Effects of Orthography on Monolingual and Bilingual Perception of Non-native Consonant Clusters

Hana Zjakic

Master of Research

Western Sydney University

27 October 2017
Acknowledgements

Firstly, thank you to my wonderful supervisors Dr. Alba Tuninetti and A/Prof Paola Escudero for their guidance, encouragement and most of all, their patience this past year. I would like to extend my appreciation to Dr. Karen Mulak for being a mentor figure to me during my time at MARCS. Thank you to my amazing family and friends who have kept me sane and who have provided me with endless love and support throughout my entire university journey. I would also like to thank the Graduate Research School and the MARCS Institute for Brain, Behaviour and Development for granting me this incredible opportunity and providing me with the resources needed to complete this project. Finally, I wish to acknowledge the students who participated in my study because without them, this thesis would not have been possible.
ORTHOGRAPHIC INFLUENCES ON WORD LEARNING

Statement of Authentication

The work presented in this is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

Hana Zjakic

27 October 2017
ORTHOGRAPHIC INFLUENCES ON WORD LEARNING

Table of Contents

List of Tables ......................................................................................................................... iii
List of Figures ........................................................................................................................ iv
Abstract ..................................................................................................................................... v

1. Introduction ...................................................................................................................... 1
   1.1 Features of Writing Systems .................................................................................... 3
   1.2 The Influence of Orthographic Features on Language Learning ......................... 4
   1.3 The Influence of L1 Orthography on Word Learning............................................ 6
   1.4 The Influence of Orthographic Input on L2 Word Learning: Does the Presentation of Spelling Actually Help? ...................................................................................... 11
   1.5 Word Learning Accuracy Between Monolinguals and Bilinguals: Is There a Difference? ............................................................................................................................. 15
   1.6 The Present Study .................................................................................................... 17

2. Method ............................................................................................................................ 20
   2.1 Research Design ....................................................................................................... 20
   2.2 Word Learning ......................................................................................................... 20
      2.2.1 Participants ....................................................................................................... 20
      2.2.2 Stimuli ............................................................................................................. 22
      2.2.3 Procedure ........................................................................................................ 24
   2.3 Phoneme Deletion ..................................................................................................... 26
      2.3.1 Participants ....................................................................................................... 26
      2.3.2 Stimuli ............................................................................................................. 26
ORTHOGRAPHIC INFLUENCES ON WORD LEARNING

2.3.3 Procedure .................................................................................................................26

3. Results.................................................................................................................................27

3.1 Data Screening ..................................................................................................................27

3.2 Word Learning Accuracy .................................................................................................28

3.3 Phoneme Deletion Accuracy ............................................................................................31

3.4 First- and Second-Phoneme Repair Strategies .................................................................38

3.5 The Link Between Word Learning and Phoneme Deletion .............................................45

4. Discussion ...........................................................................................................................47

4.1 Interpretation of Results ..................................................................................................48

4.1.1 Word Learning .............................................................................................................48

4.1.2 Phoneme Deletion ........................................................................................................52

4.1.3 The Link Between Word Learning and Phoneme Deletion ........................................56

4.2 Implications ......................................................................................................................57

References ..............................................................................................................................59

Appendix ................................................................................................................................73
ORTHOGRAPHIC INFLUENCES ON WORD LEARNING

List of Tables

Table 1. Means for Self-Rated Proficiency for Participants in the Three Language Groups .....21
Table 2. Total Number of Participants in Each Group and Condition for Word Learning and
Phoneme Deletion ........................................................................................................ 27
Table 3. Mean Word Learning Scores Across Pair Type, Condition and Group ..............31
Table 4. Mean 'Strict-Correct' First- and Second-Phoneme Deletion Scores Across Group,
Condition and Word Type ........................................................................................ 34
Table 5. Mean 'Lenient-Correct' First- and Second-Phoneme Deletion Scores Across Group,
Condition and Word Type ........................................................................................ 38
Table 6. Examples of Each Repair Strategy for the Novel-Bosnian Word 'Mrok' ..............39
Table 7. Repair Strategies for First-Phoneme Deletion Across Group, Condition and Word
Type .............................................................................................................................. 40
Table 8. Most Frequently Used Repair Strategy for Each Novel-Bosnian Word During First-
Phoneme Deletion ...................................................................................................... 42
Table 9. Repair Strategies for Second-Phoneme Deletion Across Group, Condition and Word
Type .............................................................................................................................. 43
Table 10. Most Frequently Used Repair Strategy and Change for Each Novel-Bosnian Word
During Second-Phoneme Deletion ............................................................................... 45
List of Figures

Figure 1. A typical sonority scale taken from Colantoni et al. (2015)..........................11

Figure 2. Mean word learning accuracy scores for the non-MPs, easy-MPs and difficult-MPs across group and condition. ......................................................................................................................30

Figure 3. Mean 'strict-correct' first- and second-phoneme deletion scores for the novel-Bosnian and novel-BP words across A and A+O conditions .................................................................34

Figure 4. Mean 'lenient-correct' first- and second-phoneme deletion scores for the novel-Bosnian and novel-BP words across A and A+O conditions .................................................................37

Figure 5. Mean percentage accuracy scores for word learning and phoneme deletion. ........46
ORTHOGRAPHIC INFLUENCES ON WORD LEARNING

Abstract

A critical skill for word decoding and reading development is that of orthographic processing, which refers to how orthographic symbols combine to represent word forms. Previous research has shown that the orthographic system of one’s native-language (L1) influences second-language (L2) learning, however there is conflicting evidence regarding the effects of presenting learners with written input during the initial stages of word learning. Therefore, an explicit word-learning paradigm, with auditory-only and auditory-orthography conditions, was used to investigate whether Australian-English (AusE) monolinguals and bilinguals who spoke AusE and one other language could learn novel words containing non-native onset consonant clusters – sound structures that have not been largely studied. A follow-up phoneme deletion task assessed the participants’ ability to manipulate the sounds of the newly learned words and determined whether participants used repair strategies to bring the non-native consonant clusters in line with their L1 phonotactics. Overall, word learning and phoneme deletion accuracy was similar across the language groups. In both tasks, those who received orthographic input during word learning significantly outperformed those who did not, however the facilitative effects of orthographic input varied across the word types. Participants most often perceived the two consonants in onset position as one sound or perceptually modified the first consonant so that the consonant cluster did not violate their L1 phonotactics. The results contribute to our understanding of how listeners learn foreign words and process non-native sound contrasts, and inform existing models that account for the interaction between orthography and perception. Furthermore, the results have potential implications for language learning programs in classroom settings.
1. Introduction

Understanding the way in which symbols combine to represent sounds and word forms is known as orthographic processing, with ‘orthography’ referring to the conventional writing system of a language (Varnhagen, Boechler & Steffler, 1999). This type of linguistic processing is enabled by two fundamental skills: phonological awareness, which is the ability to recognise, manipulate, segment and decode the sound structure of words; and grapheme-phoneme knowledge, which refers to the knowledge of the sound-letter correspondences in a language. These skills are critical for word decoding and reading development (Ehri, 2014; Patel, Snowling, & de Jong, 2004).

Second-language (L2) learners will often acquire an L2 with a different writing system to their native-language (L1). Furthermore, adult L2 learners are typically exposed to both written and auditory input during the initial stages of the learning process, during which phonological and orthographic processing simultaneously occurs (Bassetti, 2008; Colantoni, Steele & Escudero, 2015; Ijalba & Obler, 2015; Perfetti, Van Dyke & Hart, 2001). This is because the phonology associated with the shape of the written word is automatically activated when processing written input, and the orthographic shape associated with the sound is automatically activated when listening to auditory input (Kaushanskaya & Marian, 2009a). The accuracy, efficiency and automaticity of one’s phonological and orthographic processing abilities determine their word learning and reading success. Given the strong link between written input and L1 phonology, it is important to understand how this link may affect L2 acquisition.

Previous research has investigated the effects of orthography on L2 learning from the perspective of the learner’s L1 orthographic system as well as the orthographic system of the target language. L1 orthography has been found to influence L1, as well as L2 spelling and
reading efficiency in children (e.g., Dixon, Zhao & Joshi, 2010; Ellis & Hooper, 2001; Ellis et al., 2004; Goswami, Gombert & de Barrera, 1998; Hoxhallari, van Daal & Ellis, 2004; Seymour, Aro, & Erskine, 2003) and adults (e.g., Burki, Spinelli & Gaskell, 2012; Hamada & Koda, 2008; Ijalba & Obler, 2015; Koda, 1990; Wade-Woolley, 1999; Wang, Koda & Perfetti, 2003). Furthermore, previous research has found that during the initial stages of L2 learning, presenting learners with auditory and written input leads to conflicting results showing positive (e.g., Davidson, 2010; Erdener & Burnham, 2005; Escudero, Hayes-Harb & Mitterer, 2008; Escudero & Wanrooij, 2010; Kaushanskaya & Marian, 2009a; Steele, 2005), negative (e.g., Cutler, 2015; Davidson, 2007; Detey & Nespoulous, 2008; Escudero & Wanrooij, 2010; Hayes-Harb, Brown & Smith, 2017; Hayes-Harb, Nicol & Barker, 2010; Rafat, 2016; Young-Scholten, 1998; Young-Scholten, Akita, & Cross, 1999), and no effects (e.g., Escudero, 2015; Showalter & Hayes-Harb, 2015) on L2 word learning compared to presenting auditory stimuli alone.

While considerable research regarding the effects of orthography has focused on L1 and L2 reading comprehension, less attention has been given to the role of orthography on learning words containing complex consonant clusters. Therefore, the present study addressed this by assessing monolinguals and bilinguals on their ability to learn foreign words containing non-native onset consonant clusters, using a similar explicit word-learning paradigm to Escudero (2015). Participants then completed a phoneme deletion task, a measure of phonological awareness (Wade-Woolley, 1999; Wang et al., 2003), to determine whether they could manipulate the sounds of the newly learned words. Overall, the present study investigated whether word learning and phoneme deletion performance is influenced by (i) one’s knowledge of one or two orthographic systems, and (ii) the orthographic depth of the learners’ L1 or L2, and whether accuracy is facilitated by the presentation of both the words’ auditory and written form.
The remainder of the introduction will provide an overview of the features of different writing systems and in particular, the ones selected for the present study. Following from that, an account of previous studies investigating the effects of L1 orthography, as well as orthographic input on L2 learning will be provided. The notion of a bilingual advantage with respect to language learning will then be discussed, with reference to previous research. Finally, a summary of the present study will be provided following this background information.

