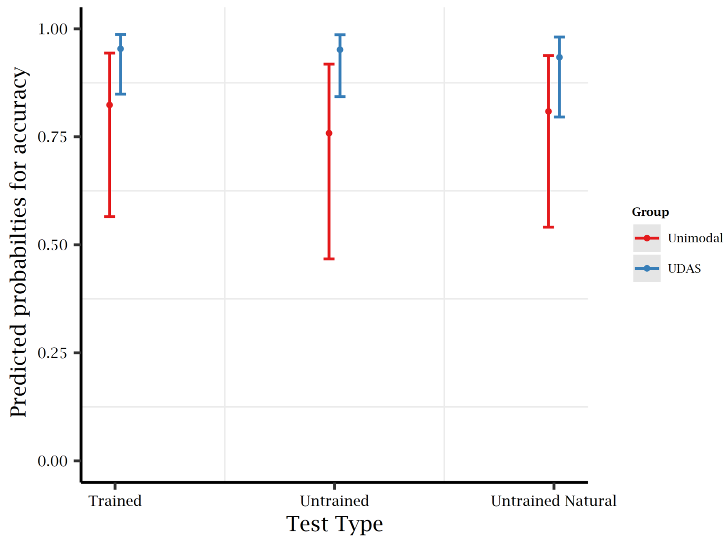


FIGURE 3.20: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.16. The probabilities are presented with 95% confidence interval.

Finally, the same analysis was replicated using response data obtained from learners who received bimodal training with or without incompatible Arabic orthography. The mixed-effects model contained the same fixed and random effects; however, the levels of GROUP were Bimodal and Bimodal Same Arabic Spelling (BSAS). A summary of the model is shown in 3.17.

TABLE 3.17: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on target different items. The reference level for GROUP is Bimodal, and the reference level for TEST TYPE is Trained.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	2.482	0.820	3.027	0.002
BSAS	-4.726	1.069	-4.419	**0.000**
Untrained	-0.256	0.707	-0.362	0.718
Untrained Natural	-0.837	0.722	-1.159	0.246
BSAS:Untrained	0.639	0.748	0.855	0.393
BSAS:Untrained Natural	2.378	0.755	3.152	**0.002**

The results of the model showed that learners who received bimodal training with incompatible orthography (BSAS) were significantly less likely to respond 'Different' to the Trained pairs compared to those who were exposed to the bimodal distribution only. The odds of discriminating the length contrast for the Bimodal group was 98.60 times higher (CI 95%: 14.05-692.30) compared to the BSAS group. However, as evidenced by the significant interaction, the odds ratio dropped to 10.46 (CI 95%: 2.55-42.95) when responding to the Untrained Natural pairs, while remaining statistically significant. An illustration of the predicted probabilities is provided in Figure 3.21.

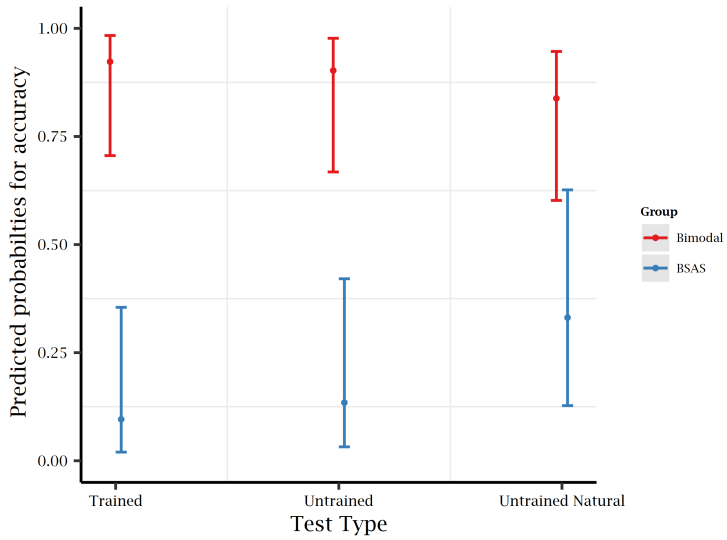


FIGURE 3.21: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.17. The probabilities are presented with 95% confidence interval.

Together, the present and the previous sections confirm that Arabic script significantly affects learners' perception of the length contrast. The effect was attested following either compatible or incompatible distributional and orthographic exposure. The following section examines if the distributional information would still make any contribution to learning in the presence of Arabic orthographic input.

3.3.2.3 Compatible versus incompatible auditory and orthographic stimuli

The purpose of the present section was to investigate whether learners relied solely on the Arabic written input in order to recover the target structure during exposure. Two models examined the performances of learners who received either the unimodal or bimodal exposure represented by the same written input. The first model compared the accuracy scores for learners who were exposed to either the unimodal or bimodal distribution paired with a single written representation for short and long consonants. The fixed predictors of the model were GROUP (Unimodal Same Arabic Spelling vs. Bimodal Same Arabic Spelling), TEST TYPE (Trained vs. Untrained and Untrained Natural) and the interaction between the two predictors. The random intercepts were SUBJECT and ITEM, with by-subject slopes for TEST TYPE. The results of the model failed to detect any difference in the likelihood of responding 'Different' between the two groups. The only significant difference obtained from the model was between the performance of the USAS group on the Untrained Natural items compared to Trained pairs. The odds ratio for accuracy on the Untrained Natural stimuli was 11.22 times greater than the accuracy on the Trained items (CI 95%: 1.7-72.22). A summary of the model is given in Table 3.18. Figure 3.22 displays the predicted probabilities obtained from the model.

TABLE 3.18: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on target different items. The reference level for GROUP is USAS (Unimodal Same Arabic Spelling), and the reference level for TEST TYPE is Trained.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-2.970	1.100	-2.701	0.007
BSAS	0.035	1.429	0.025	0.980
Untrained	1.532	0.925	1.656	0.098
Untrained Natural	2.418	0.950	2.546	**0.011**
BSAS:Untrained	-0.516	1.030	-0.501	0.616
BSAS:Untrained Natural	-0.216	1.071	-0.202	0.840

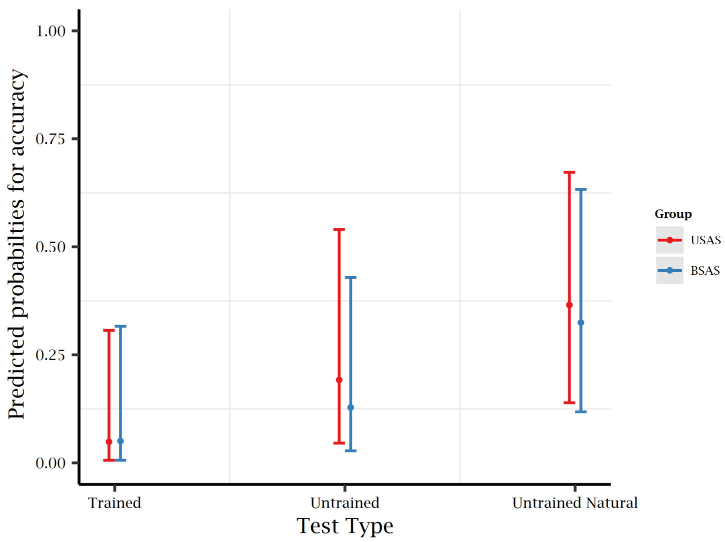


FIGURE 3.22: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.18. The probabilities are presented with 95% confidence interval.

In the same way, a second model was conducted to compare the accuracy scores of learners who were trained with either the unimodal or bimodal distribution that was represented orthographically by two distinct labels for short and long consonants. GROUP in this model contained data from the Bimodal Distinct Arabic Spelling (BDAS) and Unimodal Distinct Arabic Spelling (UDAS) groups and the same fixed and random structures were used.The only significant difference obtained from the model was between the accuracy of the BDAS group on the Untrained Natural items compared to Trained pairs. The odds ratio for accuracy on the Trained pairs was 3.81 times greater than the accuracy on the Untrained Natural items (CI 95%: 1.36-10.70). No significant differences were detected between the two groups (see Table 3.19). The predicted probabilities of the model are shown in Figure 3.23.

TABLE 3.19: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on target different items. The reference level for GROUP is BDAS (Bimodal Distinct Arabic Spelling), and the reference level for TEST TYPE is Trained.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	3.415	0.743	4.598	0.000
UDAS	0.024	0.975	0.025	0.980
Untrained	-0.997	0.538	-1.854	0.064
Untrained Natural	-1.340	0.526	-2.547	**0.011**
UDAS:Untrained	0.489	0.578	0.846	0.398
UDAS:Untrained Natural	0.134	0.561	0.239	0.811

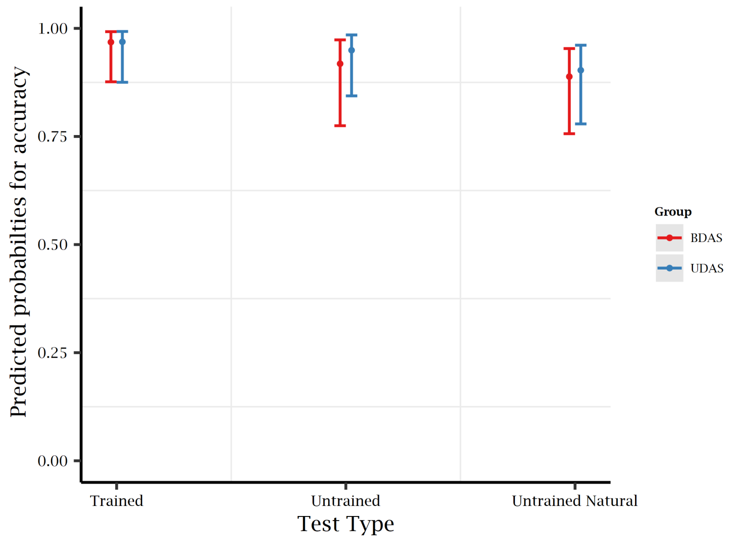


FIGURE 3.23: Predicted probabilities from the mixed-effects logistic regression results modeling the likelihood of responding 'Different' to the length contrast by GROUP, TEST TYPE and their interaction term, as presented in Table 3.19. The probabilities are presented with 95% confidence interval.

The results of these two models indicate that, when accompanied by Arabic orthographic input, distributional information had no effect on the perception of length contrast. Learners' accuracy on the AX task was determined by whether the Arabic orthographic input contained a uniform or distinct spellings of the consonant length contrast rather than by the type of the auditory distribution.

3.3.2.4 A combined analysis of all the experimental groups

The final analysis included response data from all the experimental groups and the Control group. The purpose of this analysis was to show if the performances of those exposed to Roman script and Arabic script were significantly different. To simplify the analysis, the model contained only GROUP as the fixed factor. The random effects were SUBJECT and ITEM. The results showed that compared to the Control group, the BDRS, UDAS, UDRS, BDAS, Bimodal and the Unimodal groups were significantly more likely to respond 'Different' to the length contrast. On the contrary, only the BSAS group was significantly less likely to perceive the length distinction compared to the Control group. To show the size of the difference between these learning groups and the Control group, Figure 3.24 displays the odds ratios ordered from the highest to the lowest.

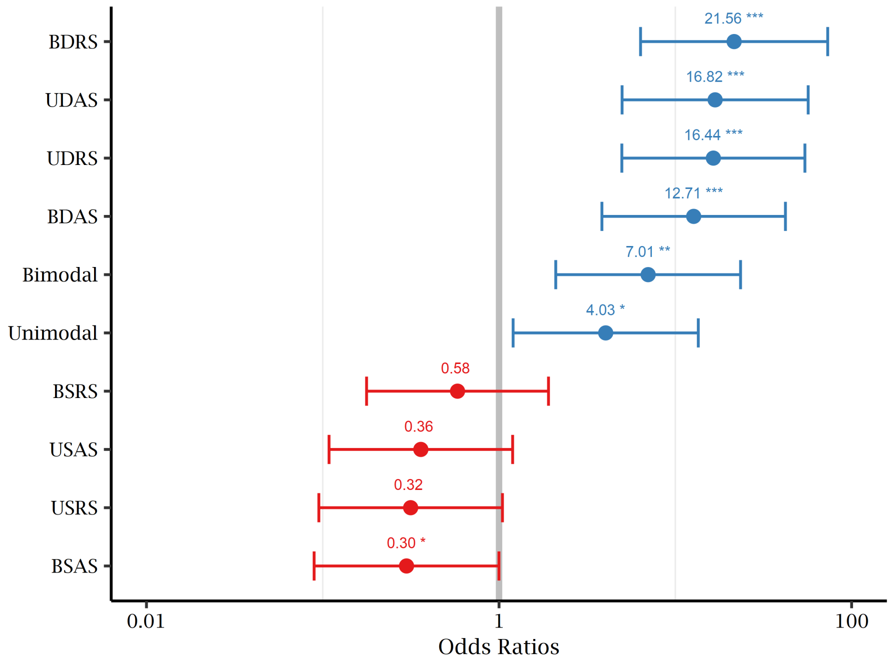


FIGURE 3.24: The odds ratios of responding 'Different' to the length contrast for the 10 experimental groups compared to the Control group (represented by the vertical grey line). Odds ratios in blue indicate that they are higher than 1; conversely, odds ratios in red indicate that they are lower than the 1. Asterisks indicate that the significance level of the odds ratio compared to the Control group. The odds ratios are presented with 95% confidence interval.

Unsurprisingly, learners who were trained with the bimodal distribution and compatible Roman script (BDRS) had the highest odds ratios (OR= 21.5, 95% CI: 6.35-73.18). Interestingly, the groups that followed were trained with unimodal distributions paired with incompatible Roman (UDRS) or Arabic script (UDAS) (OR= 16.44, 95% CI: 4.98-54.33) and (OR= 16.82, 95% CI: 4.99-56.70), respectively. Finally, the group that was exposed to the bimodal distribution and compatible Arabic spellings (BDAS) came fourth with an OR of 12.71 (95% CI: 3.83-42.13). To assess the difference between the four groups that were exposed to an orthographic input that marked the distinction between short and long consonants, multiple planned comparisons were conducted. The results revealed no significant differences in discrimination scores between the four groups.

For the groups that were exposed to an orthographic input that supported a single category perception, only learners who had the bimodal training with incompatible Arabic script (BSAS) performed significantly worse than the Control group. However, multiple planned comparisons revealed no significant differences between the BSAS and the BSRS, USAS and USRS groups.

3.3.3 Discussion

The purpose of this experiment was to verify whether native English speakers could utilize Arabic orthographic input to perceive the length contrast the same way they did with the Roman script. Specifically, given the strong orthographic effect shown in the previous experiment, it was hypothesized that L1 orthographic representation were largely responsible for the observed learning effect and, thus, exposing native English speakers to Arabic script would not produce comparable results. On the contrary, the findings of this experiment exhibited similar patterns. As with Roman script, learners trained on the unimodal distribution with compatible Arabic spellings had significantly lower accuracy scores compared to the Control group. Likewise, the exposure to Arabic orthographic input led to a significant increase in the sensitivity to the length contrast for learners who had bimodal training compared to the Control group. However, unlike Roman orthography, pairing the bimodal training with distinct Arabic spellings failed to lead to a significantly better perception of the length contrast than bimodal training alone.

Pairing auditory distributions with incompatible Arabic input produced similar results as well. Compared to the Control group, exposing learners who had unimodal training to Arabic script in which tokens 6 to 10 spelled with a diacritic significantly increased the discrimination accuracy of the length contrast, while failing to significantly decrease it for learners who were trained on bimodal distribution with spellings that contained no diacritic. Still, training learners with the bimodal distribution represented consistently without a diacritic significantly reduced the discrimination of the length contrast compared to bimodal training alone.

These results provide evidence that, like familiar script, the exposure to unfamiliar orthographic information can aide the distributional learning of nonnative speech sounds. Further, when the statistical and the orthographic information do not match, learners exhibit reliance on the unfamiliar orthographic cues to perceive L2 sounds. These results and the results of the previous experiment are discussed further in the following section.

3.4 General discussion

This chapter set out to explore an often overlooked aspect of L2 phonological development. More specifically, two experiments were conducted to understand how orthographic information interacts with the distributional learning of the consonant length contrast. To achieve this, 242 monolingual English speakers were assigned to 10 experimental groups and a control group. These groups differed on whether the auditory distribution was unimodal or bimodal, and if there was an orthographic exposure, whether it was familiar, unfamiliar, compatible or incompatible. The results of these experiments have offered clues to the research questions laid out at the outset of this chapter.

The first question concerns whether the exposure to written input facilitates the distributional learning of the consonant length contrast. The results of Experiment 1 clearly suggest that familiar orthographic information facilitates the distributional learning of the consonant length contrast. Perhaps the most striking effect of written input was the suppression of the two category perception observed when training learners with a unimodal distribution only. Indeed, pairing a unimodal exposure with a compatible Roman script led learners to a significantly lower discrimination of the length contrast than those trained with only a unimodal distribution. The

TABLE 3.20: A summary table of the difference in mean proportion of accuracy for the orthography groups compared to the Control, Unimodal and Bimodal groups on the three test types. Values in bold indicate that the difference is significant.

Test	Trained			Untrained			Untrained Natural		
Group	Control	Unimodal	Bimodal	Control	Unimodal	Bimodal	Control	Unimodal	Bimodal
USRS	**-0.21**	**-0.45**	**-0.55**	-0.14	**-0.34**	**-0.45**	-0.15	**-0.29**	**-0.34**
BDRS	**0.52**	**0.28**	**0.18**	**0.43**	**0.23**	0.12	**0.3**	0.16	0.11
UDRS	**0.45**	0.21	0.11	**0.45**	**0.25**	0.14	**0.29**	0.15	0.1
BSRS	-0.04	**-0.28**	**-0.38**	-0.04	**-0.24**	**-0.25**	-0.04	**-0.22**	**-0.17**
USAS	**-0.17**	**-0.41**	**-0.51**	-0.13	**-0.33**	**-0.44**	-0.13	**-0.27**	**-0.32**
BDAS	**0.41**	0.17	0.07	**0.36**	0.16	0.05	**0.27**	0.13	0.08
UDAS	**0.39**	0.15	0.05	**0.38**	**0.18**	0.07	**0.25**	0.11	0.06
BSAS	-0.17	**-0.41**	**-0.51**	-0.23	**-0.34**	**-0.45**	-0.13	**-0.27**	**-0.32**

facilitating effect of compatible orthographic input has also been attested in learners exposed to a bimodal distribution. Integrating a bimodal auditory distribution with compatible orthographic representations in support of two category perception resulted in a significant increase in the discrimination of the length contrast compared to bimodal exposure alone. The enhancing effect, however, was limited to the trained items.

The second research question was in relation to the competition between the distributional and orthographic properties when they convey contradictory information. It was hypothesized that learners would utilize both sources of information to perceive the length contrast and thus exhibit differential performances. However, the findings of these experiments confirm that orthographic forms, irrespective of the auditory distribution they represent, generally guide learners' perception. That is, when the distributional and orthographic information conflict, learners exhibit almost a complete reliance on the written cues to perceive the length contrast. Yet, there were a few instances in which distributional information contributed to the learning outcome. Among these was the performance of learners who were exposed to a bimodal distribution represented orthographically with compatible Roman spellings versus the performance of learners who were exposed to the same spellings but with a unimodal distribution instead. Here, the bimodal distribution seems to induce a significant added benefit to the discrimination of the length contrast.

The final research question in this chapter was whether the perception patterns established in Experiment would differ with unfamiliar orthography. The findings suggest that Arabic script can exert similar learning effects as Roman orthographic input. The perception of the length contrast for learners who were exposed to Arabic script mainly followed the patterns of learners who were presented with Roman script. These findings are contrary to previous word learning studies which have suggested that novel orthographic representations, particularly Arabic, can hinder learning (Mathieu, 2016) or do not impact the acquisition of nonnative phonological forms either positively or negatively (Showalter & Hayes-Harb, 2015). Thus, the results of the present chapter provide evidence that the mostly neutral effects of unfamiliar orthography observed in these studies is unlikely to be related to an inherent difficulty in processing novel written representations. This discrepancy between the results of these studies and the current one maybe related to a number of factors including the difference in the levels of representation investigated in these studies and the current one (pre-lexical vs. lexical), the difference in the frequency of exposure to the target orthography (learners in the current study saw the Arabic script 400 times), and finally, the difference in the relative difficulty of the target contrast (length contrast vs. /χ/- /ħ/ contrast (Mathieu, 2016) and the /k/-/q/ contrast (Showalter & Hayes-Harb, 2015)).

Together, these experiments reveal that, in general, adult learners are not always adept at learning the underlying structure of a nonnative auditory input via its distributional properties. Rather, the more explicit cues, such as orthographic information, are more readily utilized by learners to perceive nonnative speech sounds. The evidence suggests that nonnative sound categories are largely modulated by the orthographic input; learners determine the number of categories along a particular length continuum based on the written input.

Although no previous study has examined the effects of orthographic input on the distributional learning of nonnative phonemes, it is still worth comparing the results of the current study to the existing research that investigated the interaction between auditory and orthographic or visual information in nonnative speech perception. Generally, prior training studies report no measurable effect of orthography on the perception of nonnative phonemes. In Simon et al. (2010), native English speakers were trained on the French vowel /u/-/y/ contrast using novel words presented with their meanings as images. The experimental variable was whether the exposure contained the spellings of these words. A follow-up perceptual task revealed no differences between learners who were presented with orthography and those who were not. Likewise, Pytlyk (2011) taught three groups of native English speakers Mandarin contrasts using either Pinyin, Zhuyin or no orthography during three sessions of training. The perception task did not reveal any differences in discrimination scores between the three groups.

While these studies delivered the auditory input through word learning (Simon et al., 2010), or explicit instruction (Pytlyk, 2011), a similarly-designed investigation of the interaction between distributional and visual information has been reported. Hayes (2003) exposed native English speakers to multiple sound continua containing the novel pre-voiced [g] and the unaspirated voiceless [k] contrast. The visual information embedded with the auditory exposure was pictured meanings instead of orthographic input. Two groups were trained with a unimodal distribution of the contrast coupled with either a single pictured meaning that promoted a single category perception or two pictured meanings that would counter that effect. Similarly, the two other groups had a bimodal training paired with either two pictured meanings supporting the discrimination of the contrast or a single one that promoted a single category perception. Contrary to the results of the current study, the visual information had limited effect in either reinforcing or suppressing two category perception. This effect was evident only in the difference in accuracy scores between learners who had a bimodal training with two pictured meanings versus learners trained on a unimodal distributions represented with a single pictured meaning. Although it is tempting to attribute the discrepancy between these findings and the findings of the current study to the distinct role of orthographic information in determining phonemic representations, the target contrast is markedly more difficult to acquire than the length contrast, as evidenced by the comparatively low discrimination scores of the [g]-[k] distinction. Further discussion of these results and their broader implications are offered in Chapter 5.

In short, this chapter provides insights into how orthographic information interacts with frequency-based perceptual learning during the early stages of L2 category formation. In these experiments, L2 learners' perception of nonnative speech sounds was largely shaped by the orthographic forms rather than distributional or the acoustic properties of the auditory input. Both familiar and unfamiliar orthographic cues can have an additive learning effect when they are integrated with matching auditory input. They can also reorganize the phonemic representations independent of the statistical information associated with the auditory input. The next chapter extends the inquiry to the role of orthography in L2 phonology by investigating its effect on the lexical encoding of the consonant length contrast.

Chapter 4

Orthographic Input and Lexical Learning

4.1 Introduction

The previous chapter has demonstrated how orthographic input interacts with the distributional learning of consonant length contrast: Access to orthographic information shapes learners' phonemic representations regardless of the statistical and the acoustic properties of the auditory input. The level of investigation so far concerns the perceptual learning of nonnative speech categories, which constitutes a distinct domain within L2 phonological development that forms sometimes asymmetrical relationship with lexical learning. A number of studies have shown that learning to perceive a novel contrast does not necessarily indicate that the contrast can be encoded in learners' L2 lexicon (Sebastian-Galles & Baus, 2005; Díaz, Mitterer, Broersma, & Sebastián-Gallés, 2012; Amengual, 2016). Likewise, the ability to encode a nonnative contrast in the lexicon does not always require prior phonetic knowledge (Weber & Cutler, 2004; Cutler, Weber, & Otake, 2006; Escudero et al., 2008; Darcy et al., 2012).

Therefore, any conclusions drawn about the role of orthography in L2 phonology would be incomplete without considering the process of learning nonnative speech sounds as a whole. This chapter extends the inquiry to the process of lexical encoding of nonnative contrasts. Specifically, the main objective is to understand how orthographic information interacts with learners' ability to establish lexical representations for the consonant length contrast.

The bulk of the empirical research on orthographic effects cited in this dissertation has addressed this issue. The general consensus in these studies suggests that familiar orthographic input that shares learners' L1 grapheme-phoneme correspondences facilitate learning the nonnative phonological forms of novel words, and the existence of incompatible L1 and L2 grapheme-phoneme correspondences may impair the the establishment of lexical distinctions (Hayes-Harb et al., 2010; Escudero et al., 2014; Showalter, 2018). Further, unfamiliar written forms generally exert no effect in lexical learning (Showalter & Hayes-Harb, 2015; Hayes-Harb & Hacking, 2015; Mathieu, 2016); however, in some cases unfamiliar orthography may negatively interfere with the formation of distinct lexical representations (Mathieu, 2016), while in others minimally unfamiliar script may facilitate the process (Showalter & Hayes-Harb, 2013).

The current experiments differ from previous studies in the degree of orthographic depth examined. In line with the previous chapter, the orthographic exposure varies in two ways: Orthographic compatibility and orthographic familiarity. The orthographic representations of the length contrast are deemed compatible if the target stimuli has one-to-one phoneme-grapheme correspondence. Conversely, if short and long consonants that make up a phonological length distinction are represented by the same grapheme, the orthographic input is considered incompatible. In addition, orthographic familiarity is determined by whether learners' L1 shares the same writing system with L2 or not. This chapter intends to answer the following three questions: (I) Does the exposure to compatible orthographic input facilitate encoding consonant length in the lexicon? (II) if the orthography does not mark a length distinction, will the lexical encoding of the contrast be inhibited? (III) and finally, will the unfamiliarity with the Arabic orthographic input impact the lexical representations of consonant length contrast differently?

4.2 Experiment One

The first experiment examined native English speakers' ability to learn the phonological forms of new words containing a consonant length distinction. The experimental groups differed on whether their exposure included Roman orthographic input, and if the orthography supported a length distinction. The three groups were as follows: No Spelling (Control), Distinct Roman Spelling (DRS) and Same Roman Spelling (SRS).

The previous experiments in Chapter Three (3) confirmed that learners utilize the available orthographic information to assist in the perception of the length contrast. Following these results, it was expected that a similar pattern would emerge during the lexical learning of the contrast. In particular, access to compatible Roman orthography would result in an enhanced ability in encoding and retrieving words that contained the distinction. Further, the exposure to incompatible orthographic input was expected to interfere with the lexical learning of the length contrast and lead to faulty lexical representations for the new words.