1.1 Features of Writing Systems

Writing systems provide a visual analysis of a language and vary on two dimensions: orthographic representation and orthographic depth (Bassetti, 2008). Orthographic representation refers to the specific linguistic units that each symbol represents (Ziegler & Goswami, 2005). For example, while alphabetic systems (e.g., Spanish) represent phonemes (or letters), consonantal writing systems (e.g., Arabic) represent consonants, and syllabic writing systems (e.g., Japanese hiragana) and logographic writing systems (e.g., Chinese) represent syllables and morphemes (or words) respectively (Bassetti, 2008; Wang et al., 2003).

Orthographic depth, on the other hand, refers to the degree of grapheme-phoneme regularities within a language (Hamada & Koda, 2008). Languages can be classified as either ‘transparent’ or ‘opaque’, or ‘shallow’ and ‘deep’ in some literature, depending on these regularities (Bassetti, 2008; Frost, Katz & Bentin, 1987). In ‘transparent’ or ‘shallow’ orthographies (e.g., Spanish), the writing system directly represents the phonology of the words, meaning grapheme-phoneme correspondences are more consistent. For example, in languages with transparent orthographies, one letter or cluster is always pronounced in the same way (Ziegler & Goswami, 2005). Contrastingly, in ‘opaque’ or ‘deep’ orthographies (e.g., English) where grapheme-phoneme correspondences are less consistent, a letter may
represent a number of different phonemes (e.g., the ‘a’ in ‘cake’, ‘saw’ and ‘car’), or a single phoneme may be represented by various letters (e.g., the ‘k’ in ‘kite’ and the ‘c’ in ‘cat’; Frost et al., 1987; Ijalba & Obler, 2015; Ziegler & Goswami, 2005).

1.2 The Influence of Orthographic Features on Language Learning

Previous studies have shown that the orthographic system of a language, particularly the degree of transparency, affects the rate of its acquisition. More specifically, research has found that children learning a transparent language develop spelling and reading abilities earlier than children who learn an opaque language (Defior & Serrano, 2005; Ellis & Hooper, 2001; Goswami et al., 1998; Patel et al., 2004; Seymour et al., 2003; Spencer, 2007; Ziegler & Goswami, 2005). For example, Goswami et al. (1998) found that 7-year-old English-speaking children decoded monosyllabic non-words significantly worse (12% correct) than French- (53% correct) and Spanish-speaking children (94% correct) of the same age and matched for standardized reading age. By 9 years of age, the same pattern emerged reflecting the relative transparencies of the orthographies (English, 51%; French, 73%; Spanish, 92%). Similarly, the European Concerted Action on Learning Disorders as a Barrier to Human Development (reported in Seymour et al., 2003) found that by the middle of first grade, children whose languages were orthographically transparent (e.g., Greek, Finnish, German, Italian, and Spanish) were close to ceiling for real and non-word reading, while children whose languages were orthographically opaque showed lower accuracy scores (Danish, Portuguese, French, and English). The English-speaking children had the lowest scores overall for real and non-word reading (34% and 29% respectively), perhaps due to the fact that English has inconsistencies in both spelling and reading. The aforementioned studies lend support to what Katz and Frost (1992) termed the Orthographic Depth Hypothesis (ODH).
The ODH explains the ways in which orthographic depth affects phonological decoding across different writing systems. It also proposes that learners adapt their phonological and orthographic processing strategies to the orthography they are learning (Goswami et al., 1998; Katz & Frost, 1992). According to the ODH, orthographies are placed on a continuum depending on the consistency of the language’s grapheme-phoneme relations (Dixon et al., 2010). On one end, learners of transparent orthographies recognise and process words through one-to-one decoding and symbol-to-sound translations. On the other end, learners of opaque orthographies rely more on visual orthographic structures to process words (Frost et al., 1987; Hamada & Koda, 2008; Katz & Frost, 1992; Spencer, 2007). Languages with syllabic writing systems, such as Japanese Kana and Tamil, are placed somewhere in the middle of the continuum as learners will often use both phonological and visual cues to recognise and process words (Dixon et al., 2010).

These differences between languages’ orthographic systems come into play when adults are tasked with acquiring an L2 as L2 learning is believed to be directly influenced by the orthographic distance between one’s L1 and the target language (Hamada & Koda, 2008; Ijalba & Obler, 2015). Issues in L2 learning may arise when there are differences in the orthographic features of the L1 and L2 as inherent L1 phonological and orthographic rules may constrain learners (see Script Dependent Hypothesis; Abu-Rabia & Siegel, 2002; Da Fontoura & Siegel, 1995; Frost et al., 1987; Wade-Woolley, 1999; Ziegler & Goswami, 2005). For example, teaching a native-Spanish speaker the relationship between individual graphemes and phonemes would facilitate learning in Spanish due to the orthographically transparent nature of the language. However, learning would be less efficient if the same speaker applied one-to-one grapheme-phoneme decoding strategies to an orthographically opaque language such as English. For this reason, it is important to consider the orthographic system of the learner’s L1 when investigating the effects of orthography on L2 learning.
1.3 The Influence of L1 Orthography on Word Learning

Previous studies have often employed cross-linguistic comparisons to investigate the influence of L1 orthography on word learning, with a focus on L1 and L2 reading comprehension (e.g., Ellis & Hooper, 2001; Hamada & Koda, 2008; Hoxhallari et al., 2004; Ijalba & Obler, 2015; Koda, 1990; Spencer & Hanley, 2003; Wade-Woolley, 1999; Wang et al., 2003; Wimmer & Goswami, 1994). Koda (1990) found that when learning to read Sanskrit symbols, accuracy and reading speed was considerably lower for phono-graphic L1 readers (Arabic, English and Spanish) compared to morpho-graphic L1 readers (Japanese). The author suggested that the phono-graphic readers’ lower reading performance was a result of their heavy reliance on phonological information during word processing, which would not have been accessible in Sanskrit as the participants had no knowledge of the Sanskrit writing system. Wade-Woolley (1999) investigated L1 language transfer on phonological and orthographic processing during L2 reading. Low to intermediate learners of L2 English with Russian (alphabetic) and Japanese (syllabic – hiragana, or logographic – kanji) L1s were assessed on a number of English reading tasks. The Japanese readers showed greater sensitivity to orthographic information compared to the Russian readers, demonstrating better performance on tasks related to orthographic knowledge and spelling recognition. However, when asked to manipulate the individual phonemic elements of the words in a phoneme deletion task, Japanese readers showed significantly lower performance suggesting poorer phonological awareness than the Russian readers as a result of their L1 orthographic system.

Similarly, Wang and colleagues (2003) found that native Korean speakers (an alphabetic language) were more attentive to phonological information, while native Chinese speakers (a non-alphabetic language) were more attentive to orthographic information when learning new English words. As a result, orthographic similarity between the words sparked more confusion among the Chinese speakers compared to the Korean speakers, perhaps
because Korean speakers also rely on phonological cues for word identification. The Chinese
speakers also performed worse than the Korean speakers in a follow up phoneme deletion
task, due to their L1 orthographic system, which relies on visual-orthographic processing as
opposed to phonological processing. Hamada and Koda (2008) assessed L1 Korean (an
alphabetic language) and L1 Chinese (a logographic language) learners of L2 English on a
non-word picture naming task with phonologically regular and irregular conditions. The
Korean group pronounced the non-words significantly faster and more accurately compared
to the Chinese group, reinforcing the notion that similarities between a learner’s L1 and L2
facilitates L2 learning.

Learning to read in a transparent orthography has been found to impose fewer
constraints on a reader than learning to read in an opaque orthography (Ellis & Hooper, 2001;
Ellis et al., 2004; Hoxhallari et al., 2004; Ijalba & Obler, 2015). For example, Ellis and
Hooper (2001) reported that after approximately two years of reading instruction, Welsh-
educated bilingual children (transparent orthography) were able to accurately read aloud
significantly more of their language (61% of tokens, 1821 word types) than English-educated
monolingual children (52% of tokens, 716 word types) who were matched for reading
instruction. Hoxhallari et al. (2004) compared reading performance of Albanian children to
those tested in Ellis and Hooper (2001) and found that the Albanian speakers (a more
transparent language than Welsh) were able to accurately read significantly more words than
the other two language groups, suggesting a linear trend with respect to orthographic
transparency of a language and one’s reading efficiency (i.e. the more transparent a language
is, the easier it is to acquire). Supporting this notion, Ellis et al. (2004) tested the reading
efficiency of children whose native languages were Japanese hiragana, Albanian, Greek,
English or Japanese kanji (in decreasing order of orthographic transparency, such that
Japanese kanji is the most opaque). The authors found that reading accuracy correlated with
orthographic depth: accuracy was highest for hiragana, then Albanian, Greek, English, and kanji. Ijalba and Obler (2015) found that native Spanish readers (transparent orthography) and native English readers (opaque orthography) learned a novel transparent writing system faster than a novel opaque writing system regardless of their L1. In a recall test, the Spanish readers had difficulty learning the novel opaque words compared to the novel transparent words, while the English readers presented no difference between the word types, providing further evidence for the influence of orthographic depth on word learning and reading efficiency.

L1 orthography has also been found to influence L2 spelling. For example, Dixon et al. (2010) found that Mandarin Chinese speakers (a morpho-syllabic language) had overall better spelling accuracy and made more real-word substitutions than Malay (an alphabetic language) and Tamil speakers (a syllabic language) when spelling words in their L2 (i.e. English for all three groups). The authors proposed that the use of high-frequency words in the study meant that the speakers of Mandarin Chinese were able to utilise their visual-orthographic processing, despite their weak processing of phonological cues. L1 orthography has also been found to influence listeners’ processing of unfamiliar phonotactic structures. For example, Burki et al. (2012) assessed native French speakers on their ability to spell twenty novel French words, each with a two letter consonant cluster in either word initial or internal position. Half of the words contained a schwa ‘e’ between the two consonants while half did not. The authors found lower spelling error rates for novel words with phonotactic structures found more frequently in French (e.g., words with an internally attested cluster and an ‘e’ in the spelling, e.g., relassier, or with an initially attested cluster without an ‘e’ in the spelling, e.g., plour) than for words with phonotactic structures found less frequently in French (e.g., words with an internally attested cluster without an ‘e’ in the spelling, e.g., rlassier, or with an initially attested cluster and an ‘e’ in the spelling, e.g., pelour), suggesting
that L1 orthographic knowledge can assist in constructing phonological representations of novel words.