4.2.1 Methods

4.2.1.1 Participants

Seventy-five monolingual English speakers (55 females and 20 males) were recruited from both the University of Alberta Department of Linguistics participant pool and the Edmonton area to take part in the study. Four of the participants received $10 for their participation, and the remaining participants received course credit. The mean age was 22.5 (*SD*= 7.7, range= 18-61). Participants reported no formal or informal experience in languages that have phonemic length contrast. None of the participants reported any hearing or speech impairments.

4.2.1.2 Stimuli

The experimental materials consisted of 12 Arabic pseudo-words produced by a 38 year-old male native Arabic speaker from Dammam, Saudi Arabia. Six of the 12 words contained singleton consonants and the remaining six contained geminates. The target consonants were the same ones used in the previous chapter (/m/ and /n/). Each target consonant was embedded in a CVCV structure where the word initial consonant was either /b/, /d/, /t/, /k/, /s/ or /l/; the vowels preceding the target consonants were either /a/ or /u/; and the vowels following the target consonants were either /aː/ or /uː/.

As previously noted, the acoustic correlates of gemination may include secondary temporal manipulations of the neighboring vowels besides the primary durational cue associated with geminate consonants. In Japanese, vowels preceding geminate consonants are typically longer while vowels following geminates have shorter durations (Idemaru & Guion, 2008). Perceptually, the length of vowels preceding geminate stops is correlated with enhanced sensitivity; longer vowels tend to enhance the perception of geminate stops (Takeyasu & Giriko, 2017). In Lebanese Arabic, the pattern is reversed; only phonologically long vowels become shorter when they precede geminates, while short vowels lengthen following long consonants (Al-Tamimi & Khattab, 2015). Based on the acoustic measurements of the training stimuli in this experiment, the preceding vowel is slightly shorter before geminate consonants, while the following vowel is lengthened. However, the duration of the consonant remains the most consistent cue for gemination. Table 4.1 summarizes the duration of the target consonants in addition to the preceding and the following vowels.

Each word was randomly paired with a visual referent displaying an object. The pictured-meanings were retrieved from The Bank of Standardized Stimuli (BOSS) (Brodeur, Guérard, & Bouras, 2014). In addition to pictured-meanings, the Roman spellings of these words were provided to the orthography groups: For the Distinct Roman Spelling group geminates were spelled with double letters and singletons

TABLE 4.1: The duration (in milliseconds) of the segments in singleton and geminate contexts. **V1** refers to the vowel preceding the target consonant; **C** refers to target consonant; and **V2** refers to the vowel following the target consonant.

Singleton	V1	C	V2	Geminate	V1	C	V2
/la**ma**/	100	79	285	/la**mm**a/	65	234	316
/ta**na**/	121	66	375	/ta**nn**a/	110	257	383
/su**ma**/	84	87	305	/su**mm**a/	82	276	286
/ka**mu**/	94	98	275	/ka**mm**u/	81	271	331
/du**na**/	120	77	331	/du**nn**a/	105	263	349
/ba**nu**/	123	76	264	/ba**nn**u/	104	297	340

were spelled with a single letter, and for the Same Roman Spelling group both geminate and singleton consonants were spelled with a single letter.

4.2.1.3 Training procedure

The design of the training procedure followed the word-learning paradigm implemented in multiple studies that investigated the role of orthographic input in the lexical encoding and retrieval of words containing nonnative speech sounds (Hayes-Harb et al., 2010; Showalter & Hayes-Harb, 2013, 2015; Mathieu, 2016; Hayes-Harb & Cheng, 2016; Showalter, 2018). The training consisted of two parts: (1) Learning phase followed by a (2) Criterion test. During the learning phase, all participants were presented with the auditory forms of the minimal pairs along with their pictured-meanings. Each word was played while its corresponding image simultaneously appeared on the display screen and remained for 3 seconds before the next word was presented. In the orthography groups, participants were additionally exposed to the Roman spellings of these words, whereas the Control group saw instead <XXX> along with each auditory token and its visual referent. In the compatible orthography group, geminates were spelled with double letters, while both geminates and singletons were spelled with single letters in the incompatible orthography group (see Table 4.2 for an illustration). The spelling of each the auditory item was placed below the corresponding image. The participants went through four blocks of training and in each block the 12 words were presented randomly once, bringing the total number of learning trials to 48. An example of a trial is presented in Figure 4.1.

TABLE 4.2: An example of the auditory and visual stimuli presented for each learning group.

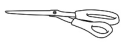
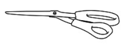
Group	Auditory item	Pictured-meaning	Spelling
Control	/kammu/		XXX
DRS	/kammu/		kammu
SRS	/kammu/		kamu

FIGURE 4.1: An example of a trial presented to the participants in the Distinct Roman Spelling group during the learning phase. Participants heard the auditory form /tanna/ while simultaneously saw its pictured-meaning and spelling.

Once the learning phase concluded, participants performed a criterion test to examine their ability to associate the auditory items with their corresponding images. The purpose of this test was to ensure that the phonological forms of the these words were successfully learned. During each trial, participants heard a word and saw an image on screen and had to indicate whether the auditory item matched the pictured-meaning by pressing the 'Yes' or 'No' buttons. The number of trials in the criterion test was 24: 12 Matched items and 12 Mismatched items. The Matched items consisted of auditory words that were paired with their correct pictured-meanings; while the Mismatched items contained auditory items that were paired with pictured-meanings of other auditory words that differed by multiple phonemes. For instance, the auditory item /bannu/ was paired with the image of the word /duna/ (See Table 4.3). Once the auditory item was presented with the pictured-meaning, the picture stayed on the screen for three seconds, during which a response had to be made. If no response was registered, the next stimuli were presented. Participants who reached 90% accuracy on the criterion test advanced to the final test, and those who failed had to restart the learning phase and retake the test until they reached the target score.

The experiment took place in a sound-attenuated booth where participants sat in front of a computer screen and were fitted with over-the-ear headphones. Prior to the start of the learning phase, the participants were told that they would hear multiple words in an unfamiliar language and see their meanings as images on the screen and also see their spellings in the orthography groups. The experiment was conducted using the E-Prime 2.0 software (Psychology Software Tools, 2012). The training and testing phases lasted an average of 30 minutes.

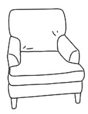
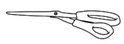
TABLE 4.3: An example of the auditory and visual stimuli pairing presented during the criterion test.

List	**Auditory item**	**Pictured-meaning**	**Correct response**
Matched	/bannu/		Yes
Mismatched	/duna/		No

4.2.1.4 Final test

The final test aimed to examine learners' ability to lexically encode the consonant length contrast following exposure. Similar to the criterion test, learners were presented with an auditory item and a visual referent and had to decide if they matched or not. However, the Mismatched list in the final test contained pairings of words that were minimally contrastive. That is, each pair consisted of two items that only differed in whether the word-medial consonant was a singleton or a geminate. For example, in one of the trials, participants heard the word /bannu/ and saw the pictured-meaning of the word /banu/. Table 4.4 provides an illustration. Participants completed 24 randomly presented trials (12 Matched; 12 Mismatched).

TABLE 4.4: An example of the auditory and visual stimuli pairing presented during the final test.

List	Auditory item	Pictured-meaning	Correct response
Matched	/banu/		Yes
Mismatched	/banu/		No

4.2.2 Results

First, the three experimental groups required similar numbers of learning cycles to reach the 90% threshold in the criterion test. The average number of learning cycles for the Control group was 2.52 (*SD*= 1.62, range=1-6), while the average number of cycles for the Distinct and Same Roman Spelling groups was 2.64 (*SD*= 1.58, range=1-8) and 2.41 (*SD*= 1.45, range=1-8), respectively.

For the final test, the accuracy scores for the three groups on Matched and Mismatched pairs are displayed in Figure 4.2. On Matched items, the proportions of correct responses for the three groups were similar. Both the Control and Distinct

Roman Spelling groups had the same mean accuracy score of 0.88 (*SD*= 0.32), while the mean accuracy for the Same Roman Spelling group was slightly higher (*M*= 0.90, *SD*= 0.30).

On Mismatched pairs, the accuracy scores of the three groups display different distributions. For the Control group, the accuracy scores exhibits a unimodal distribution, with the mean proportion of correct responses at 0.28 (*SD*= 0.19, range= 0-0.73). The distribution of the accuracy scores for the Distinct Roman Spelling group has a higher range (0.09-1) and seems to display two modes; one at around 0.35 and the other at 0.75, with the mean proportion of correct responses occurring at 0.46 (*SD*= 0.25). Finally, the distribution for the Same Roman Spelling group displays a range similar to the Control group (0-0.73); however, the distribution is thicker at the lower end and almost shows a uniform shape until it narrows at around 0.35, with the mean at 0.25 (*SD*= 22).

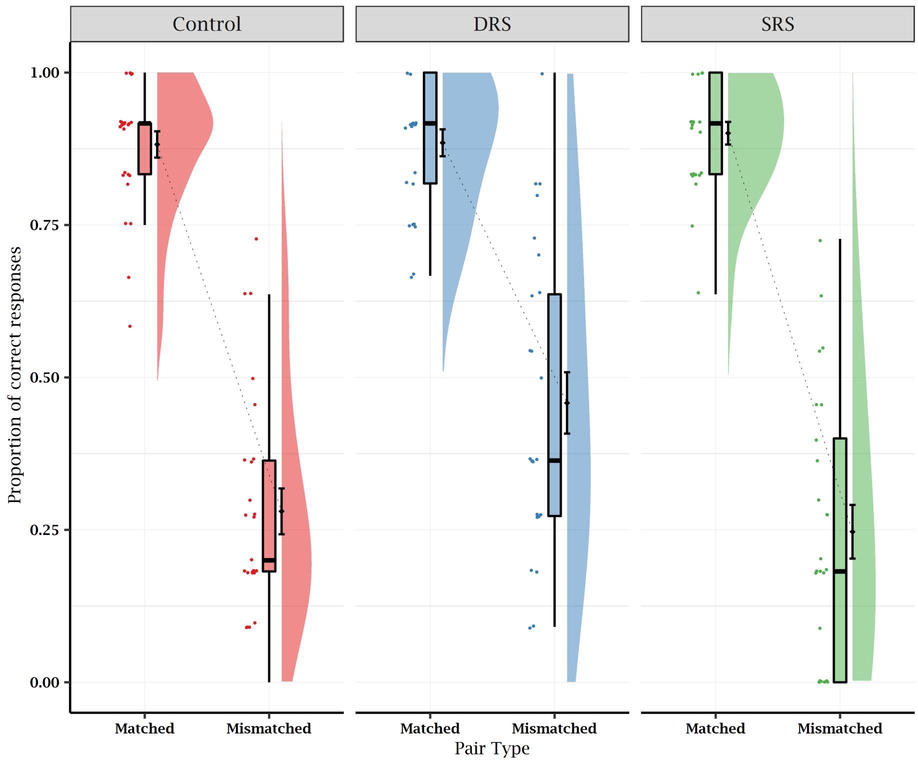


FIGURE 4.2: A combined plot representing the accuracy of the Control group, the Distinct Roman Spelling group (DRS) and the Same Roman Spelling group (SRS) on Matched and Mismatched items. The plot contains probability distributions (split-half violin); box-plots and means with standard errors for each group. Data from each individual are represented as points.

To examine whether the exposure to familiar orthographic information was associated with the lexical encoding of the length contrast, a mixed-effect logistic regression model was fitted in R using the lme4 package (Bates et al., 2015). The outcome variable was the response to the auditory-picture matching task (YES vs. NO) for Mismatched items[1]. The best fitting model, which was determined by the likelihood ratio test, contained only GROUP (Control vs Distinct Roman Spelling and Same Roman Spelling) as the fixed effect. CYCLE (the number of learning cycles required to advance to the final test) did not improve the model fit and, thus, was eliminated. The random effects were SUBJECT and ITEM. A summary of the model is provided in Table 4.5.

TABLE 4.5: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on the auditory-picture matching task. The reference level for GROUP is Control and the reference level for PAIR TYPE is Matched

.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.079	0.237	-4.554	0.000
DRS	0.887	0.326	2.722	**0.006**
SRS	-0.254	0.336	-0.757	0.449

The model revealed that learners who had access to compatible orthographic representations of the stimuli were significantly more likely to respond "NO" to the Mismatched auditory and pictured-meaning than the Control group. The odds of accurately distinguishing between the minimal pairs for the Distinct Roman Spelling group (DRS) were 2.43 times that of the Control group (CI 95%: 1.82-4.6). The model did not show any significant difference between the performances of learners in Same Roman Spelling and the Control groups. To examine the difference between the two orthography groups, the reference level for GROUP was set to Same Roman Spelling (SRS). The results showed that learners who were exposed to compatible spellings were significantly more likely to respond "NO" to the Mismatched items than learners who were presented with incompatible spellings (β= 1.14, SE= 0.33, p<0.001). The odds of accuracy for the Distinct Roman Spelling group (DRS) was

[1]Analysis of Matched items revealed no significant difference between the three groups.

3.13 times greater compared to the Same Roman Spelling group (SRS) (CI 95%: 1.64-6.01). Figure 4.3 displays the predicted probabilities for each learning group.