The linguistic properties of one’s L1 have also been shown to influence the processing of non-native sounds and contrasts, also known as cross-linguistic influences (CLI) (Colantoni et al., 2015; Wade-Wooley, 1999). In fact, models of L2 speech perception have long held the notion that listeners hear non-native sounds and contrasts through the phonotactics of their L1 (see Perceptual Assimilation Model (Best, 1995) and Second-Language Linguistic Perception Model (Escudero, 2005, 2009; van Leussen & Escudero, 2015)). Non-native linguistic structures such as consonant clusters are particularly difficult for L2 learners to perceive and produce. Previous studies investigating the influence of L1 orthography on the perception and production of these complex structures have focused on the principle of syllable structure modification. Given that syllable structures differ across languages, L2 learners will often perceive or produce forms that differ from the auditory or orthographic input to which they are exposed. Furthermore, L2 learners will often apply perceptual modifications (or repair strategies) to conform non-native consonant clusters to their L1 phonotactics (Colantoni et al., 2015). For example, Spanish learners of English will often add a vowel before the word ‘Spain’ (pronouncing it as ‘Espain’) as the syllable structure in Spanish does not allow the sequence /sp/ in word-onset position (Bassetti, 2008). Likewise, stop-liquid clusters in English (e.g., drain /dren/; play /ple/) may be perceived via vowel insertion (e.g., drain as [dren]) or deletion of one of the consonants (e.g., play as [pe]) when L2 learners have an L1 with a dominant consonant-vowel (CV) or CVC syllable structure (e.g., Mandarin). According to Colantoni et al. (2015), the four main types of repair strategies include: (1) epenthesis, the insertion of a vowel or consonant; (2) deletion of one or more of the segments; (3) substitution of one or more of the segments; and (4) metathesis or re-ordering of the segments. Orthography is believed to be another main source of perceptual
modifications (as well as syllable structure), particularly for perceptual epenthesis (Colantoni et al., 2015). For example, orthographically opaque languages may consist of silent consonantal graphemes (e.g., the French <e> can correspond to unpronounced schwas such as *appeler* /aplə/ ‘to call’) which may trigger epenthesis, particularly when learners are exposed to the word’s written form before or at the same time as being exposed to the spoken form (Colantoni et al., 2015). For this reason, orthographically transparent languages may provide an advantage to L2 learners as the written input directly represents the phonology of the words.

Using minimal pair discrimination and word identification tasks, previous studies have shown that the aforementioned repair strategies are employed by L2 learners for consonant clusters in word-onset, medial and final position (e.g., Davidson & Shaw, 2012; Dupoux, Kakehi, Hirose, Pallier & Melher, 1999; Dupoux, Parlato, Frota, Hirose & Peperkamp, 2011; Halle, Segui, Frauenfelder & Meunier, 1998; Kabak & Idsardi, 2007; Pitt, 1998). Halle and colleagues (1998) found that native French speakers would perceive the ‘illegal’ French onset consonant cluster /dl/ as the ‘legal’ onset clusters /gl/ and /kl/.

Similarly, Pitt (1998) found that native English speakers would perceive illegal English onset consonant clusters as legal sequences (e.g., the illegal /tl/ as the legal /tr/). Furthermore, Dupoux et al. (1999) found that Japanese listeners would perceive illusionary epenthetic vowels within consonantal sequences that violated the phonotactic properties of Japanese. Likewise, Davidson and Shaw (2012) assessed native English speakers on discriminating novel words containing illegal onset consonant clusters in English. The authors found that listeners confused fricative-initial sequences with prosthesis (e.g., /zmatsu/-/ezmatu/), stop-stop sequences with vowel insertion (e.g., /ktase/-/ketase/) and stop-nasal sequences with a change or deletion of the first consonant (e.g., /bmafa/-/pmafa/ or /bmafa/-/mafa/), reinforcing the notion that L2 speech processing is influenced by L1 phonotactic knowledge.
It has been argued that across languages, there is stronger dispreference for consonant clusters when the consonants are similar in sonority, which relates to the acoustic intensity (i.e., the loudness) of the speech sounds (Colantoni et al., 2015; Moreton, 2002).

**Least sonorous**

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stops</td>
<td>Fricatives</td>
<td>Nasals</td>
<td>Liquids</td>
<td>Semi-consonants</td>
<td>Vowels</td>
</tr>
</tbody>
</table>

/p t k b d g .../  /f s ʃ v z ʒ .../  /m n ɲ ŋ .../  /l ɹ .../  /w j ɥ .../  /a e i o u .../

Fig. 1. A typical sonority scale taken from Colantoni et al. (2015). From left to right, speech sounds increase in sonority.

Most often, sonority increases from the syllable margins (the onset and coda) towards the nucleus; however, in some languages or language families (e.g., Slavic languages), sonority sequencing violations occur (Colantoni et al., 2015). Due to L1-based CLI, these complex forms are some of the most challenging linguistic structures for L2 learners to master. For this reason, it is important to consider whether orthographic input aids L2 learning during the beginning stages of acquisition, particularly for complex, non-native structures.

### 1.4 The Influence of Orthographic Input on L2 Word Learning: Does the Presentation of Spelling Actually Help?

The interaction of L1 phonology and orthography with L2 orthographic and auditory input is believed to affect learners’ L2 phonological representations, influencing L2 production and performance in word learning and phonological awareness tasks (Bassetti, 2008). However, previous research has revealed conflicting evidence showing positive (e.g., Davidson, 2010; Erdener & Burnham, 2005; Escudero, Hayes-Harb & Mitterer, 2008; Escudero & Wanrooij, 2010; Kaushanskaya & Marian, 2009a; Steele, 2005), negative (e.g.,
Cutler, 2015; Detey & Nespoulous, 2008; Escudero & Wanrooij, 2010; Hayes-Harb et al., 2017; Hayes-Harb et al., 2010; Rafat, 2016; Young-Scholten, 1998; Young-Scholten et al., 1999), and no effects (e.g., Escudero, 2015; Showalter & Hayes-Harb, 2015) of orthographic input on L2 word learning.

There is evidence that orthographic input aids learners with the perception of target phonemes, syllables and words. For example, Steele (2005) found that L1 Chinese speakers of beginner-level French were more accurate at perceiving the uvular fricative /ʁ/ in a consonant cluster (e.g., traineau) when they were presented with both orthographic and auditory input. In the absence of orthographic input, listeners were more likely to perceive the consonant cluster as a consonant followed by aspiration (i.e., [tʰ]). Another study (Erdener & Burnham, 2005) assessed monolingual Australian-English (AusE) and Turkish adult speakers on their ability to learn Spanish (transparent orthography) and Irish (opaque orthography) words. The results demonstrated a significant facilitative effect of the auditory-orthographic condition compared to the auditory-only condition for both language groups; however, the effect was stronger for the Turkish group, perhaps due to its degree of orthographic transparency compared to AusE.

Escudero et al. (2008) assessed native Dutch speakers on their ability to learn the L2 English contrast /æ/ and /ɛ/ in novel English words (e.g., /bæskəl/ and /bɛstət/) and to associate them with their corresponding visual referents. All participants received the auditory form of the words, while only some participants received the written forms of the words as well (e.g., <bestet> for /bɛstət/ and <baskle> for /bæskəl/). An eye tracker determined whether participants could identify the target word from a set of visual referents. While participants in the auditory-only condition could not distinguish between the /æ/ and /ɛ/ tokens, participants who received orthographic input only fixated on the pictures associated with the /ɛ/ words when presented with the /ɛ/ tokens, and the pictures associated
with the /æ/ and /ɛ/ words when presented with the /æ/ tokens, suggesting that orthographic input can enhance L2 lexical encoding of novel contrasts. Furthermore, Escudero and Wanrooij (2010) found that pairing auditory stimuli with their orthographic labels assisted L1 Spanish learners of Dutch (opaque orthography) in categorising vowel tokens of novel Dutch words, but only for specific vowel contrasts, suggesting that orthographic input can have both facilitative and hindering effects on L2 word learning.

Previous research has found that orthographic input can also lead to non-target-like productions (such as phoneme additions, omissions and substitutions) which would probably not occur if learners were only exposed to auditory input (Bassetti, 2008). For instance, Young-Scholten (1998; Young-Scholten et al., 1999) tested native English speakers on their ability to learn words containing Polish consonant clusters and then recall them in a picture naming task (with no prior knowledge of the Polish language). Interestingly, results showed that participants would insert epenthetic vowels when presented with difficult consonant clusters, particularly when presented with the words’ orthographic and auditory input simultaneously. The author argued that adults prefer epenthesis to consonant omission, as they want to retain all of the consonants present in the orthographic input (Young-Scholten, 1998). Contrastingly, participants would omit a consonant from the cluster when presented with the auditory input only. Similarly, Detey and Nespoulous (2008) assessed native Japanese learners of French on their ability to parse different types of consonant clusters presented in auditory, auditory-visual and visual modes. Higher rates of epenthesis were observed for all cluster types in the two visual conditions, suggesting a possible influence of orthography on L2 representations. The authors argue that participants in the auditory-visual condition may have focused on only one modality to reach a phonological judgement. The authors also argue that participants may have been sensitive to the phonetic properties of the auditory stimuli, while the visual stimuli may have activated phonological representations.
that were more in line with their L1 phonological and orthographic structures, leading to the higher rates of epenthesis in the visual conditions (Detey & Nespoulous, 2008).

Rafat (2016) examined the effects of orthographic input on L1 phonological transfer in Canadian-English learners of Spanish (L1 phonological transfer defined as the production of an English phoneme when the grapheme corresponds to a different phoneme in Spanish). A picture-naming task revealed that L1 phonological transfer was highest for those who received orthographic input during training. Furthermore, exposure to the same orthographic input during training and testing had no facilitative effects on L2 production. Similarly, Hayes-Harb et al. (2017) investigated native English speakers’ acquisition of German final obstruent devoicing and found that participants who were exposed to both the auditory and written forms of novel German words (e.g., /ʃtaɪt/ and /ʃtaɪd/, both pronounced [ʃtaɪt] but written as <Steit> and <Steid>) would produce final voiced obstruents in a follow up picture naming task. As this was also the case when they were given explicit information that the written word forms were misleading, the findings lend support to the notion that orthographic input can hinder the acquisition of target forms.

Finally, orthographic input has also been found to have no effect on L2 word learning compared to auditory input alone. For example, Escudero (2015) tested native English (opaque orthography) and Spanish (transparent orthography) speakers on the ability to learn novel Dutch words in the form of non-minimal pairs (non-MPs; i.e. sets of words that differ in two or more of their segments) and minimal pairs (MPs; i.e. sets of words that differ in only one segment, either their consonant or vowel) in an explicit word-learning task. During learning, participants were explicitly taught associations between the novel words and their corresponding visual referents. Participants in the auditory + orthography condition were also presented with the orthographic input of the word during the learning phase. During the test phase, participants were assessed on their ability to associate the words and objects that they
had been exposed to during learning. Results showed no effect of orthographic input on learning both types of words regardless of their L1. Likewise, Showalter and Hayes-Harb (2015) tested native English speakers on their ability to learn minimal novel pairs distinguished by the Arabic velar-uvular contrast (e.g., [kubu], [qubu]) and manipulated the orthographic input presented to participants. Results showed that regardless of whether participants were exposed to orthographic input or not, they were consistently unable to associate the phonemes with the novel words.

Based on the literature presented, it is apparent that the effects of orthographic input on L2 word learning remain contested. Of particular interest is whether the effects of orthographic input on word learning differs for bilinguals who have knowledge of two orthographies compared to monolinguals who have experience with only one orthographic system. The following section discusses previous studies that have investigated word learning between monolingual and bilingual populations.

1.5 Word Learning Accuracy Between Monolinguals and Bilinguals: Is There a Difference?

Word learning studies comparing monolingual and bilingual performance have shown conflicting results (e.g., Escudero, Mulak, Fu & Singh, 2016; Escudero, Mulak & Vlach, 2016; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009a; Kaushanskaya & Marian, 2009b; Kaushanskaya & Rechtzigel, 2012; Mattock, Polka, Rvachew, & Krehm, 2010; Zjakic, Tuninetti & Escudero, 2016). Using an explicit word learning paradigm, Kaushanskaya and Marian (2009b) taught monolingual English speakers, Spanish-English bilinguals (languages that are orthographically and phonologically similar) and Mandarin-English bilinguals (languages that are orthographically and phonologically different) novel auditory words. After participants heard each word, they were presented with the English orthographic translation for each word and were then tested on their ability to map the
auditory words to their correct orthographic translation. The authors found that both bilingual groups outperformed the monolingual group. Similarly, Kaushanskaya (2012) found that Spanish-English bilinguals outperformed native English monolinguals when the novel words consisted of phonemes from both languages.

Using an implicit version of the word-learning task (during which learners infer word-object associations by tracking word-object co-occurrences over a number of learning trials), Escudero, Mulak, Fu et al. (2016) found that simultaneous bilingual Singaporean English-Mandarin speakers were overall more accurate at learning novel word-referent mappings than their monolingual AusE counterparts. However, using similar implicit word learning paradigms, Escudero, Mulak and Vlach (2016) and Zjakic et al. (2016) found that participants were able to learn non-MPs and MPs above chance level regardless of their language background. Importantly, these studies tested mixed-bilingual groups, which may have accounted for their results.

Despite conflicting results regarding word learning performance in bilingual populations, previous research has suggested that bilingualism may foster a range of cognitive advantages that may facilitate word learning, such as greater performance compared to monolinguals on tests of executive function (e.g., Adesope, Lavin, Thompson & Ungerleider, 2010; Bialystok & Craik, 2010; Bialystok, Craik & Luk, 2008; Bialystok, Craik & Ryan, 2006; Bialystok & Viswanathan, 2009). In bilinguals, the brain areas associated with executive control and language control are believed to overlap significantly (Pliatsikas & Luk, 2016). Despite the belief that bilinguals experience slower performance in word learning as a result of increased activation of competitor words (as bilinguals have a lexicon in each language), the fact that bilinguals have outperformed monolinguals in novel word-learning tasks (e.g., Escudero, Mulak, Fu et al., 2016; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009b) suggests that they can suppress competitor activation in the non-target language.
As bilinguals use the phonotactic and orthographic knowledge of their two languages to assist with categorising non-native speech sounds, bilinguals are believed to have enhanced phonological awareness compared to monolinguals (Kaushanskaya & Marian, 2009a; 2009b). In fact, bilinguals have been found to outperform monolinguals on measures of phonological awareness, particularly when their L1 and L2 share similar phonological and orthographic structures. For example, Bialystok, Majumder and Martin (2003) found that Spanish-English bilinguals outperformed monolinguals on a phoneme segmentation task, while Chinese-English bilinguals performed worse, reinforcing the notion that phonological awareness is enhanced when the learner’s L1 and L2 share orthographic and phonotactic features. Similarly, Bialystok, Luk and Kwan (2005) found that bilingual children whose languages both had alphabetic writing systems were better at phonological awareness and decoding tasks than bilinguals whose languages had dissimilar writing systems.

1.6 The Present Study

While previous research has found that the orthographic system of one’s L1 influences L2 processing, there is conflicting evidence regarding the effects of orthographic input on L2 word learning. Furthermore, relatively few studies have investigated the effects of orthography on learning words containing consonant clusters. Therefore, the present study did this by using an explicit word-learning task (similar to Escudero, 2015) to assess participants’ ability to learn 16 novel transparent words, eight of which contained onset consonant clusters in CCVC form (e.g., dlok). The chosen consonant clusters conformed to Bosnian phonotactics as Bosnian allows onset consonant clusters, which violate English phonotactics. Another eight novel words without consonant clusters (CVCV form) were
included as filler words (e.g., fafe). The novel words together formed non-MPs (e.g., fafe-dlok), easy-MPs (e.g., dlok-mrok) or difficult-MPs (e.g., dlok-dvok).

Considering the conflicting evidence regarding a bilingual advantage for word learning, the present study compared task performance between monolingual and bilingual populations. All participants had native-like proficiency in AusE, while the bilingual participants also had intermediate to native-like auditory proficiency in a language other than English (LOTE) and reported some exposure to their LOTE orthography. The bilinguals were divided into two groups (transparent or opaque) based on the orthographic depth of their LOTE. Thus, the present study investigated whether word learning is influenced by the knowledge of one or two orthographic systems and by the orthographic depth of the monolinguals’ L1 or the bilinguals’ L1 and L2. In order to determine the effects of orthographic input on novel word learning, half of the participants received the words’ auditory and written form during word learning, while the other half received auditory input only. Participants also completed a phoneme deletion task to determine whether they were able to manipulate the sounds of the newly learned words, or whether they would implement repair strategies to bring the non-native consonant clusters in line with their L1 phonotactics.

Although the bilinguals in the present study reported minimal exposure to their LOTE orthography, it was hypothesised that the bilinguals would outperform monolinguals in both word learning and phoneme deletion, as bilinguals are generally believed to have enhanced phonological awareness skills and have previously been found to outperform monolinguals in similar tasks. More specifically, it was predicted that the bilinguals in the transparent group would show higher accuracy overall due to the orthographic similarities between their transparent language and the transparent orthography of the stimuli. While presenting learners with orthographic input has been found to have conflicting effects on word learning, it was predicted that in this case, participants who receive both auditory and written input
would outperform those who receive auditory input only (for both word learning and phoneme deletion) as the words in the present study differ in their consonants and consonants are believed to be more salient in word recognition (Nespor, Pena & Mehler, 2003). As learning MPs requires listeners to encode fine phonetic detail, it was hypothesised that participants would have greater accuracy scores for learning non-MPs compared to MPs and easy-MPs compared to difficult-MPs. Furthermore, as the consonant clusters used in the present study violate English phonotactics, it was predicted that participants would have higher first- and second-phoneme deletion accuracy for the novel-BP words compared to the novel-Bosnian words.
2. Method

2.1 Research Design

A 3 (Group: monolingual (M) vs. bilingual transparent (BT) vs. bilingual opaque (BO)) x 2 (Condition: audio only (A) vs. audio + orthography (A+O)) x 2 (Task: word learning vs. phoneme deletion) overall mixed design was employed. Participants were assigned to a language group based on their auditory comprehension of either one or two languages. The M group comprised native speakers of AusE, while the participants in the BT group understood a LOTE with a transparent orthography (i.e. a language with more consistent grapheme-phoneme correspondences than AusE). Participants in the BO group understood a LOTE with an opaque orthography (i.e. a language where grapheme-phoneme correspondences are less consistent such as in AusE).

Within each language group, half of the participants were randomly assigned to the A or the A+O condition. Participants in the A+O condition received the orthographic representation of the words along with the auditory input during the learning phase of the word-learning task. Regardless of whether the participants were assigned to the A or A+O condition, each participant completed two tasks: a word-learning task and a phoneme deletion task. The participants’ accuracy scores for each task (word learning and phoneme deletion) were calculated and analysed.

2.2 Word Learning

2.2.1 Participants

Participants included 41 undergraduate students from Western Sydney University, ranging from 18-38 years of age ($M = 20.98, SD = 4.51$), with no hearing, speaking or language learning impairments. Prior to the experiment, all participants provided informed consent according to the Western Sydney University Human Research Ethics Committee and
completed a language background questionnaire focusing on the participants’ language exposure from infancy to adulthood in various domains (see Appendix 1). All participants reported either near-native to native auditory proficiency in AusE, and intermediate to native auditory proficiency in their LOTE if they were bilingual. Almost all participants reported native-like reading and writing proficiency in English. Although the bilinguals’ median level of reading and writing proficiency in their LOTE was low, all participants had some exposure to their LOTE orthography. The transparent languages included Amharic: \( n = 1 \) (Abdi & Therrien, 2016), Bengali: \( n = 1 \) (Sircar & Nag, 2013), Farsi: \( n = 2 \) (Panah & Padakannaya, 2008), Hindi: \( n = 3 \) (Rao, Vaid, Srinivasan & Chen, 2011), Italian: \( n = 1 \) (Filippo et al., 2005), Spanish: \( n = 1 \) (Ijalba & Obler, 2015), Thai: \( n = 1 \) (Liao, Kuo, Deenang & Mok, 2016), Turkish: \( n = 1 \) (Erdener & Burnham, 2005) and Vietnamese: \( n = 3 \) (Petrescu, Helms-Park & Dronjic, 2017). The opaque languages included Arabic: \( n = 12 \) (Rao et al., 2011), Cantonese: \( n = 2 \) (Wang & Geva, 2003) and Urdu: \( n = 2 \) (Rao et al., 2011). Thus, there were 11 M, 14 BT and 16 BO participants. Table 1 illustrates the mean self-ratings of the participants’ language proficiency in their L1 and L2 if they were bilingual.