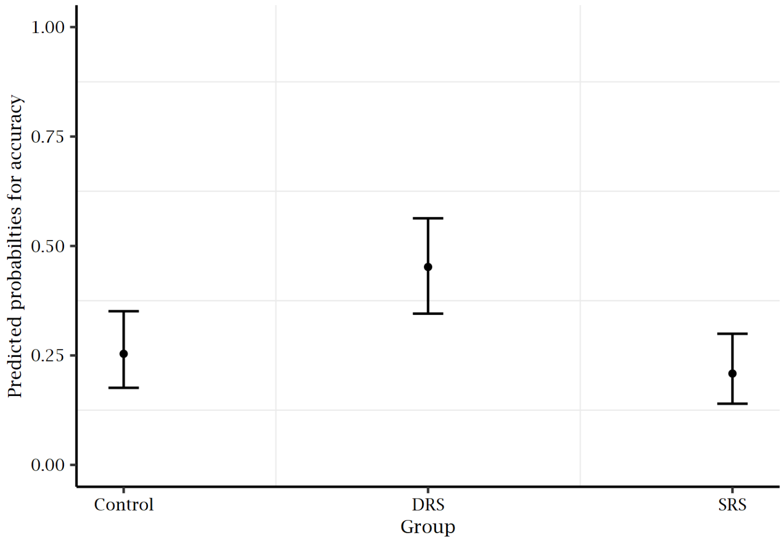


FIGURE 4.3: Predicted probabilities from the mixed-effects logistic results modelling the likelihood of performance accuracy on Mismatched items for the three groups, as presented in Table 4.5. The probabilities are presented with 95% confidence interval.

4.2.3 Discussion

The purpose of this experiment was to determine if familiar orthographic input impacts learners' ability to establish and subsequently use distinct lexical representations for the length contrast. Specifically, the experiment sought to examine whether presenting the auditory forms and pictured-meanings with compatible and incompatible spellings influences learning the phonological forms of new words. It was predicted that the impact of orthography on lexical learning would parallel its impact on perceptual learning. That is, access to compatible Roman orthography would enhance learners' accuracy in learning the new words and the exposure to incompatible orthography would interfere with learners' ability to accurately match the auditory forms with their corresponding images. These predictions were partially confirmed.

First, as predicted, the presence of a familiar orthographic input that marked the length distinction increased learners' ability to establish lexical representations for the length contrast. Learners who received such exposure significantly outperformed those who received no orthographic input or received spellings that did not mark the contrast. Second, contrary to prediction, presenting learners with identical spellings of the length contrast led to performances similar to the Control group. The effect of incompatible orthographic input was neutral; it neither helped nor inhibited the lexical encoding and retrieval of the length contrast.

However, upon closer inspection of the data from the two groups, individual performances within each group vary greatly. Almost all participants who completely failed to associate the auditory forms with their correct meanings on Mismatched items belonged to the Same Roman Spelling group (7 out of 8 participants). This indicates that learners who received identical spellings of the length contrast may have relied on different strategies to resolve the learning difficulty. It is plausible that the incompatible orthographic input was partially responsible for the failure of these learners to establish separate lexical representations for the length contrast. Further discussion is provided in Section 4.4.

4.3 Experiment Two

The second experiment investigated the the effects of unfamiliar orthographic information on the lexical encoding of the consonant length contrast. In particular, instead of the familiar Roman script, learners were exposed to Arabic spellings of the auditory forms. To this end, two additional experimental groups were designed: Distinct Arabic Spelling (DAS) and Same Arabic Spelling (SAS). The previous experiment revealed that learners utilized the congruent Roman script to establish distinct lexical representations for the length contrast; however, when the script did not support the contrast, access to the the orthographic input failed to induce any significant learning effect.

In light of these results and the results of the previous chapter, it was predicted that a similar pattern would emerge. As shown in Chapter 3, learners could exploit the Arabic script to aid the perception of the length distinction. In fact, both Roman and Arabic scripts had similar effects on learners' perceptual sensitivity to the length contrast following exposure. Thus, it was expected that compatible Arabic orthography would enhance learners' ability to lexcially encode the contrast. For incompatible Arabic spellings, it was predicted that the presence of identical spellings for the minimal pairs would hinder learners' ability to learn the phonological forms of the new words.

4.3.1 Method

4.3.1.1 Participants

Fifty monolingual English speakers (42 females and 8 males) were recruited from the University of Alberta Department of Linguistics participant pool and received course credit for their participation. The mean age of the participants was 19.8 (*SD*=2.8, range= 17-34). A post-experiment questionnaire confirmed that none of the participants had any significant experience with languages that utilize length contrast phonemically or expressed any knowledge of the Arabic writing system. There were no reports of speech or hearing impairments.

4.3.1.2 Stimuli

The same pseudo-words and their pictured-meanings were used in this experiment. Instead of Roman spellings, the participants were presented with the Arabic written representations of these words. In order to avoid further confusion, short vowels were not presented orthographically—Arabic short vowels are optionally represented with diacritics. In Arabic, geminate consonants are orthographically marked with the addition of a diacritic placed above the letters. For instance, the auditory form /bannu/ is spelled in Arabic as <بنّو>, whereas /banu/ is spelled as <بنو>.

4.3.1.3 Training procedure

The training phase in this experiment was similar to the one described in 4.2.1.3. The only difference was the type of script presented to learners in both groups. For the Distinct Arabic Spelling group, the difference between the long and short consonants was indicated by adding a diacritic above the long consonant. In contrast, both long and short consonants were spelled identically for learners in the Same Roman Spelling group. For example, the auditory forms /bannu/ and /banu/ were spelled as <بنو> (Table 4.6 provides another example). Prior to the experiment, participants were instructed about the direction of the writing system. Participants were required to achieve 90% accuracy on the criterion test before they could advance to the final test.

TABLE 4.6: An example of the auditory and visual stimuli presented for each learning group.

Group	**Auditory item**	**Pictured-meaning**	**Spelling**
DAS	/kammu/		كمّو
SRS	/kammu/		كمو

4.3.1.4 Final test

The final test was identical to the one described in the previous experiment.

4.3.2 Results

On average, learners in the Distinct Roman Spelling group required 2.64 learning cycles to advance to the final test (*SD*= 0.84, range= 2-5), while it took learners in the Same Roman Spelling group an average of 3.15 learning cycles (*SD*= 1.75, range= 1-8) to reach 90% in the criterion test. As reported in the previous experiment, the average number of learning cycles for the Control group was 2.52 (SD= 1.62, range=1-6).

The distributions of accuracy scores for the both groups along with the Control group—the same group from Experiment One—on Matched and Mismatched pairs are shown in Figure 4.4. For the Matched items, all groups display higher performances compared to Mismatched items. The mean accuracy score for the Control, Distinct and Same Arabic Spelling groups is 0.88 (*SD*= 0.32), 0.90 (*SD*= 0.10) and 0.87 (*SD*=0.09), respectively. On Mismatched items, the distributions of both groups show higher ranges of accuracy and more variability than the distribution of the Control group. When comparing their distributions, both orthography groups exhibit slightly different patterns; the distribution of the accuracy scores for the Distinct Arabic Spelling group is denser than that of the Same Arabic Spelling group at the upper end of the distribution. In addition, the mean accuracy score of the Distinct Arabic Spelling group (*M*= 0.32, *SD*= 0.27) is slightly higher than both the Same Arabic Spelling group (*M*= 0.30, *SD*= 0.26) and the Control group (*M*= 0.28, *SD*= 0.19).

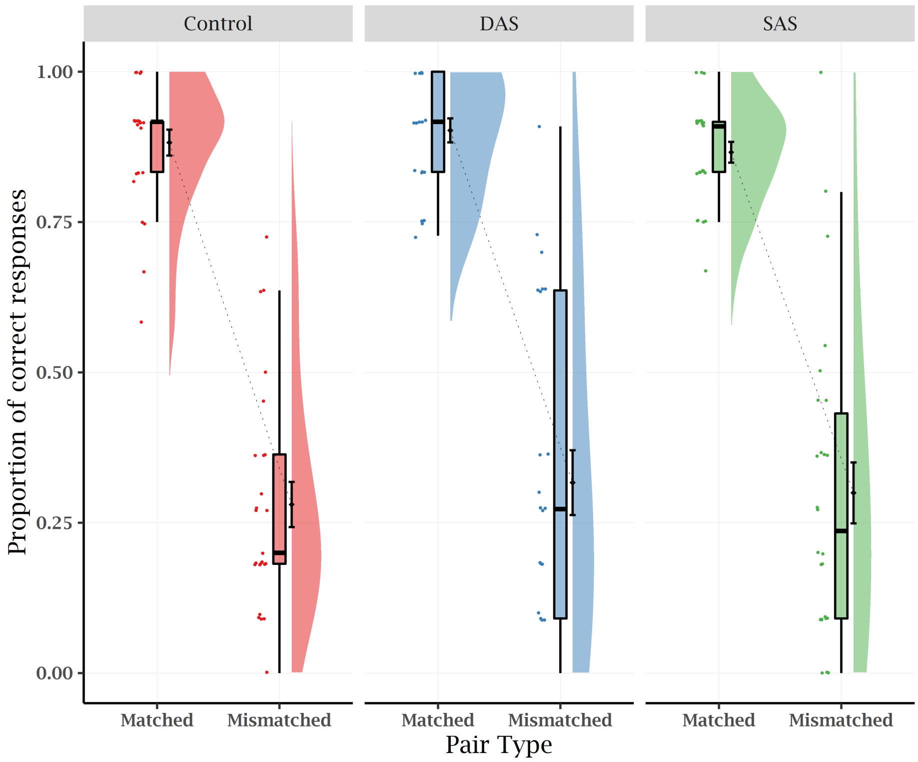


FIGURE 4.4: A combined plot representing the accuracy of the Control group, the Distinct Arabic Spelling group (DAS) and the Same Arabic Spelling group (SAS) on Matched and Mismatched items. The plot contains probability distributions (split-half violin); box-plots and means with standard errors for each group. Data from each individual are represented as points.

To examine whether the Arabic orthographic exposure was associated with the lexical encoding and the subsequent retrieval of words containing the length contrast, the responses to auditory-picture matching task for the three groups on Mismatched items [2] were analyzed using a mixed-effects logistic regression model. The final model contained GROUP (Control vs. Distinct Roman Spelling and Same Roman Spelling), CYCLE (the number of learning cycles) and their interaction as fixed effects. Including CYCLE and its interaction with GROUP in the model was determined by the likelihood ratio test which confirmed that the number of learning cycles impacted the performances of these groups differently. SUBJECT and ITEM served as random effects. In order to simplify the interpretation of the model coefficients, the value for CYCLE was mean-centered. Table 4.7 summarizes the model.

The model reveals no significant differences in accuracy between learners who were exposed to either compatible or incompatible Arabic script and the Control

[2] No significant difference was found between the three groups on Matched items.

TABLE 4.7: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy of participants on the auditory-picture matching task. The reference level for GROUP is Control. The value for CYCLE is mean-centered
.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-1.088	0.238	-4.578	0.000
DAS	0.220	0.336	0.653	0.514
SAS	-0.095	0.340	-0.278	0.781
Cycle	-0.044	0.147	-0.297	0.766
DAS:Cycle	0.748	0.311	2.408	**0.016**
SAS:Cycle	0.107	0.199	0.539	0.590

group. However, When inspecting the interaction between GROUP and CYCLE, the model revealed that the number of learning cycles was significantly associated with the accuracy scores for the Distinct Arabic Spelling group. A one unit increase in the number of learning cycles led to a significant increase in the likelihood of responding "NO" on Mismatched items compared to the Control group. That is, for every additional learning cycle, the odds of an accurate response for the Distinct Arabic Spelling group increased by a factor of 2.11 (CI 95%: 1.15-3.88) compared to those who received no orthographic input. An illustration of the interaction is shown in 4.5.

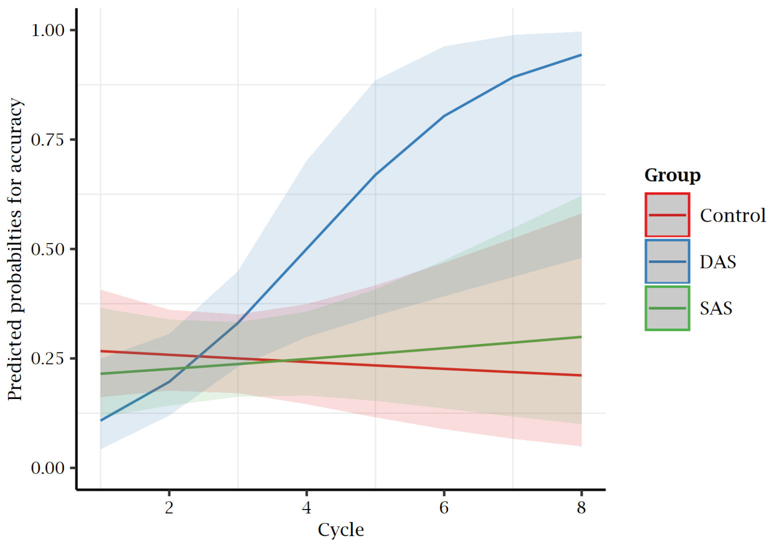


FIGURE 4.5: Predicted probabilities from the mixed-effects logistic results modelling the interaction between GROUP and CYCLE, as presented in Table 4.7. The probabilities are presented with 95% confidence interval.

4.3.3 Discussion

The aim of the current experiment was to determine if the exposure to unfamiliar orthographic input would have the same effect observed with Roman script in the previous experiment. The participants were trained on the same stimuli, except that the auditory forms were represented orthographically in Arabic. Based on the results obtained in Chapter 3, it was predicted that learners would not experience difficulty with the Arabic script and thus would use the available orthographic input to learn the new words. The findings of the current experiment paint a complex picture. On the face of it, the effects of Arabic script look to be insignificant in either aiding or hindering the lexical learning of the length contrast. The performances of both Arabic groups matched that of the Control group.