Table 1

<table>
<thead>
<tr>
<th>Language Group</th>
<th>M</th>
<th>BT</th>
<th>BO</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 Comprehension</td>
<td>5.96 (0.21)</td>
<td>5.94 (0.24)</td>
<td>5.93 (0.26)</td>
</tr>
<tr>
<td>L1 Production</td>
<td>5.83 (0.58)</td>
<td>5.83 (0.51)</td>
<td>5.78 (0.65)</td>
</tr>
<tr>
<td>L1 Writing</td>
<td>5.74 (0.75)</td>
<td>5.80 (0.63)</td>
<td>5.73 (0.84)</td>
</tr>
<tr>
<td>L1 Reading</td>
<td>5.87 (0.34)</td>
<td>5.86 (0.36)</td>
<td>5.78 (0.69)</td>
</tr>
<tr>
<td>L2 Comprehension</td>
<td>-</td>
<td>4.67 (1.17)</td>
<td>4.60 (1.19)</td>
</tr>
<tr>
<td>L2 Production</td>
<td>-</td>
<td>4 (1.32)</td>
<td>4 (1.34)</td>
</tr>
<tr>
<td></td>
<td>L2 Writing</td>
<td>L2 Reading</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.29 (1.46)</td>
<td>2.30 (1.42)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.50 (1.67)</td>
<td>2.57 (1.61)</td>
</tr>
</tbody>
</table>

*Note. Standard deviations are shown in parentheses. Self-rated proficiency is on a 6 point scale with 6 indicating native-like proficiency and 1 indicating very low proficiency.*

Participants were recruited from Western Sydney University research participation pool, and were reimbursed with credit points for their participation. Another four participants were tested but were excluded from the sample before the stage of data analysis due to the following: technical issues during the experiment (*n* = 3) and reporting problems with reading development on the language background questionnaire (*n* = 1). Importantly, participants had no experience with any Slavic languages, most of which allow the onset consonant clusters that were used in the present study.

### 2.2.2 Stimuli

Stimuli comprised of 16 novel words and their corresponding visual referents. Eight novel words had onset consonant clusters with a CCVC structure that conformed to Bosnian phonotactics, and the eight novel words without consonant contrasts had a CVCV structure that conformed to Brazilian Portuguese (BP) phonotactics. The novel words were produced by native speakers of these two languages respectively. The Bosnian novel words were recorded by two female and two male native speakers of Bosnian who were born in Sarajevo, Bosnia and Herzegovina (*M* age = 21.1, *SD* = 2.13). The voice recordings were conducted in a quiet room at the Sarajevo School of Science and Technology, using a Zoom H4n microphone and 2.1.2 version of the computer software program Audacity®. The eight novel-Bosnian words were embedded into two Bosnian carrier sentences so that they were produced naturally in continuous speech. The carrier sentences included: “Ovo je [word]” and “Pritisnite na [word]” translating to “This is a [word]” and “Click on the [word]”. During the recording session, the speakers were presented with a PowerPoint where each slide
contained the two carrier sentences along with one of the eight words (so that there were eight slides in total). Each speaker repeated the two carrier sentences one after the other, five times, for each of the eight words. All four speakers received monetary compensation for their time.

Using the computer software Praat (Boersma & Weenink, 2017), two tokens for each novel-Bosnian word were segmented from the carrier sentences produced by one Bosnian male- and female speaker. These tokens were used in the learning phase of the word-learning task. For the test phase of the word-learning task, one token for each word was segmented from the carrier sentences produced by the other Bosnian male and female speaker. Each novel-Bosnian word contained a different onset consonant cluster non-native to English, with ‘-ok’ following the two-letter consonant cluster. The eight novel-Bosnian words included: dlok, mnok, hvok, ptok, zfok, mrok, dnok, and dvok. The eight novel-BP words were fafe, pipe, sase, kuke, popo, teko, seso, and koko, which were recorded by two native male and two native female speakers. They were selected from the Escudero, Boersma, Rauber and Bion (2009) corpus. The novel-Bosnian and novel-BP words together formed non-minimal pairs (non-MPs; e.g., dnok-seso), where the word pair differed in most or all of their segments, while the novel-Bosnian words together formed either easy-MPs (e.g., dlok-mrok), where the word pair differed in both of the consonants in the cluster, and difficult-MPs (e.g., dnok-dvok), where the word pair differed in only one of the consonants in the cluster.

To ensure consistency across the sound stimuli, the average scale intensity for each sound file was adjusted to 65 dB SPL using the computer software Pratt (Boersma & Weenink, 2017). Each word had a different corresponding black and white line drawing taken from the Shatzman and McQueen (2006) corpus, which were the same line drawings used in Zjakic et al. (2016).
2.2.3 Procedure

Using the computer software program E-prime 2.0 (Psychology Software Tools Inc.), participants were individually tested on a word-learning task in a quiet test booth. The experiment was completed on a laptop, with the sound stimuli presented over closed headphones. The volume of the stimuli was held constant across participants. Participants received oral instructions for each part of the experiment as well as instructions presented to them on the laptop screen.

Once consent was obtained, the participants completed a word-learning task consisting of a learning phase and a test phase similar to the procedure in Escudero (2015). During learning, participants were told that they would be learning new words along with their corresponding images and that they would be tested on the words in a follow up task. During each learning trial, participants heard one of the novel words and viewed its corresponding image in the centre of a laptop screen in front of them for a duration of 1000 ms. In the A+O condition, the spelling of the word was also presented directly above the image in bold, size 18 font. Immediately afterwards, participants heard the same word and saw two images appear side by side on the screen (i.e. the target image, along with a distractor image, which corresponded to one of the other novel words). Participants had to indicate whether the word represented the left or the right image by pressing the right or left ALT key on the keyboard. The next learning trial started 1 s after the key press. The learning phase consisted of 128 trials, during which each of the 16 novel words were randomly presented as a target word eight times. The image pairs formed 96 non-MP trials, 21 easy-MP trials, and 11 difficult-MP trials. Each novel-Bosnian word was presented by a native male speaker four times and a native female speaker four times, and each novel-BP word was presented by a native male speaker four times and a native female speaker four times. The
order of the images was counterbalanced on the left and right side of the screen, as well as the number of times a word appeared as a distractor.

Participants completed the test phase, which consisted of 240 trials presented in a random order, immediately after the learning phase. In the test phase, participants were assessed on their ability to make word-object associations for the 16 novel words presented in the learning phase. In each test trial, participants heard the audio for one of the 16 words and viewed two images side by side on the laptop screen (i.e. the target image and a distractor image). The image pairs formed 184 non-MP trials, 40 easy-MP trials, and 16 difficult-MP trials. As in the learning phase, participants were instructed to select the image that the word corresponded to by pressing the right or left ALT key. The orthographic representations of the words were not presented in the test phase. The next test trial commenced 500 ms after the participant made their response. During the test phase, each word was presented as a target 15 times. The words in the test phase were produced by different speakers to those in the learning phase to add variation to the stimuli and to ensure that the task was more difficult for the participants. Half of the participants in the A and A+O conditions were presented with a native female Bosnian speaker and a native male BP speaker, while the other half were presented with a native male Bosnian speaker and a native female BP speaker. The order of the images was counterbalanced on the left and right side of the screen, and the number of times a word appeared as a distractor was also counterbalanced. Each word was paired with each of the other 15 words once. Participants took approximately 20 minutes to complete the word-learning task, including both the learning and test phase.
2.3 Phoneme Deletion

2.3.1 Participants

All participants completed the phoneme deletion task after word learning except one participant from the A+O condition/BT group who did not complete the task due to task complexity.

2.3.2 Stimuli

The stimuli included the 16 novel words from the word-learning task. The words were produced by the four speakers in the test phase of the word-learning task (i.e. one native male- and female speaker of Bosnian, and one native male- and female speaker of BP).

2.3.3 Procedure

Immediately after word learning, participants were tested on their ability to manipulate the sound structures of the newly learned words by completing a phoneme deletion task (created using E-prime 2.0; Psychology Software Tools Inc.). Participants received oral instructions for the task prior to commencing, as well as written instructions on the laptop screen. The task was made up of four blocks – within each block, each of the 16 words were presented to participants twice in a random order (once by the male speakers of Bosnian and BP, and once by the female speakers of Bosnian and BP).

In two of the four blocks, participants were presented with the audio for each word and were required to delete the first sound of the word and type the remaining word using the keyboard (e.g., ‘dlok’ = ‘lok’, ‘sase’ = ‘ase’). In the other two blocks, participants were presented with the audio for each word and were required to delete the second sound of each word and type the remaining word using the keyboard (e.g., ‘dlok’ = ‘dok’, ‘sase’ = ‘sse’). The four blocks were randomised and counterbalanced so that participants could receive the four blocks in any order. After the participants heard each word, they would receive a prompt
in the centre of the computer screen to remind them which sound they had to delete (i.e. “Delete the FIRST sound” or “Delete the SECOND sound”). Each trial began after the participant typed their response and pressed the enter key. Regardless of whether participants were in the A or A+O condition during word learning, participants were not presented with written input during the phoneme deletion task. The phoneme deletion task took approximately 10 minutes for participants to complete.

3. Results

3.1 Data Screening

Initial data screening for outliers revealed that two participants in the BO group (one in the A+O condition, one in the A condition) performed at chance (50%) for non-MPs which have previously had high learning percentages (e.g., Escudero, 2015; Escudero, Mulak, Fu et al., 2016; Escudero, Mulak & Vlach, 2016). These two participants were therefore excluded from the data analysis, resulting in a total of 39 participants for word learning and 38 participants for phoneme deletion. The table below illustrates the number of participants in each group and condition for both experimental tasks.

Table 2

<table>
<thead>
<tr>
<th>Group/Condition</th>
<th>M</th>
<th>BT</th>
<th>BO</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+O</td>
<td>6</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>A</td>
<td>7</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>A+O</td>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total Number of Participants in Each Group and Condition for Word Learning and Phoneme Deletion
3.2 Word Learning Accuracy

Participants’ correct and incorrect responses for word learning were analysed in a mixed-effects logit model, with group (M, BT, BO), condition (A, A+O) and pair type (non-MPs, easy-MPs, difficult-MPs) as fixed effects, and subject number, trial number, speaker group, sound file, and picture 1 and 2 (the paired images presented to the learner during the test phase) as random effects. A mixed-effects logit model was an appropriate statistical model to use as differences in sample sizes across groups (which is the case in the present sample) do not influence the model’s outcome (Pinheiro & Bates, 2000), and because the results comprised categorical data (Baayen, Davidson, & Bates, 2008). Analysis revealed no main effect of group ($\chi^2 (2, n = 9360) = .66, p = .72$), but a main effect of condition ($\chi^2 (1, n = 9360) = 24.51, p < .001$) as participants in the A+O condition had significantly better word learning accuracy scores compared to those in the A condition ($\beta = .13$, 95% CI [.08, .19], $p < .001$). A main effect of pair type was also found ($\chi^2 (2, n = 9360) = 130.85, p < .001$) as participants were overall better at non-MPs compared to easy-MPs ($\beta = .07$, 95% CI [.05, .09], $p < .001$) and difficult-MPs ($\beta = .18$, 95% CI [.14, .22], $p < .001$), and easy-MPs compared to difficult-MPs ($\beta = .11$, 95% CI [.07, .15], $p < .001$).

Although there was no interaction between condition and group ($\chi^2 (2, n = 9360) = .97, p = .62$), there was an interaction between condition and pair type ($\chi^2 (2, n = 9360) = 17.54, p < .001$), such that the A+O condition had significantly higher accuracy scores for non-MPs and easy-MPs compared to the A condition (non-MPs: $\beta = .11$, 95% CI [.07, .16], $p < .001$; easy-MPs: $\beta = .17$, 95% CI [.10, .24], $p < .001$). Within both conditions, participants had higher accuracy scores for non-MPs compared to easy-MPs (A: $\beta = .10$, 95% CI [.05, .15], $p < .001$; A+O: $\beta = .04$, 95% CI [.02, .07], $p < .001$) and difficult-MPs (A: $\beta = .14$, 95%
ORTHOGRAPHIC INFLUENCES ON WORD LEARNING

CI [.09, .20], p < .001; A+O: β = .19, 95% CI [.14, .24], p < .001), while the A+O condition had higher accuracy scores for easy-MPs compared to difficult-MPs (β = .15, 95% CI [.09, .20], p < .001). An interaction between group and pair type was also found (χ² (4, n = 9360) = 12.88, p = .01), as within each group, participants were significantly better at non-MPs compared to easy-MPs (M: β = .05, 95% CI [.00, .10], p = .04; BT: β = .07, 95% CI [.03, .11], p < .001; BO: β = .08, 95% CI [.05, .12], p < .001) and difficult-MPs (M: β = .13, 95% CI [.09, .16], p < .001; BT: β = .22, 95% CI [.14, .30], p < .001; BO: β = .19, 95% CI [.11, .27], p < .001), and easy-MPs compared to difficult-MPs (M: β = .08, 95% CI [.03, .13], p = .001; BT: β = .15, 95% CI [.07, .24], p = .001; BO: β = .10, 95% CI [.02, .18], p = .01).

The word learning results are best explained by the significant interaction between group, condition and pair type (χ² (4, n = 9360) = 21.07, p < .001). Fisher’s LSD post hoc pairwise comparisons revealed that for the bilingual groups, the A+O condition performed significantly better than the A condition for non-MPs (BO: β = .11, 95% CI [.05, .18], p = .001; BT: β = .14, 95% CI [.04, .23], p = .006) and easy-MPs (BO: β = .22, 95% CI [.14, .29], p < .001; BT: β = .20, 95% CI [.09, .31], p < .001) but not for difficult-MPs (BO: β = .00, 95% CI [-.14, .15], p = .97; BT: β = .04, 95% CI [-.15, .24], p = .66). Contrastingly, for the M group, the A+O condition performed significantly better than the A condition for difficult-MPs (β = .15, 95% CI [.04, .26], p = .008) but not for non-MPs (β = .08, 95% CI [-.01, .17], p = .08) and easy-MPs (β = .09, 95% CI [-.04, .22], p = .19).

Within the A+O condition, all three groups performed significantly better for non-MPs compared to difficult-MPs (BO: β = .23, 95% CI [.10, .36], p = .001; BT: β = .24, 95% CI [.15, .33], p < .001; M: β = .10, 95% CI [.05, .14], p < .001), while the BO and M groups also performed significantly better for non-MPs compared to easy-MPs (BO: β = .04, 95% CI [.01, .07], p = .007; M: β = .05, 95% CI [.01, .09], p = .03). The bilingual groups also showed significantly higher results for easy-MPs compared to difficult-MPs (BO: β = .18, 95% CI
ORTHOGRAPHIC INFLUENCES ON WORD LEARNING

[.05, .31], p = .006; BT: β = .21, 95% CI [.10, .31], p < .001), while the M group did not (β = .05, 95% CI [-.02, .12], p = .13). Within the A condition, the BO and BT groups had higher scores for non-MPs compared to easy-MPs (BO: β = .15, 95% CI [.08, .22], p < .001; BT: β = .10, 95% CI [.05, .16], p < .001) and difficult-MPs (BO: β = .12, 95% CI [.02, .21], p = .02; BT: β = .15, 95% CI [.04, .26], p = .007), while there was no significant difference in their performance between easy-MPs and difficult-MPs (BO: β = -.03, 95% CI [-.11, .05], p = .44; BT: β = .05, 95% CI [-.06, .16], p = .38). The M group showed significantly better performance for non-MPs compared to difficult-MPs (β = .17, 95% CI [.12, .22], p < .001) and easy-MPs compared to difficult-MPs (β = .11, 95% CI [.05, .17], p < .001), however there was no significant difference between non-MPs and easy-MPs (β = .05, 95% CI [-.05, .15], p = .29). The group, condition and pair type interaction is illustrated in Figure 2 and the descriptive statistics are provided in Table 3.

![Mean word learning accuracy scores for the non-MPs, easy-MPs and difficult-MPs across group and condition.](image-url)
Table 3

Mean Word Learning Scores Across Pair Type, Condition and Group

<table>
<thead>
<tr>
<th>Condition</th>
<th>Group</th>
<th>Pair Type</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>non-MPs</td>
<td>easy-MPs</td>
<td>difficult-MPs</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>M</td>
<td>.83 (.04)</td>
<td>.78 (.06)</td>
<td>.66 (.05)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BT</td>
<td>.83 (.05)</td>
<td>.73 (.05)</td>
<td>.68 (.09)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BO</td>
<td>.84 (.03)</td>
<td>.70 (.03)</td>
<td>.73 (.03)</td>
<td></td>
</tr>
<tr>
<td>A+O</td>
<td>M</td>
<td>.91 (.02)</td>
<td>.86 (.03)</td>
<td>.81 (.03)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BT</td>
<td>.97 (.01)</td>
<td>.93 (.03)</td>
<td>.72 (.05)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BO</td>
<td>.96 (.01)</td>
<td>.91 (.03)</td>
<td>.73 (.07)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses.

3.3 Phoneme Deletion Accuracy

A mixed-effects logit model was used to analyse participants’ correct and incorrect responses for the phoneme deletion task, with group (M, BT, BO), condition (A+O, A), deletion (first, second) and word type (novel-Bosnian, novel-BP) as fixed effects, and subject, trial number, speaker, sound file and order as random effects. Responses were classified as ‘strict-correct’ if participants deleted the correct sound (i.e., first or second) and then accurately identified the other sound in the cluster (e.g., if participants typed ‘dok’ when instructed to delete the second phoneme of the novel word ‘dlok’). Results showed no main effect of group ($\chi^2 (2, n = 4864) = 2.05, p = .36$), but a main effect of deletion ($\chi^2 (1, n = 4864) = 24.30, p < .001$) as participants were significantly better at deleting the first phoneme of the novel words compared to the second phoneme ($\beta = .13, 95\% \text{ CI} [.08, .18], p < .001$). A main effect of condition was also found ($\chi^2 (1, n = 4864) = 30.12, p < .001$) as participants in
the A+O condition performed significantly better than those in the A condition ($\beta = .30$, 95% CI [.20, .41], $p < .001$). Furthermore, participants performed better overall for the novel-BP words compared to the novel-Bosnian words ($\beta = .27$, 95% CI [.22, .32], $p < .001$) as shown by the main effect of word type ($\chi^2 (1, n = 4864) = 114.98$, $p < .001$).

There were no interactions between deletion and group ($\chi^2 (2, n = 4864) = 1.66$, $p = .44$), or deletion and word type ($\chi^2 (1, n = 4864) = 1.27$, $p = .26$), but an interaction between deletion and condition was found ($\chi^2 (1, n = 4864) = 8.85$, $p = .003$) as deletion accuracy was significantly higher in the A+O condition compared to the A condition (first-phoneme: $\beta = .31$, 95% CI [.20, .43], $p < .001$; second-phoneme: $\beta = .27$, 95% CI [.15, .39], $p < .001$).

Within the A+O condition, participants were significantly better at deleting the first phoneme of the novel-words and correctly identifying the second phoneme than deleting the second phoneme ($\beta = .12$, 95% CI [.08, .16], $p < .001$). Contrastingly, there was no difference between first- and second phoneme deletion accuracy within the A condition ($\beta = .07$, 95% CI [-.02, .16], $p = .13$).

No interactions were found between group and condition ($\chi^2 (2, n = 4864) = .76$, $p = .69$), or group and word type ($\chi^2 (2, n = 4864) = 3.69$, $p = .16$), however there was an interaction between condition and word type ($\chi^2 (1, n = 4864) = 24.50$, $p < .001$) such that participants showed significantly higher accuracy scores for both word types in the A+O condition compared to the A condition (novel-Bosnian: $\beta = .24$, 95% CI [.12, .37], $p < .001$; novel-BP: $\beta = .28$, 95% CI [.16, .39], $p < .001$). Within both conditions, participants’ overall accuracy was significantly higher for the novel-BP words compared to the novel-Bosnian words (A: $\beta = .20$, 95% CI [.11, .29], $p < .001$; A+O: $\beta = .24$, 95% CI [.17, .31], $p < .001$).