However, as the number of learning cycles increased, learners who were exposed to compatible Arabic spellings were more successful in establishing lexical representations for the length contrast. A possible explanation of this pattern points to an initial difficulty experienced by learners in processing the Arabic script, and with every additional learning cycle the orthographic regularities of the Arabic script became less complex. This way, learners who went through multiple learning cycles had more chance to associate the orthographic labels with the auditory forms. Further discussion is provided in the next section.

4.4 General discussion

This chapter investigated the role of orthographic information in the formation of nonnative phonemic representations at the lexical level. In line with the previous chapter, two experiments were designed to examine the effects of orthographic compatibility and familiarity on the lexical learning of consonant length contrast by monolingual English speakers.

The first research question asked whether access to compatible familiar orthography facilities the lexical encoding of the length contrast. The findings of Experiment One confirmed that the availability of distinct Roman orthographic labels for singletons and geminates (e.g., the auditory form /banu/ was spelled as <banu>, while /bannu/ was spelled as <bannu>) significantly improved the accuracy of learners in matching the auditory forms with their corresponding pictured-meanings. The improvement was measured in comparison to both learners who had no orthographic exposure and those who were presented with identical Roman spellings for short and long consonants. This suggests that lexical representations of nonnative speech sounds can be reinforced by familiar orthographic input. This finding mirrors those of the previous studies that have examined the enhancing effect of congruent familiar orthography on the lexical encoding and retrieval of nonnative speech sounds (Escudero et al., 2008; Escudero et al., 2014; Showalter, 2018).

More relevant to the current investigation, this outcome, however, is contrary to Mathieu (2014) who found that access to Roman spellings that signaled the presence of a short consonant with a single letter and a long consonant with double letters failed to lead to a more accurate lexical representations of the length contrast. This inconsistency is likely related to the differences in duration ratios used in both studies. In Mathieu (2014), the ratio between singleton and geminate consonants ranged from 2.2 to 3 (average= 2.5); while the consonant duration ratio used in the current experiments ranged from 2.7 to 3.9 (average= 3.4). Of more significance, the durations of the geminate consonants in the current experiments were longer by an average of 84 (ms). This slightly amplified temporal contrast—compared to Mathieu (2014)—between singleton and geminate consonants likely made the durational difference more salient and subsequently facilitated the establishment of distinct lexical representations for the length contrast in this study.

The second research question was whether the presence of incompatible Roman script would interfere with the lexical learning of the length contrast. In experiment one, a group of learners were exposed to identical Roman orthographic representations of both short and long consonants (both presented with single letters). The findings reveal, overall, no significant difference in the accuracy of encoding the contrast between these learners and learners who had no orthographic exposure, which indicates that incompatible Roman script exerts no effect on the lexical representations of the length contrast. This observed neutral effect of the orthographic forms may be due to their competition with more reliable visual cues (pictured meanings) that informed learners about the nature of the auditory forms. In the experiment, learners were tasked with matching the auditory forms with their corresponding meanings, and since the orthographic forms contained uninformative cues, learners chose to rely on them less. However, a sizable number of learners who were presented with identical Roman orthographic labels for the length contrast seems to exhibit a complete dependency on the orthographic cues during exposure. A closer look at the accuracy scores show that 7 participants failed to successfully recognize any of the Mismatched items compared to only one in the Control group. This may indicate that learners adopt different strategies when processing multiple cues. It is plausible that those who failed to register any correct response on the Mismatched items were mainly driven by the absence of distinctive orthographic labels for the length contrast.

The final research question in this chapter was whether a similar pattern would emerge if learners were exposed to unfamiliar script. Experiment two was designed to train learners on the same words using Arabic script. One group of learners were presented with distinct Arabic spellings of the length contrast in which long consonants were distinguished by the addition of a diacritic placed over letter (e.g., the auditory form /banu/ was spelled as <بنو>, whereas /bannu/ was spelled as <بنّو>). The other group was presented with identical Arabic spellings for both short and long consonants (both spelled without the diacritic). The pattern shown in this experiment is more complex than the one established in Experiment One. Overall, the findings suggest that learners ignored the orthographic information as evidenced by the similar accuracy scores achieved by both Arabic script groups in comparison to those who had no orthographic access. While this pattern answers the question straightforwardly, the interaction between the length of the exposure and learning group provides a more nuanced account of the effect of compatible Arabic script on learners' lexical representations of the length contrast. The results of the analysis indicated that in general learners exposed to distinct Arabic spellings for singleton and geminates failed to accurately encode the contrast; but unlike the other groups, the performance of the compatible Arabic orthography group significantly improved with every additional learning cycle.

There are two possible scenarios that could account for this significant association. First, it might be the case that learners attempted to utilize the unfamiliar orthographic cues from the onset of exposure, but initially experienced confusion when trying to decode the Arabic script and use it to learn the new words, subsequently failing to pass the criterion test. With more exposure, learners started to recognize the Arabic orthographic patterns better, including the diacritic that signaled the presence of long consonants. This added exposure eventually helped them encode the length contrast more accurately.

The second possibility is that, overall, learners ignored the Arabic orthographic cues and attempted to learn the novel words by associating the auditory forms with their corresponding pictured-meanings. Those who experienced persistent failure to advance to the final test were forced to resort to the orthographic cues in order to strengthen those associations. This might explain why no differences were detected between the two Arabic groups and the Control group when measuring the accuracy scores at the mean number of learning cycles. Those who required fewer learning cycles may have relied on the auditory forms and their pictured-meanings to perform the task rather than the orthographic representations of these words. This pattern has not been reported before since most of the data analysis conducted in previous studies has not considered the number of exposure cycles required as a predictor for the accuracy scores on the auditory-picture matching task. Prior research has shown that unfamiliar orthographic input in general does not exert any influence in the lexical learning of nonnative contrasts (Mathieu, 2014; Showalter & Hayes-Harb, 2015; Hayes-Harb & Hacking, 2015), and in some cases may in fact negatively interfere with learning (Mathieu, 2016). The current findings reveal that a subset of learners who had access to Arabic script displayed a significantly enhanced ability to encode the contrast lexically. This further suggests that, as shown in Chapter 3, the unfamiliar Arabic script poses no inherent difficulties or learning constraints for native English speakers. Thus, it is possible to think that having a basic prior knowledge about the Arabic orthographic system or increased exposure would exert more influence on learners' ability to establish accurate lexical representations for the length contrast than what have been observed in experiment two.

In conclusion, this chapter examined the effects of orthographic input on the acquisition of consonant length contrast at the lexical level. In line with the previous chapter, two orthographic variables were manipulated: Familiarity and compatibility. The findings suggest that familiar orthographic information facilitates the lexical encoding and the subsequent retrieval of words containing the length contrast, while the effects of incompatible orthographic input are mostly neutral and may negatively affect some learners. Further, unfamiliar orthographic input may require more exposure in order to enhance the lexical representations of the contrast.

Chapter 5

Conclusion

5.1 Introduction

Learning the phonology of a second language is a highly complex process. Different variables—both linguistic and non-linguistic—combine to influence the acquisition of nonnative speech sounds. This dissertation has examined the relative contribution of the orthographic input in the acquisition of nonnative speech sounds. Specifically, a systematic analysis of the interaction between orthographic information and nonnative speech sounds has been conducted to evaluate native English speakers' ability to perceive and lexically encode the length contrast. Multiple experiments were designed to address the following questions:

1. Does the exposure to orthographic cues bolster the distributional and lexical learning of the consonant length contrast?

2. If the orthographic information contradicts other distributional or lexical cues, does orthography override the effect of these cues in a similar way?

3. Does the level of familiarity with L2 orthography determine the learning outcome? If the target language employs an unfamiliar alphabet with diacritics, does the availability of diacritics neutralize or even reverse the negative influence of unfamiliar orthography in L2 phonemes?

The following section summarizes the main findings and attempts to answer the research questions posed in this dissertation.

5.2 Summary of the main findings

5.2.1 Orthographic influence at two different levels of representation

The findings in this dissertation suggest that orthographic input affects both the perceptual and lexical representations of the consonant length contrast. Chapter 3 demonstrates that the availability of Roman orthographic labels aided the learning of novel sound categories (a summary table of the results in Chapter 3 is provided in 5.1). The effect was evident in both the formation of a single category perception and the two-category length distinction. Regarding the former, as demonstrated in Chapter 2, the unimodal distribution alone was insufficient in guiding learners to perceive the singleton and geminate consonants as members of the same category. Rather, learners' performance indicated a high degree of discrimination between short and long consonants. Only when combining the unimodal training with matching Roman orthographic information was the discrimination of the length contrast significantly poorer, indicating a single length category had been established. For the latter, pairing the bimodal distribution with Roman orthographic input boosted the discrimination of the consonant length contrast. In this case, learners utilized both the distributional and orthographic information to robustly form two distinct length categories.

Chapter 4 reveals that access to Roman orthographic input enhances the lexical representation of the length contrast. In an auditory-picture matching task, learners who were exposed to auditory forms with their pictured meanings performed significantly worse than those who were additionally presented with the Roman spellings of these forms. Together, these results indicate that familiar orthographic

TABLE 5.1: A summary of the performance of the experimental groups in the perceptual learning task. Accuracy level is divided into four categories: Low, medium, high, very high. A group performance level was determined to be low if the accuracy is significantly lower than the Control and the Unimodal groups; and very high if the accuracy is significantly higher than the Control and the Bimodal groups.

Group	Distribution	Orth. Compatibility	Orth. Familiarity	Accuracy
Control	None	None	None	Medium
Unimodal	Unimodal	None	None	High
Bimodal	Bimodal	None	None	High
USRS	Unimodal	Compatible	Familiar	Low
BDRS	Bimodal	Compatible	Familiar	Very high
UDRS	Unimodal	Incompatible	Familiar	High
BSRS	Bimodal	Incompatible	Familiar	Medium
USAS	Unimodal	Compatible	Unfamiliar	Low
BDAS	Bimodal	Compatible	Unfamiliar	High
UDAS	Unimodal	Incompatible	Unfamiliar	High
BSAS	Bimodal	Incompatible	Unfamiliar	Medium

information facilitates the formation of accurate perceptual and lexical representations for the length contrast.

5.2.2 The influence of contradictory orthographic information in the perceptual and lexical representations of the length contrast

The results of this dissertation indicate that contradictory orthographic information affects the distributional and lexical learning of the consonant length contrast differently. The first experiment in Chapter 3 manipulated the consistency between the distributional information and Roman orthographic input, resulting in two conditions: One group was exposed to a unimodal distribution in which the target consonant in tokens 1-5 was spelled with a single letter and in tokens 6-10 was spelled with double letters, while the other group was exposed to a bimodal distribution represented orthographically with the same spelling. The conflict between the two sources of information was mainly resolved through orthographic input; learners established their perceptual categories based on whether the available written representations supported a single category or two-category length perception. Those who received a unimodal distribution paired with two distinct orthographic labels for the auditory tokens perceived the length contrast with higher accuracy, and those who were trained on a bimodal distribution represented with the a single orthographic label for the target consonant discriminated the length contrast with significantly less accuracy. However, the categories established via the misalignment between distributional information and orthographic input appeared to be slightly less robust than the ones created when both sources of information were aligned. For instance, in groups that were presented with orthography which denoted the existence of a single length category (i.e., both long and short consonants had the same spellings), only those who had the unimodal training exhibited significantly less accuracy in discriminating the length contrast than the Control group. Likewise, in groups that were exposed to orthography which indicated the existence of a length distinction (i.e., long and short consonants had two distinct Roman spellings), those who had the bimodal exposure showed a significantly higher accuracy in the discrimination of the length contrast than those who received the unimodal training (only on trained items). These results indicate that, overall, the influence of statistical information in guiding learners perception is relatively limited when competing with explicit information. The presence of orthographic cues seems to constrain the distributional learning of nonnative speech sounds.

This tendency for orthographic information to regulate distributional learning relates to the differential learning mechanisms that are dominant in infant and adult language acquisition. While infants construct their perceptual categories implicitly (Pierrehumbert, 2003; Bergmann, Tsuji, & Cristia, 2017), possibly by utilizing the available statistical information (Maye & Gerken, 2000; Maye et al., 2008; Yoshida et al., 2010; Cristià et al., 2011), adults seem to rely more on explicit information to establish nonnative sound categories (Archila-Suerte, Zevin, Bunta, & Hernandez, 2012). Unlike infants, adults' mature cognitive abilities allow the processing of high-level information during learning (White, Hutka, Williams, & Moreno, 2013). Thus, when the two types of cues interact, explicit information initiates the "registration of pattern recognizers for constructions" which are subsequently fine-tuned by implicit information (N. C. Ellis, 2005). In the context of the current results, orthographic cues are considered high-level/explicit information relative to the distributional cues. Recognizing orthographic patterns does not require extensive exposure the way recognition of distributional patterns does. In addition, learning the orthographic cues of the auditory forms is a conscious visual process in contrast to the unconscious nature of the distributional learning of these auditory forms. Learners exposed to orthographic information along with distributional cues derive their knowledge about the structure of the target contrast from the orthographic cues and, subsequently, proceed to utilize distributional information to refine the acquired categories. That is, orthographic cues appear to trigger the formation of L2 categories by determining whether a single or two separate categories should be established for the target stimuli. If the written representations provide no support for a length contrast, learners regard length duration as irrelevant within-category information despite what the distributional information conveys. However, if the orthographic information provides two distinct written representations for the length stimuli, learners encode the length contrast phonemically even if the most frequent auditory items represented by two distinct orthographic labels are perceptually indistinguishable. The dominance of orthographic cues in guiding perceptual learning does not entail that the role of distributional input is inconsequential in nonnative phonological development. As previously illustrated, compatible distributional and orthographic information leads to more robust and refined perceptual categories than incompatible distributional and orthographic input. This has been attested in both single and two-category perception. Thus, while orthographic cues may determine the number and the shape of the acquired perceptual categories, low-level matching distributional information continuously fine-tunes these categories.