The results are best explained by the significant interaction between deletion, condition and word type ($\chi^2 (1, n = 4864) = 16.61$, $p < .001$). Fisher’s LSD post hoc pairwise
comparisons revealed that for both word types, participants had significantly higher accuracy scores in the A+O condition compared to the A condition for first-phoneme deletion (novel-Bosnian: $\beta = .15, 95\% \text{ CI} [.04, .27], p = .01$; novel-BP: $\beta = .37, 95\% \text{ CI} [.20, .54], p < .001$), as well as second-phoneme deletion (novel-Bosnian: $\beta = .32, 95\% \text{ CI} [.18, .46], p < .001$; novel-BP: $\beta = .16, 95\% \text{ CI} [.06, .26], p = .001$). Within the A+O condition, participants had significantly higher scores for the novel-BP words compared to the novel-Bosnian words for both types of deletion (first-phoneme: $\beta = .20, 95\% \text{ CI} [12, .27], p < .001$; second-phoneme: $\beta = .26, 95\% \text{ CI} [17, .35], p < .001$). Within the A condition, there was no significant difference between the first-phoneme deletion scores for the novel-Bosnian and novel-BP words ($\beta = .02, 95\% \text{ CI} [-.13, .18], p = .76$), while participants were significantly better at deleting the second phoneme of the novel-BP words compared to the novel-Bosnian words ($\beta = .41, 95\% \text{ CI} [.34, .49], p < .001$).

Across both conditions, participants’ deletion accuracy was significantly higher for the first phoneme compared to the second phoneme for the novel-Bosnian words (A: $\beta = .30, 95\% \text{ CI} [.23, .38], p < .001$; A+O: $\beta = .13, 95\% \text{ CI} [.07, .20], p < .001$). Within the A condition, there was no significant difference between first- and second-phoneme deletion for the novel-BP words ($\beta = -.14, 95\% \text{ CI} [-.29, .02], p = .08$), while in the A+O condition, participants were significantly more accurate at deleting the first phoneme than the second phoneme for the novel-BP words ($\beta = .07, 95\% \text{ CI} [.03, .11], p = .001$). The deletion, condition and word type interaction is illustrated in Figure 3 and descriptive statistics are illustrated in Table 4.
Fig. 3. Mean 'strict-correct' first- and second-phoneme deletion scores for the novel-Bosnian and novel-BP words across A and A+O conditions.

Table 4

Mean 'Strict-Correct' First- and Second-Phoneme Deletion Scores Across Group, Condition and Word Type

<table>
<thead>
<tr>
<th>Condition</th>
<th>Group</th>
<th>First</th>
<th>Second</th>
<th>First</th>
<th>Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>M</td>
<td>.63 (.07)</td>
<td>.33 (.08)</td>
<td>.48 (.19)</td>
<td>.75 (.08)</td>
</tr>
<tr>
<td></td>
<td>BT</td>
<td>.70 (.06)</td>
<td>.28 (.06)</td>
<td>.66 (.14)</td>
<td>.78 (.07)</td>
</tr>
<tr>
<td></td>
<td>BO</td>
<td>.59 (.09)</td>
<td>.41 (.12)</td>
<td>.69 (.11)</td>
<td>.72 (.10)</td>
</tr>
<tr>
<td>A+O</td>
<td>M</td>
<td>.72 (.09)</td>
<td>.64 (.09)</td>
<td>.96 (.03)</td>
<td>.86 (.06)</td>
</tr>
<tr>
<td></td>
<td>BT</td>
<td>.83 (.06)</td>
<td>.63 (.10)</td>
<td>.98 (.02)</td>
<td>.93 (.03)</td>
</tr>
<tr>
<td></td>
<td>BO</td>
<td>.80 (.06)</td>
<td>.70 (.07)</td>
<td>.99 (.01)</td>
<td>.94 (.02)</td>
</tr>
</tbody>
</table>

Note. Standard errors are in parentheses.
A second analysis was conducted to determine whether any group differences would arise if the classification of correct responses were less stringent. A mixed-effects logit model, with the same fixed and random effects, was used to analyse participants’ correct and incorrect responses. This time, responses were classified as ‘lenient-correct’ if participants deleted the correct sound regardless of whether the other sound was accurately identified (e.g., if participants typed ‘gok’ when instructed to delete the second phoneme in ‘dlok’). Differences in accuracy between the ‘strict-correct’ and ‘lenient-correct’ scores would suggest the implementation of repair strategies such as those proposed by Davidson and Shaw (2012). Results showed no main effect of group ($\chi^2 (2, n = 4864) = .58, p = .75$) or word type ($\chi^2 (1, n = 4864) = .05, p = .83$), but a main effect of deletion ($\chi^2 (1, n = 4864) = 46.74, p < .001$) as participants were significantly better at deleting the first phoneme of the novel words compared to the second phoneme ($\beta = .09, 95\%\ CI [.05, .13], p < .001$). A main effect of condition was also found ($\chi^2 (1, n = 4864) = 27.73, p < .001$) as participants in the A+O condition significantly outperformed those in the A condition ($\beta = .17, 95\%\ CI [.10, .25], p < .001$).

There were no interactions between group and deletion ($\chi^2 (2, n = 4864) = 4.71, p = .10$), group and condition ($\chi^2 (2, n = 4864) = 2.17, p = .34$), or group and word type ($\chi^2 (2, n = 4864) = 3.70, p = .16$), but an interaction between deletion and condition was found ($\chi^2 (1, n = 4864) = 6.70, p = .01$). Deletion accuracy was significantly higher in the A+O condition compared to the A condition (first-phoneme: $\beta = .14, 95\%\ CI [.08, .20], p < .001$; second-phoneme: $\beta = .20, 95\%\ CI [.09, .30], p < .001$). Within each condition, participants were significantly better at deleting the first phoneme of the novel words compared to the second phoneme (A: $\beta = .11, 95\%\ CI [.04, .18], p = .003$; A+O: $\beta = .05, 95\%\ CI [.02, .09], p = .002$). An interaction between deletion and word type ($\chi^2 (1, n = 4864) = 24.39, p < .001$) revealed that participants were significantly better at deleting the first phoneme of the novel-Bosnian
words compared to the novel-BP words ($\beta = .06, 95\% CI [.01, .11], p = .03$), while second-phoneme deletion was significantly better for the novel-BP words compared to the novel-Bosnian words ($\beta = .15, 95\% CI [.10, .21], p < .001$). Participants were significantly better at deleting the first phoneme of the novel-Bosnian words compared to the second phoneme ($\beta = .20, 95\% CI [.14, .27], p < .001$), while there was no difference between first- and second-phoneme deletion for the novel-BP words ($\beta = -.01, 95\% CI [-.06, .04], p = .72$).

An interaction between condition and word type was also found ($\chi^2 (1, n = 4864) = 10.76, p = .001$) such that participants showed significantly higher accuracy scores for both word types in the A+O condition compared to the A condition (novel-Bosnian: $\beta = .11, 95\% CI [.04, .17], p = .001$; novel-BP: $\beta = .24, 95\% CI [.13, .36], p < .001$). Within the A+O condition, participants’ overall deletion accuracy was significantly better for the novel-BP words compared to the novel-Bosnian words ($\beta = .03, 95\% CI [.00, .05], p = .02$), while there was no significant difference between the word types for the A condition ($\beta = -.11, 95\% CI [-.22, .00], p = .06$).

The phoneme deletion results are best explained by the significant interaction between deletion, condition and word type ($\chi^2 (1, n = 4864) = 5.67, p = .02$). Fisher’s LSD post hoc pairwise comparisons revealed that for both word types, participants had significantly higher accuracy scores in the A+O condition compared to the A condition for first-phoneme deletion (novel-Bosnian: $\beta = .03, 95\% CI [.00, .07], p = .05$; novel-BP: $\beta = .37, 95\% CI [.20, .54], p < .001$), as well as second-phoneme deletion (novel-Bosnian: $\beta = .25, 95\% CI [.13, .37], p < .001$; novel-BP: $\beta = .14, 95\% CI [.04, .24], p = .006$). Within the A+O condition, no significant differences were found between first-phoneme deletion of the novel-Bosnian and novel-BP words ($\beta = .00, 95\% CI [-.02, .02], p = .81$), but second-phoneme deletion was significantly better for the novel-BP words than the novel-Bosnian words ($\beta = .10, 95\% CI [.03, .16], p = .002$). Within the A condition, first-phoneme deletion was significantly better
for the novel-Bosnian words compared to the novel-BP words ($\beta = .33$, 95% CI [.16, .51], $p < .001$), while participants were significantly better at deleting the second phoneme of the novel-BP words compared to the novel-Bosnian words ($\beta = .21$, 95% CI [.12, .30], $p < .001$).

Across both conditions, participants’ deletion accuracy was significantly higher for the first phoneme than the second phoneme for the novel-Bosnian words ($A$: $\beta = .33$, 95% CI [.25, .42], $p < .001$; $A+O$: $\beta = .11$, 95% CI [.04, .18], $p = .001$). Contrastingly, within the $A$ condition, participants’ deletion accuracy for the novel-BP words was significantly higher for the second phoneme than the first phoneme ($\beta = .21$, 95% CI [.06, .36], $p = .006$), while there was no difference between first- and second-phoneme deletion of the novel-BP words in the $A+O$ condition ($\beta = .02$, 95% CI [.00, .05], $p = .11$). The deletion, condition and word type interaction is illustrated in Figure 4 and descriptive statistics are illustrated in Table 5.

Fig. 4. Mean 'lenient-correct' first- and second-phoneme deletion scores for the novel-Bosnian and novel-BP words across $A$ and $A+O$ conditions.
Table 5

Mean 'Lenient-Correct' First- and Second-Phoneme Deletion Scores Across Group, Condition and Word Type

<table>
<thead>
<tr>
<th>Condition Group</th>
<th>novel-Bosnian First</th>
<th>novel-Bosnian Second</th>
<th>novel-BP First</th>
<th>novel-BP Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>A M</td>
<td>.98 (.01)</td>
<td>.67 (.08)</td>
<td>.48 (.19)</td>
<td>.82 (.09)</td>
</tr>
<tr>
<td>BT</td>
<td>.95 (.03)</td>
<td>.56 (.06)</td>
<td>.66 (.14)</td>
<td>.86 (.08)</td>
</tr>
<tr>
<td>BO</td>
<td>.88 (.06)</td>
<td>.61 (.11)</td>
<td>.69 (.11)</td>
<td>.79 (.10)</td>
</tr>
<tr>
<td>A+O M</td>
<td>.94 (.03)</td>
<td>.88 (.09)</td>
<td>.96 (.03)</td>
<td>.95 (.04)</td>
</tr>
<tr>
<td>BT</td>
<td>.99 (.01)</td>
<td>.86 (.06)</td>
<td>.98 (.02)</td>
<td>.96 (.02)</td>
</tr>
<tr>
<td>BO</td>
<td>.99 (.01)</td>
<td>.87 (.03)</td>
<td>.99 (.01)</td>
<td>.97 (.01)</td>
</tr>
</tbody>
</table>

Note: Standard errors are in parentheses.