This reliance on orthography to induce learning is particularly contrasted when looking at the effects of orthographic cues on the lexical learning of consonant length contrast. In Experiment 4.2 of Chapter 4, a group of learners was exposed to multiple words containing short and long consonants along with their pictured-meanings. Crucially, the orthographic representations of these words conveyed no useful information about the nature of the target contrast; both short and long consonants had the same spellings. The picture-auditory matching task revealed no significant difference between those who received such exposure and those who were presented with only the auditory forms and their pictured-meanings, indicating that incompatible orthography does not interfere with the lexical encoding of the length contrast—contrary to its influence in constraining distributional learning. A plausible explanation of this apparent lack of effect is that orthographic cues compete with high-level lexical information in informing the learners about the status of the length feature. Through minimal pair analysis, learners realize that the length feature is important in the target language because it cues meaning distinctions, and since the orthographic input does not distinguish between short and long consonants, they generally choose to ignore it.

However, there are still other alternative explanations for the discrepancy in the role of orthography at the perceptual and lexical levels. First, the two learning tasks put different demands on the working memory. In perceptual learning, subjects listened passively to tokens containing the length contrast, and their assessment involved judging whether two auditory items belong to one or two separate categories. The word learning task, on the other hand, involved actively learning to associate nonnative auditory forms with their meanings presented as visual referents, which requires higher cognitive resources.

Another possible cause for the differential influence of contradictory orthographic input in perceptual and lexical learning relates to the frequency with which orthographic cues were encountered in each task. In the perceptual learning, orthographic input was present in 400 trials, while the number of trials in which orthographic cues were shown ranged from 24 to 192 in the word learning task. This considerably lower exposure time, particularly for those who required fewer learning cycles to pass the criterion test, may have contributed to the lack of interference of the orthographic information with the lexical encoding of the consonant length contrast.

Finally, the difference in the visual prominence of the orthographic stimuli in both experiments could explain the limited effect of the contradictory orthography in the lexical learning task. In the distributional learning task, the only visual stimulus present was the orthographic input which occupied the center of the screen during the learning phase (see Figure 3.3 for an illustration). In the lexical learning task, however, the orthographic input occupied the bottom of the screen due to the presence of the pictured-meanings as shown in Figure 4.1. The less prominent position as well as the competition with real life visual referents could have constrained the influence of the the contradictory orthographic input in the lexical encoding of the length contrast.

5.2.3 Orthographic familiarity

A main research question of this dissertation concerns whether unfamiliar orthographic input has an adverse impact on learners' ability to perceive and lexically encode the consonant length contrast. The findings suggest otherwise. Arabic orthographic input paired with distributional cues produced a pattern similar to the familiar Roman orthographic input. The difference was one of degree: The impact of the unfamiliar Arabic input on learners' perceptual ability was less robust, especially in supporting a two-category length distinction. Learners who were exposed to bimodal distribution that was represented orthographically with matching Roman spellings were significantly better at discriminating the length contrast than those who received bimodal distribution only, a feat that was not achieved with

Arabic orthography. In other cases, the effects of Arabic spellings matched the effects of Roman written cues on learners' perception of the length contrast. For instance, presenting learners with a bimodal distribution paired with incompatible Arabic spelling produced the same learning outcome obtained with incompatible Roman spellings: Learners inferred a single length category. In lexical learning, the impact of unfamiliar orthography on learners' ability to establish accurate lexical representations is not as straightforward. Learners may require more exposure to the unfamiliar orthography in order to utilize it when learning nonnative speech sounds. Experiment Two of Chapter 4 (4.4) investigated whether Arabic orthography would facilitate learning the phonological forms of novel words containing the length distinction. Overall, the results showed no improvement in the lexical encoding and retrieval of the contrast compared to the Control group. However, there was a significant association between the length of exposure and the accuracy scores: Learners who had an extended exposure performed significantly better at matching auditory forms containing the length contrast with their correct pictured-meanings. These findings suggest that unfamiliar orthography in general poses no inherent difficulties during the course of phonological development. Learners are expected to overcome the initial confusion and rapidly learn to map the unfamiliar orthographic labels with their phonemic referents. This result is contrary to earlier studies that reported no improvement gained by the exposure to Arabic orthography in the lexical encoding of nonnative speech sounds (Showalter & Hayes-Harb, 2015; Mathieu, 2016). A possible confounding factor in these studies was the relative difficulty of the nonnative contrasts (/k/-/q/ in Showalter and Hayes-Harb, 2015 and /ħ/-/χ/ in Mathieu, 2016), which might have rendered the access to the Arabic spellings less helpful. One additional point to note in these studies is the lack of measure of the potential association between the length of exposure —the number of learning cycles required to pass the criterion test—and the accuracy scores, which might have masked any advantage for the increased exposure in learning the phonological forms of new words.

5.2.4 Individual differences

A major characteristic of the results in this dissertation was the considerable amount of variation observed among learners both in the perceptual and lexical learning of the consonant length contrast. Adult learners vary greatly in their ability to successfully acquire a second language due to multiple factors including cognitive abilities (Darcy, Park, & Yang, 2015; Lev-Ari & Peperkamp, 2013) (see R. Ellis, 2004, for a general review).

Of relevance to the current research is the variability in performance observed among the experimental groups that were exposed to contradictory orthographic input. Specifically, different patterns emerged mirroring the ways learners utilized the available cues to resolve the learning difficulty. In the perceptual learning task, those who were trained on the bimodal distribution of the consonant length contrast and simultaneously were presented with Roman written representations indicative of a single length category exhibited more variability in their accuracy scores. Table 5.2 shows the percentages of learners in each group whose accuracy scores were either very low (0-0.25), low (0.26-0.5), intermediate (0.51-0.75) or high (0-76-1). As can be noted, 52.4% of the learners in the BSRS group (Bimodal Incompatible Roman Orthography) showed very low discrimination of the length contrast (57.2% scored below or at chance level), while 38% achieved high accuracy scores in their perception of the contrast (42.8% scored above chance level).

This wide spread of accuracy scores in which the majority of learners exhibited either very low or high accuracy is generally not reflected in other groups. For instance, the majority of learners who were exposed to the same orthography coupled instead with a unimodal distribution performed at or below chance level (80%). Thus, it is reasonable to assume that the variability of the accuracy scores in the BSRS group is specifically related to learners' interaction with the language input rather than other sources of inter-individual differences. Despite the contradictory

TABLE 5.2: Distribution of the percentages of learners in each groups across four accuracy intervals: Very low (0-0.25), Low (0.26-0.5), Intermediate (0.51-0.75) and High (0.76-1).

Group	**Very low** (0-0.25)	**Low** (0.26-0.5)	**Intermediate** (0.51-0.75)	**High** (0.76-1)
Control	44%	17%	2.4%	36.6%
Unimodal	15%	20%	15%	50%
Bimodal	15%	0%	15%	70%
USRS	60%	20%	5%	15%
BDRS	0%	10%	0%	90%
UDRS	0%	0%	15%	85%
BSRS	52.4%	4.8%	4.8%	38%
USAS	55%	20%	10%	15%
BDAS	0%	14.3%	9.5%	76.2%
UDAS	4.8%	4.8%	23.8%	66.6%
BSAS	60%	20%	0%	20%

orthographic cues, a subset of learners exhibited sensitivity to the distributional information, leading to more accurate discrimination of the length contrast. The other subset derived their knowledge about the status of the auditory stimuli exclusively from the orthographic cues, undermining the influence of distributional information. This level of variance in the accuracy scores provides evidence that learners respond differently to contradictory input; while for the majority of learners high level information determines the learning outcome, low level forms can influence or even shape the learned categories for a considerable number of learners.

Although it is beyond the scope of this dissertation to precisely describe the cognitive mechanisms underlying the observed pattern, it is worth revisiting the explanation offered in the previous section about the interaction between written cues and distributional information. In general, adults' mature cognitive abilities entail that high level information takes precedence over lower level cues (White et al., 2013); however, as the results suggest, not all learners exhibit the same learning dynamic. Specifically, the extent to which orthographic information shapes distributional input is different for these learners. That is, distributional learning seems to drive the establishment of two distinct categories for the length contrast, thereby blocking orthographic input that would typically lead to a single length category perception. This observation indicates that learners exposed to the same input may fundamentally differ in the way they process conflicting information. The same learning mechanism infants employ when learning their first language may also be employed by adults learning nonnative speech sounds. This is particularly interesting given that these learners have simultaneous access to orthographic cues which offer a more straightforward pathway to learning the target speech sounds. It can be thus suggested that while high-level information dominates learning in nonnative phonological development, for some learners, distributional learning still operates unaffected by higher level information such as orthographic cues. The co-existence of both learning mechanisms supports the view that adult learners utilize both high-level and low-level information when exposed to varied input, with some learners showing propensity to form nonnative categories based on low level cues (White et al., 2013).

Interestingly, the same degree of variation among learners exposed to contradictory input has not been attested in groups that were exposed to unimodal distribution along with two distinct orthographic labels for the length contrast. This discrepancy could be attributed to learners' weak sensitivity to the unimodal distribution as evidenced by their high accuracy on the discrimination task following training. The only exposure that resulted in a somewhat similar pattern was the bimodal distribution represented orthographically with a uniform Arabic spelling. While the majority of learners (60%) scored very low on the discrimination task indicating that they based their perception on the orthography, a small subset of learners (20%) displayed high accuracy.

For the lexical learning task, the analysis of the individual differences indicates a similar pattern; learners who were exposed to conflicting orthographic input displayed variability in the lexical encoding and retrieval of the length contrast. In particular, although learners in the Control group and those who were presented with contradictory Roman orthography and auditory forms along with their pictured meanings showed similar accuracy scores, 7 learners in the contradictory orthography group failed to lexically encode any test pair compared to only one in the Control group. This comparatively high number cannot be attributed only to individual differences unrelated to the exposure. Clearly, the inconsistent orthographic information negatively interfered with learners' ability to establish any lexical representation for the length contrast. For these learners, orthographic cues were given prominence in determining what auditory segments were encoded in the lexicon. However, this pattern has not been replicated in groups where the exposure contained Arabic spellings. Both groups with consistent and inconsistent spellings produced the same numbers of learners who failed to establish any lexical representation for the length contrast (3), which suggests that this particular pattern may have emerged as a result of the exposure to unfamiliar orthography in general rather than whether the orthography was consistent or inconsistent with the auditory forms. The only difference was that increased exposure was positively associated with accuracy for learners who received consistent Arabic spellings, while there was no association either positive or negative between increased exposure and accuracy for those who were presented with inconsistent Arabic spellings.

In sum, the patterns of individual differences described in this section indicate that learners respond differently to orthographic stimuli, with some exhibiting a preference for learning from written cues over auditory, distributional or lexical ones. It should be stressed, however, that these emerging patterns of individual variation represent an early stage of phonological development. It remains unclear whether learners who show bias towards orthographic information, even when it leads to faulty representation, adapt to the demands of the target input and adjust their learning mechanism accordingly.

5.3 Implications for current models of L2 phonology

Despite the strong empirical evidence, there are no models yet that consider—either implicitly or explicitly—orthographic representations as important variables in L2 phonological development. This dissertation and previous research have demonstrated that orthographic information influences almost all aspects of adult L2 phonology including speech perception (Escudero & Wanrooij, 2010; Simon et al., 2010; Pytlyk, 2011; Mok et al., 2018), production (Rafat, 2015; Bassetti & Atkinson, 2015; Young-Scholten & Langer, 2015; Hayes-Harb et al., 2018; Han & Kim, 2017; Bassetti, 2017; Bassetti et al., 2018), lexical encoding (Hayes-Harb et al., 2010; Simon et al., 2010; Showalter & Hayes-Harb, 2013; Escudero et al., 2014; Showalter & Hayes-Harb, 2015; Hayes-Harb & Hacking, 2015; Mathieu, 2016; Hayes-Harb & Cheng, 2016; Simonchyk & Darcy, 2018; Showalter, 2018) and L2 spoken word recognition (Veivo & Järvikivi, 2013; Veivo et al., 2016; Qu et al., 2018; Veivo et al., 2018). The most commonly quoted models of second language phonology, the Speech Learning Model (SLM) (Flege, 1995; Flege, Schirru, & MacKay, 2003) and the Perceptual Assimilation Model (PAM and PAM2) (Best, 1995; Best & Tyler, 2007), are predominately concerned with predicting the relative difficulty of perceiving and producing nonnative speech sounds based on the differences between the sound systems of learners' L1 and the target language. Although the research in this area has contributed greatly to the understanding of the perceptual and learning problems facing L2 learners, the predictive power of these models cannot account for learning difficulties related to the orthographic input. First, as has been demonstrated in this dissertation and previous research, the exposure to orthographic information can drastically alter the shape of the perceptual categories and may interfere with the lexical representations of nonnative speech sounds. In addition, the mismatch between the grapheme-phoneme correspondences of L1 and L2 in languages with shared Roman orthography often leads to persistent errors in producing L2 sounds (Young-Scholten & Langer, 2015), or the failure to encode nonnative speech sounds in the lexicon (Hayes-Harb et al., 2010; Escudero et al., 2014). Finally, it has been shown that L2 learners whose native languages employ shallow orthographies—where there is one-to-one grapheme-phoneme correspondences—experience difficulties if the relationship between the orthographic and phonemic systems of the target language is inconsistent. This may result in persistent production errors even among advanced learners (Bassetti & Atkinson, 2015; Bassetti, 2017; Bassetti et al., 2018).

These findings warrant a reexamination of the current L2 models to account for the role of orthography in second language phonology. A straightforward way to incorporate orthographic input is to assume that along with the phonological system of the native language, learners arrive to the task of learning a second language with their native language grapheme-phoneme correspondences. If L1 and L2 share the same writing system, a mismatch between the orthographic representations of the native and the target languages is predicted to result in either a delayed learning or even persistent difficulty that does not resolve with increased language experience. Another prediction that can be generated through the integration of orthography as a variable in these models states that the discrepancy in the level of orthographic depth between L1 and L2 may interfere with the acquisition of the target phonology. This prediction is partially supported by empirical evidence, particularly in production. For instance, Italian and Spanish learners of English have been shown to commit production errors reflective of the transparent nature of the orthography of these languages as opposed to English (Bassetti & Atkinson, 2015; Vokic, 2011). It is not clear yet whether the reverse situation, one in which the native language utilizes a deep orthography and the target language employs a transparent one, would induce similar learning difficulties when acquiring L2 speech sounds. The only available evidence suggests that when exposed to transparent orthography, learners whose L1 has consistent grapheme-phoneme correspondences commit fewer errors than learners whose native language has inconsistent orthographic representations (Erdener & Burnham, 2005). Another part of the prediction that requires further testing is whether the degree of orthographic depth also applies to unfamiliar orthography. This describes a case where both the native and the target languages differ in the level of orthographic familiarity as well as orthographic depth. This prediction is motivated by the assumption that learners' expectations about the target orthography are influenced by the type of relationship between L1 graphemes and phonemes.

Beyond predicting the relative influence of L1 phonology in the acquisition of L2 speech sounds, neither SLM nor PAM2 provides a comprehensive theory that explains the underlying learning mechanisms guiding L2 phonological development. In addition, both models were originally intended to account for phonological acquisition in language immersion settings where other modes of learning are less utilized (for application of these models in instructional settings see Piske (2007) for SLM, and Tyler (2019) for PAM2). Therefore, the written representation of spoken language has not been considered an essential factor in L2 phonological developments. This dissertation has explored the influence of orthographic input in the perceptual and lexical learning of nonnative speech sounds. The evidence suggests that learning nonnative speech categories can be achieved implicitly through attending to the distributional properties available in the input. For some learners, the sensitivity to the distributional information is not affected by the availability of orthographic cues, but for the majority, the learned perceptual categories are shaped by the orthographic stimuli. In lexical learning, encoding and subsequent retrieval of words containing nonnative speech sounds in the lexicon is significantly enhanced by the availability of transparent orthographic input. When the target orthographic system lacks one-to-one grapheme-phoneme correspondences, a subset of learners may completely rely on the orthography and subsequently fail to learn the nonnative phonological forms of new words. The strong orthographic effect observed in perceptual learning is possibly due to the dominant role of top-down/explicit learning mechanism in adult language acquisition as a result of brain maturation, however; the analysis of individual differences clearly reveals that some learners still exhibit sensitivity to the distributional information in the presence of higher order cues.

In lexical learning, mapping between high level lexical information and low level auditory forms proceeds in a top-down manner that can be facilitated by matching orthographic cues. Mismatching orthographic information does not induce the the same effect observed in perceptual learning due to its competition with higher-level lexical cues—although a subset of learners may exhibit orthographic bias and subsequently fail to encode nonnative contrasts lexically. These learning patterns could help explain some of the individual variation that is present in all stages of L2 phonological development. As has been shown, learners respond differently to varied input, particularly when multiple sources of information convey contradictory message.

Modeling L2 phonological development requires not only considering the effects of L1 transfer on learners' ability to perceive and produce nonnative speech sounds, rather a comprehensive model should account for the ways multiple sources of information, including orthography, interact to influence learners' ability to acquire the target phonology. An ideal model should be equally able to explain any systematic individual differences observed among L2 learners instead of regarding variation as an afterthought that does not warrant further scrutiny.

5.4 Limitations and future directions

The findings in this dissertation provide strong evidence for orthographic effect both at the perceptual and lexical level. However, caution should be exercised before generalizing these findings to other contexts as they offer a specific account of the interaction between auditory and visual cues under a highly controlled environment which may not fully reflect the complex nature of L2 phonological acquisition. As with all experimental research, there are a number of limitations that should be noted. First, designing short learning experiments often requires using relatively less complex training stimuli. As demonstrated by the earlier versions of the distributional learning experiments, exposing learners to difficult stimuli may fail to yield any rapid learning effect. Second, in the lexical learning task, the position of the orthographic stimuli was not consistent with the distributional learning task (for an illustration see Figures 3.3 and 4.1). In the latter, the spelling of the auditory form was positioned in learners' line of sight in the center of the screen, and since the pictured-meaning in the lexical learning task occupied the center, the spelling was positioned at the bottom of the screen. The difference here is not trivial insofar as it may partially explain why orthographic effects were more prominent in distributional learning than in lexical learning. In fact, some participants reported that they focused less on the spellings of the auditory forms during the word learning task. Finally, learning the meanings of new words is not limited to picture-word association; instead L2 learners utilize multiple ways to derive the meanings of novel words, including translation from L2 to L1, learning from context, dictionary definitions, etc. It remains unclear whether the observed orthographic effect in the lexical learning experiment can be generalized to other ways of word learning.

Some of these limitations can be addressed in future research. Specifically, in the word-learning task, in order to give the same emphasis to orthographic cues, the presentation of the spellings of the auditory forms should be in learners' line of sight. This can be accomplished by placing the spelling directly below or above the pictured-meaning, or by displaying the two visual cues on different slides. Future studies should also explore the effects of orthography on the lexical encoding of nonnative speech sounds through other means of word learning. It is safe to assume that pairing orthographic cues with non-visual cues of novel words' meanings may affect the lexical encoding of nonnative speech sounds differently. In such case, learners are predicted to attend to the orthographic input more simply because visual cues are more salient. Expanding the scope of the inquiry to include other methods of word learning will further our understanding of the role of orthography in nonnative speech learning.

Another important avenue for future research is understanding how orthographic representations change over time. The present dissertation tested orthographic effects at the onset of L2 phonological development. It is expected that, with increased exposure, some of the negative effects of the over-reliance on orthographic cues will be resolved. Therefore, future research will attempt to examine under which conditions negative orthographic interference subsides or persists thorough L2 phonological development. Further research is also needed to find which aspect of L2 phonology is more susceptible to orthographic effects. Experimental studies on advanced L2 learners have revealed persistent orthographic errors in production (Bassetti & Atkinson, 2015; Bassetti, 2017; Bassetti et al., 2018). It is unclear whether experienced learners exhibit a great degree of orthographic interference only in production, or whether the negative influence extends to L2 perception, lexical representations and word recognition.

One crucial area that requires further investigation is understanding the basis of the observed individual variation. As has been demonstrated, learners vary greatly in their utilization of the orthographic input, particularly when it is paired with conflicting auditory input. The explanation offered here for the observed patterns suggests that learners exhibit different degrees of dependence on the learning mechanisms implemented to process the target input. Whereas some learners apply a top-down learning through mapping the orthographic cues onto the auditory forms, a subset of learners utilize low-level information such as distributional properties to establish nonnative speech sounds, thus bypassing the available high-level information (orthographic input). In matching input, optimal learning condition requires utilizing both high-level and low-level information. Future research will examine if there are any neural correlates for the observed patterns of individual variation. Functional magnetic resonance imaging (fMRI) offers a neuroimaging technique that records brain activity when a subject is engaging in an experimental task. Along with the behavioral data, neural images of the brain will advance our understanding of the underlying cognitive processes involved in the acquisition of nonnative speech sounds. More importantly, the data will reveal if there is any neural basis for the individual differences that occur in response to conflicting input.

5.5 Conclusion

This dissertation has highlighted the potential role of orthographic representations in L2 phonology. The analysis conducted here has provided a nuanced view of the interaction between written cues and auditory forms both in the distributional and the lexical learning of the consonant length contrast. The findings indicate that orthographic input, irrespective of its level of familiarity or depth, considerably influences the perception and the lexical encoding of the length contrast. This influence, however, varies on the individual level as a function of the nature of the input. Overall, the results provide original contributions to the body of work on the interaction between orthography and phonology, as well as offer theoretical implications to the models of L2 phonology and practical applications in foreign language teaching. It is hoped that some of the intriguing findings in this dissertation will continue to inspire more research on the intricate influence of orthographic information in the development of L2 phonology.

References

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. (2018). Raincloud plots: A multi-platform tool for robust data visualization. *PeerJ Preprints*, *6*, e27137v1.

Amengual, M. (2016). The perception of language-specific phonetic categories does not guarantee accurate phonological representations in the lexicon of early bilinguals. *Applied Psycholinguistics*, *37*(5), 1221–1251.

Archila-Suerte, P., Zevin, J., Bunta, F., & Hernandez, A. E. (2012). Age of acquisition and proficiency in a second language independently influence the perception of non-native speech. *Bilingualism: Language and Cognition*, *15*(1), 190–201.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Baese-Berk, M. M. (2010). *An examination of the relationship between speech perception and production* (Doctoral dissertation, Northwestern University).

Bassetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Language and Literacy*, *9*(1), 95–114.

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.

Bassetti, B. & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics*, *36*(01), 67–91.

Bassetti, B., Soković-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.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, Articles*, *67*(1), 1–48.

Bergmann, C., Tsuji, S., & Cristia, A. (2017). Top-down versus bottom-up theories of phonological acquisition: A big data approach. In *Interspeech 2017* (pp. 2103–2107).

Best, C. T. (1995). A direct-realist view of cross-language speech perception. *Speech perception and linguistic experience: Theoretical and methodological issues*, 171–204.

Best, C. T. & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. *Language experience in second language speech learning: In honor of James Emil Flege*, 13–34.

Boersma, P. & Weenink, D. (2018). Praat: Doing phonetics by computer. Retrieved from http://www.praat.org/

Brewer, J. (2008). *Phonetic Reflexes of Orthographic Characteristics in Lexical Representation* (Doctoral dissertation, University of Arizona).

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of standardized stimuli (BOSS) Phase II: 930 new normative photos. *PLOS ONE*, *9*, 1–10.

Brown, C. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. *Second language acquisition and linguistic theory*, *1*, 4–64.

Cowan, N. & Morse, P. A. (1986). The use of auditory and phonetic memory in vowel discrimination. *The Journal of the Acoustical Society of America*, *79*(2), 500–507.

Cristià, A., McGuire, G. L., Seidl, A., & Francis, A. L. (2011). Effects of the distribution of acoustic cues on infants' perception of sibilants. *Journal of Phonetics*, *39*(3), 388–402.

Crowder, R. G. (1982). Decay of auditory memory in vowel discrimination. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(2), 153–162.

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*, *34*(2), 269–284.

Darcy, I., Dekydtspotter, L., Sprouse, R. A., Glover, J., Kaden, C., McGuire, M., & Scott, J. H. (2012). Direct mapping of acoustics to phonology: On the lexical encoding of front rounded vowels in L1 English– L2 French acquisition. *Second Language Research*, *28*(1), 5–40.

Darcy, I., Park, H., & Yang, C. L. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences*, *40*, 63–72.

Díaz, B., Mitterer, H., Broersma, M., & Sebastián-Gallés, N. (2012). Individual differences in late bilinguals' L2 phonological processes: From acoustic-phonetic analysis to lexical access. *Learning and Individual Differences*, *22*(6), 680–689.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition*, *27*(2), 305–352.

Ellis, R. (2004). Individual differences in second language learning. *The handbook of applied linguistics*, 525–551.

Erdener, V. D. & Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning*, *55*(2), 191–228.

Escudero, P. (2015). Orthography plays a limited role when learning the phonological forms of new words: The case of Spanish and English learners of novel Dutch words. *Applied Psycholinguistics*, *36*(01), 7–22.

Escudero, P., Benders, T., & Wanrooij, K. (2011). Enhanced bimodal distributions facilitate the learning of second language vowels. *The Journal of the Acoustical Society of America*, *130*(4), 206–212.

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*, 384–395.

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

Escudero, P. & Williams, D. (2014). Distributional learning has immediate and long-lasting effects. *Cognition*, *133*(2), 408–413.

Flege, J. E. (1988). The production and perception of foreign language speech sounds. *Human communication and its disorders: A review*, 224–401.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Theoretical and methodological issues*, 233–277.

Flege, J. E., Schirru, C., & MacKay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech communication*, *40*(4), 467–491.

Gulian, M., Escudero, P., & Boersma, P. (2007). Supervision hampers distributional learning of vowel contrasts. In *Proceedings of the 16th international congress of phonetic sciences* (Vol. 2, pp. 1893–1896). Saarbrücken, Germany.

Han, J. I. & Kim, J. Y. (2017). The Influence of Orthography on the Production of Alphabetic, Second-Language Allophones by Speakers of a Non-alphabetic Language. *Journal of Psycholinguistic Research*.

Hancin-Bhatt, B. (1994). Segment transfer: a consequence of a dynamic system. *Second Language Research*, *10*(3), 241–269.

Hayes, R. L. (2001). The perception of novel phoneme contrasts in a second language: A developmental study of native speakers of English learning Japanese singleton and geminate consonant contrasts. *Coyote Papers: Working Papers in Linguistics, Language in Cognitive Science*.

Hayes, R. L. (2003). *How are second language phoneme contrasts learned?* (Doctoral dissertation, The University of Arizona.).

Hayes-Harb, R. (2005). Optimal L2 speech perception: Native speakers of English and Japanese consonant length contrasts. *Journal of Language and Linguistics*, *4*(1), 1–29.

Hayes-Harb, R. (2007). Lexical and statistical evidence in the acquisition of second language phonemes. *Second Language Research*, *23*(1), 65–94.

Hayes-Harb, R., Brown, K., & Smith, B. L. (2018). Orthographic input and the acquisition of German final devoicing by native speakers of English. *Language and Speech*, *61*(4), 547–564.

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. & Hacking, J. F. (2015). The influence of written stress marks on native English speakers' acquisition of Russian lexical stress contrasts. *Slavic & East European Journal*, *36*(1), 185–200.

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

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

Hirata, Y. (2004). Training native English speakers to perceive Japanese length contrasts in word versus sentence contexts. *The Journal of the Acoustical Society of America*, *116*(4), 2384–2394.

Hirata, Y., Whitehurst, E., & Cullings, E. (2007). Training native English speakers to identify Japanese vowel length contrast with sentences at varied speaking rates. *The Journal of the Acoustical Society of America*, *121*(6), 3837–3845.

Hisagi, M. & Strange, W. (2011). Perception of Japanese temporally-cued contrasts by American English listeners. *Language and Speech*, *54*(2), 241–264.

Idemaru, K. & Guion, S. G. (2008). Acoustic covariants of length contrast in Japanese stops. *Journal of the International Phonetic Association*, *38*(2), 167–186.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language*, *59*(4), 434–446.

Kato, H. & Tajima, K. (2002). Native and non-native perception of phonemic length contrasts in Japanese: A categorization study. *The Journal of the Acoustical Society of America*, *112*(5), 2387–2387.

Lado, R. (1957). *Linguistics across cultures. Applied linguistics for language teachers*. University of Michigan Press.

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.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the acoustical society of America*, *94*, 1242–1255.

Mathieu, L. (2014). *The influence of unfamiliar orthography on L2 phonolexical acquisition* (Doctoral dissertation, The University of Arizona).

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

Maye, J. & Gerken, L. (2000). Learning phonemes without minimal pairs. In *Proceedings of the 24th Annual Boston University Conference on Language Development* (Vol. 2, pp. 522–533). Somerville, MA: Cascadilla Press.

Maye, J. & Gerken, L. (2001). Learning phonemes: how far can the input take us. In *Proceedings of the 25th Annual Boston University Conference on Language Development* (Vol. 1, pp. 480–490). Somerville, MA: Cascadilla Press.

Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science*, *11*(1), 122–134.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3).

Mok, P. P. K., Lee, A., Li, J. J., & Xu, R. B. (2018). Orthographic effects on the perception and production of L2 Mandarin tones. *Speech Communication*, *101*, 1–10.

Morrison, G. S. (2007). Logistic regression modelling for first and second language perception data. *Amsterdam studies in the theory and history of linguistic science series IV*, *282*, 219–236.

Motohashi-Saigo, M. & Hardison, D. M. (2009). Acquisition of L2 Japanese geminates: Training with waveform displays. *Language Learning & Technology*, *13*(2), 29–47.

Okuno, T. (2013). *Acquisition of L2 vowel duration in Japanese by native English speakers* (Doctoral dissertation, Michigan State University).

Ong, J. H., Burnham, D., & Escudero, P. (2015). Distributional learning of lexical tones: A comparison of attended vs. unattended listening. *PLoS ONE*, *10*(7).

Ong, J. H., Burnham, D., Escudero, P., & Stevens, C. J. (2017). Effect of linguistic and musical experience on distributional learning of nonnative lexical tones. *Journal of Speech, Language, and Hearing Research*, *60*(10), 2769–2780.

Ong, J. H., Terry, J., & Escudero, P. (2016). Can Australian English listeners learn non-native vowels via distributional learning? In *Proceedings of the 16th Australasian International Conference on Speech Science and Technology* (pp. 289–292). Parramatta, Australia.

Pajak, B. (2012). *Inductive inference in non-native speech processing and learning* (Doctoral dissertation, University of California, San Diego).

Pajak, B. & Levy, R. (2011). Phonological generalization from distributional evidence. In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2673–2678).

Peperkamp, S., Pettinato, M., & Dupoux, E. (2003). Allophonic variation and the acquisition of phoneme categories. In *Proceedings of the 27th Annual Boston University Conference in Language Development* (Vol. 2, pp. 650–661). Sommerville, MA: Cascadilla Press.

Perfors, A. & Dunbar, D. (2010). Phonetic training makes word learning easier. In *Cognition in flux: Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (Vol. 32, pp. 1613–1618).

Pickett, J. M. & Decker, L. R. (1960). Time factors in perception of a double consonant. *Language and Speech*, *3*(1), 11–17.

Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and speech*, *46*(2-3), 115–154.

Piske, T. (2007). Implications of James E. Flege's research for the foreign language classroom. *Language experience in second language speech learning. In honor of James Emil Flege*, 301–314.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Attention, Perception, & Psychophysics*, *13*(2), 253–260.

Porretta, V. J. & Tucker, B. V. (2015). Perception of non-native consonant length contrast: The role of attention in phonetic processing. *Second Language Research*, *31*(2), 239–265.

Psychology Software Tools, I. (2012). E-Prime 2.0. Retrieved from http://www.pstnet.com

Pytlyk, C. (2011). Shared orthography: Do shared written symbols influence the perception of L2 sounds? *Modern Language Journal*, *95*(4), 541–557.

Qu, Q., Cui, Z., & Damian, M. F. (2018). Orthographic effects in second-language spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 1325–1332.

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria. Retrieved from https://www.r-project.org

Rafat, Y. (2015). The interaction of acoustic and orthographic input in the acquisition of Spanish assibilated/fricative rhotics. *Applied Psycholinguistics*, *36*(01), 43–66.

Sebastian-Galles, N. & Baus, C. (2005). On the relationship between perception and production in L2 categories. *Twenty-first century psycholinguistics: Four cornerstones*, 279–292.

Shea, C. & Curtin, S. (2005). Learning allophones from the input. *University of Calgary*.

Sheldon, A. & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, *3*, 243–261.

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

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.

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*(01), 23–42.

Simon, E., Chambless, D., & Kickhöfel Alves, U. (2010). Understanding the role of orthography in the acquisition of a non-native vowel contrast. *Language Sciences*, *32*(3), 380–394.

Simonchyk, A. & Darcy, I. (2018). The effect of orthography on the lexical encoding of palatalized consonants in L2 Russian. *Language and speech*, *61*(4), 522–546.

Tajima, K., Kato, H., Rothwell, A., Akahane-Yamada, R., & Munhall, K. G. (2008). Training English listeners to perceive phonemic length contrasts in Japanese. *The Journal of the Acoustical Society of America*, *123*(1), 397–413.

Takeyasu, H. & Giriko, M. (2017). Effects of duration and phonological length of the preceding/following segments on perception of the length contrast in Japanese.

In *The phonetics and phonology of geminate consonants*. Oxford: Oxford University Press.

Al-Tamimi, J. & Khattab, G. (2015). Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: The case of fricative consonants. *The Journal of the Acoustical Society of America*, *138*(1), 344–360.

Terry, J., Ong, J. H., & Escudero, P. (2015). Passive distributional learning of non-native vowel contrasts does not work for all listeners. In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*. Glasgow, Scotland.

Tsukada, K., Cox, F., & Hajek, J. (2014). Cross-language perception of Japanese singleton and geminate consonants: Preliminary data from non-native learners of Japanese and native speakers of Italian and Australian English. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH* (pp. 1288–1292).

Turk, A., Nakai, S., & Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide. *Methods in empirical prosody research*, *3*, 1–28.

Tyler, M. D. (2019). PAM-L2 and phonological category acquisition in the foreign language classroom. *A Sound Approach to Language Matters: In Honor of Ocke-Schwen Bohn*, 607–630.

Van Hessen, A. J. & Schouten, M. E. H. (1992). Modeling phoneme perception. II: A model of stop consonant discrimination. *The Journal of the Acoustical Society of America*, *92*(4), 1856–1868.

Veivo, O. & Järvikivi, J. (2013). Proficiency modulates early orthographic and phonological processing in L2 spoken word recognition. *Bilingualism: Language and Cognition*, *16*(4), 864–883.

Veivo, O., Järvikivi, J., Porretta, V., & Hyönä, J. (2016). Orthographic activation in L2 spoken word recognition depends on proficiency: Evidence from eye-tracking. *Frontiers in psychology*, *7*, 1120.

Veivo, O., Porretta, V., Hyönä, J., & Järvikivi, J. (2018). Spoken second language words activate native language orthographic information in late second language learners. *Applied Psycholinguistics*, 1–22.

Vokic, G. (2011). When alphabets collide: Alphabetic first-language speakers' approach to speech production in an alphabetic second language. *Second Language Research*, *27*(3), 391–417.

Wanrooij, K. & Boersma, P. (2013). Distributional training of speech sounds can be done with continuous distributions. *The Journal of the Acoustical Society of America*, *133*(5), 398–404.

Wanrooij, K., Boersma, P., & Van Zuijen, T. L. (2014). Distributional vowel training is less effective for adults than for infants. A study using the mismatch response. *PLoS ONE*, *9*(10).

Wanrooij, K., Escudero, P., & Raijmakers, M. E. (2013). What do listeners learn from exposure to a vowel distribution? An analysis of listening strategies in distributional learning. *Journal of Phonetics*, *41*(5), 307–319.

Weber, A. & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25.

Werker, J. F. & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, *37*(1), 35–44.

Werker, J. F. & Tees, R. C. (1984a). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63.

Werker, J. F. & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *The Journal of the Acoustical Society of America*, *75*(6), 1866–1878.

White, E. J., Hutka, S. A., Williams, L. J., & Moreno, S. (2013). Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. *Frontiers in systems neuroscience*, *7*, 90.

Yoshida, K. A., Pons, F., Maye, J., & Werker, J. F. (2010). Distributional phonetic learning at 10 months of age. *Infancy*, *15*(4), 420–433.

Young-Scholten, M. (2002). Orthographic input in L2 phonological development. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: papers in honour of henning wode* (pp. 263–279).

Young-Scholten, M. & Langer, M. (2015). The role of orthographic input in second language German: Evidence from naturalistic adult learners' production. *Applied Psycholinguistics*, *36*(01), 93–114.

Appendix A

Summary table of the mixed effect model reported in section 3.3.2.4

TABLE A.1: A summary of the mixed-effect logistic regression model for variables predicting the response accuracy for participants trained on either a unimodal or bimodal distribution paired with mismatching orthography. The reference level for CONDITION is 'Control'

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-0.049	0.392	-0.126	0.900
Unimodal	1.393	0.617	2.258	0.024
Bimodal	1.948	0.616	3.163	0.002
USRS	-1.154	0.612	-1.887	0.059
BDRS	3.071	0.624	4.924	0.000
UDRS	2.799	0.610	4.589	0.000
BSRS	-0.542	0.606	-0.894	0.371
USAS	-1.021	0.612	-1.670	0.095
BDAS	2.542	0.612	4.157	0.000
UDAS	2.823	0.620	4.552	0.000
BSAS	-1.208	0.616	-1.962	0.050

Appendix B

Stimuli used for experiments in Chapter 4

TABLE B.1: The auditory, visual and orthographic stimuli presented to the Control group in the word-learning experiment.

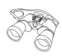
Auditory form	Pictured-meaning	Spelling
/banu/		XXX
/bannu/		XXX
/duna/		XXX
/dunna/		XXX
/kamu/		XXX
/kammu/		XXX
/lama/		XXX
/lamma/		XXX
/suma/		XXX
/summa/		XXX
/tana/		XXX
/tanna/		XXX

TABLE B.2: The auditory, visual and orthographic stimuli presented to the Distinct Roman Spelling group (DRS) in the word-learning experiment.

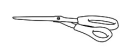
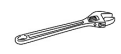
Auditory form	**Pictured-meaning**	**Spelling**
/banu/		banu
/bannu/		bannu
/duna/		duna
/dunna/		dunna
/kamu/		kamu
/kammu/		kammu
/lama/		lama
/lamma/		lamma
/suma/		suma
/summa/		summa
/tana/		tana
/tanna/		tanna

TABLE B.3: The auditory, visual and orthographic stimuli presented to the Same Roman Spelling group (SRS) in the word-learning experiment.

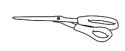
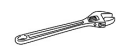
Auditory form	**Pictured-meaning**	**Spelling**
/banu/		banu
/bannu/		banu
/duna/		duna
/dunna/		duna
/kamu/		kamu
/kammu/		kamu
/lama/		lama
/lamma/		lama
/suma/		suma
/summa/		suma
/tana/		tana
/tanna/		tana

TABLE B.4: The auditory, visual and orthographic stimuli presented to the Distinct Arabic Spelling group (DAS) in the word-learning experiment.

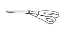
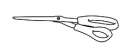
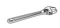
Auditory form	**Pictured-meaning**	**Spelling**
/banu/		بنو
/bannu/		بنّو
/duna/		دنا
/dunna/		دنّا
/kamu/		كمو
/kammu/		كمّو
/lama/		لما
/lamma/		لمّا
/suma/		سما
/summa/		سمّا
/tana/		تنا
/tanna/		تنّا

TABLE B.5: The auditory, visual and orthographic stimuli presented to the Same Arabic Spelling group (SAS) in the word-learning experiment.

Auditory form	**Pictured-meaning**	**Spelling**
/banu/		بنو
/bannu/		بنو
/duna/		دنا
/dunna/		دنا
/kamu/		كمو
/kammu/		كمو
/lama/		لما
/lamma/		لما
/suma/		سما
/summa/		سما
/tana/		تنا
/tanna/		تنا