3.4 First- and Second-Phoneme Repair Strategies

Participants’ written responses for each word were analysed to determine whether any repair strategies were used as proposed by Davidson and Shaw (2012). The analysis only focused on the repair strategies employed for the first two sounds of each word as participants were only instructed to delete these sounds. The repair strategies in the analysis included: change of the first consonant (C1 change), change of the second consonant (C2 change), two sounds perceived as one, vowel insertion, and prosthesis (the addition of a letter or syllable at the start of a word). ‘C2 change’ was only relevant to the novel-Bosnian words, while the other repair strategies were applicable to both word types. Participants’ responses were labelled as incorrect if the response differed entirely from the target word or if the
participant deleted the wrong sound. Table 6 illustrates what each repair strategy would look like in the first- and second-phoneme deletion blocks for the novel-Bosnian word ‘mrok’.

Table 6

*Examples of Each Repair Strategy for the Novel-Bosnian Word 'Mrok'*

<table>
<thead>
<tr>
<th>Repair strategy</th>
<th>First-phoneme deletion</th>
<th>Second-phoneme deletion</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 change</td>
<td>[V]brok</td>
<td>b[r]ok</td>
</tr>
<tr>
<td>C2 change</td>
<td>[m]lok</td>
<td>m[V]lok</td>
</tr>
<tr>
<td>2 sounds perceived as 1</td>
<td>[mr]ok</td>
<td>m[ro]k / [m][r][o]k</td>
</tr>
<tr>
<td>Vowel insertion</td>
<td>[m]erok</td>
<td>m[V]rok</td>
</tr>
<tr>
<td>Prosthesis</td>
<td>[V]mrok</td>
<td>e[m]rok</td>
</tr>
</tbody>
</table>

*Note.* Phoneme/s deleted/inserted/modified by participant are in brackets. V = vowel. Some examples consist of more than one repair strategy.

Across all of the trials for the first-phoneme deletion blocks (2,432 in total), 339 trials (13.94%) included one or more repair strategies. Of those 339 trials, there were 318 instances (13.08%) of one repair strategy being used, and 21 instances (0.86%) of two repair strategies being implemented in a single response (across all participants and both word types). 1,823 trials (74.96%) comprised correct responses which meant that no repair strategy was employed. The remainder of the trials (n = 270, 11.10%) were classified as incorrect responses. Concerning the second-phoneme deletion blocks (2,432 trials in total), 571 trials (23.48%) included one or more repair strategies. There were 483 instances (19.86%) of one repair strategy being used, 86 (3.54%) instances of two repair strategies being used, and two (0.08%) instances of three repair strategies in a single response. 1,599 trials (65.75%) comprised correct responses, while the remainder of the trials (n = 261, 10.73%) comprised incorrect responses.
The frequencies for each repair strategy across group, condition and word type are presented in Table 7 for the first-phoneme deletion blocks. Frequencies should be interpreted with caution as there were unequal numbers of participants across the groups and conditions.

Table 7

*Repair Strategies for First-Phoneme Deletion Across Group, Condition and Word Type*

<table>
<thead>
<tr>
<th>Repair strategy</th>
<th>M</th>
<th>BT</th>
<th>BO</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 change</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A+O</td>
<td>1</td>
<td>-</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td></td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>A+O</td>
<td>-</td>
<td>6</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>6</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>C2 change</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A+O</td>
<td>13</td>
<td>6</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>A</td>
<td>12</td>
<td></td>
<td>16</td>
<td>24</td>
</tr>
<tr>
<td>A+O</td>
<td>24</td>
<td></td>
<td>24</td>
<td>48</td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td>30</td>
<td>40</td>
<td>145</td>
</tr>
<tr>
<td>2 sounds perceived as 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A+O</td>
<td>23</td>
<td>44</td>
<td>25</td>
<td>40</td>
</tr>
<tr>
<td>A</td>
<td>26</td>
<td></td>
<td>26</td>
<td>52</td>
</tr>
<tr>
<td>A+O</td>
<td>54</td>
<td></td>
<td>54</td>
<td>108</td>
</tr>
<tr>
<td>Total</td>
<td>212</td>
<td>93</td>
<td>111</td>
<td>416</td>
</tr>
<tr>
<td>Vowel insertion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A+O</td>
<td>8</td>
<td>8</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td></td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>A+O</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>18</td>
<td>18</td>
<td>57</td>
</tr>
<tr>
<td>Prosthesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A+O</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td></td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>A+O</td>
<td>22</td>
<td></td>
<td>22</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>16</td>
<td>16</td>
<td>73</td>
</tr>
<tr>
<td>Correct entry (no repair strategy)</td>
<td>139</td>
<td>100</td>
<td>160</td>
<td>156</td>
</tr>
<tr>
<td>Incorrect entry</td>
<td>7</td>
<td>2</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>185</td>
<td>76</td>
<td>188</td>
<td>148</td>
</tr>
</tbody>
</table>
Note. Single listener responses consisting of two or more repair strategies were tallied separately. The bottom row labelled ‘Total’ represents the overall number of repair strategy instances for that group/condition.

<table>
<thead>
<tr>
<th>Incorrect entry</th>
<th>7</th>
<th>84</th>
<th>4</th>
<th>76</th>
<th>1</th>
<th>79</th>
<th>251</th>
</tr>
</thead>
</table>

The table above illustrates that for the first-phoneme deletion blocks, repair strategies were employed for the novel-Bosnian words to bring the non-native consonant clusters in line with English phonotactics. The data revealed that listeners most frequently perceived the consonant clusters in the novel-Bosnian words as one sound, which meant that they either perceptually deleted one of the consonants in the cluster (e.g., ‘ptok’ as ‘tok’) or they misperceived the two consonants as a different single consonant (e.g., ‘ptok’ as ‘dok’). More repair strategies were used in the A conditions compared to the A+O conditions across the three language groups. Within the A condition, the BO group used the highest number of repair strategies ($n = 108$), compared to the BT group ($n = 69$) and the M group ($n = 60$). Contrastingly, in the A+O condition, the M group had the highest number of repair strategies ($n = 49$), compared to the BO group ($n = 40$) and the BT group ($n = 34$). Table 8 shows the most frequently used repair strategy for each novel-Bosnian word in the first-phoneme deletion blocks.

As the repair strategies proposed by Davidson and Shaw (2012) relate to consonant clusters, the absence of repair strategies for the novel-BP words is not surprising. Interestingly, the number of incorrect responses for the novel-BP words was much higher than the novel-Bosnian words. Furthermore, the number of incorrect responses for the novel-BP words were much higher in the A conditions compared to the A+O conditions. When analysing the incorrect responses for the novel-BP words, a pattern of syllable deletion emerged. Of the 251 incorrect responses for the novel-BP words, participants deleted the first syllable rather than the first phoneme 71.31% of the time (179 trials; $A = 174$ trials, $A+O = 5$...
trials). This pattern was similar across the three language groups ($M = 61$ trials, $BT = 59$ trials, $BO = 59$ trials).

Table 8

*Most Frequently Used Repair Strategy for Each Novel-Bosnian Word During First-Phoneme Deletion*

<table>
<thead>
<tr>
<th>Repair strategy</th>
<th>2 sounds perceived as 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>dlok</td>
<td>(17.11%)</td>
</tr>
<tr>
<td>ptok</td>
<td>(38.82%)</td>
</tr>
<tr>
<td>dnok</td>
<td>(7.24%)</td>
</tr>
<tr>
<td>dvok</td>
<td>(12.5%)</td>
</tr>
<tr>
<td>zvok</td>
<td>(22.37%)</td>
</tr>
<tr>
<td>mnok</td>
<td>(13.82%)</td>
</tr>
<tr>
<td>hvok</td>
<td>(25.66%)</td>
</tr>
<tr>
<td>mrok</td>
<td>Prosthesis (7.24%)</td>
</tr>
</tbody>
</table>

*Note. Words are listed from least to most difficult based on the total number of incorrect responses across all participants.*

*Note. Percentages represent the total number of instances of that repair strategy across the total number of trials for that word (each word had a total number of 152 trials across all participants).

The frequencies for each repair strategy across group, condition and word type are shown in Table 9 for the second-phoneme deletion blocks.
Table 9

Repair Strategies for Second-Phoneme Deletion Across Group, Condition and Word Type

<table>
<thead>
<tr>
<th>Repair strategy</th>
<th>M</th>
<th>BT</th>
<th>BO</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A+O</td>
<td>A</td>
<td>A+O</td>
<td>A</td>
</tr>
<tr>
<td>C1 change</td>
<td>43</td>
<td>61</td>
<td>45</td>
<td>69</td>
</tr>
<tr>
<td>C2 change</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>2 sounds perceived as 1</td>
<td>15</td>
<td>30</td>
<td>15</td>
<td>49</td>
</tr>
<tr>
<td>Vowel insertion</td>
<td>5</td>
<td>13</td>
<td>1</td>
<td>26</td>
</tr>
<tr>
<td>Prosthesis</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>19</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>65</td>
<td>108</td>
<td>65</td>
<td>167</td>
</tr>
<tr>
<td>Correct entry (no repair strategy)</td>
<td>122</td>
<td>52</td>
<td>121</td>
<td>62</td>
</tr>
<tr>
<td>Incorrect entry</td>
<td>14</td>
<td>21</td>
<td>13</td>
<td>21</td>
</tr>
</tbody>
</table>

**Novel-BP**

<table>
<thead>
<tr>
<th>Repair strategy</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
</tr>
<tr>
<td>C1 change</td>
<td>17</td>
</tr>
<tr>
<td>C2 change</td>
<td>-</td>
</tr>
<tr>
<td>2 sounds perceived as 1</td>
<td>-</td>
</tr>
<tr>
<td>Vowel insertion</td>
<td>-</td>
</tr>
<tr>
<td>Prosthesis</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>-</td>
</tr>
<tr>
<td>Correct entry (no repair strategy)</td>
<td>165</td>
</tr>
<tr>
<td>Incorrect entry</td>
<td>10</td>
</tr>
</tbody>
</table>

*Note.* Single listener responses consisting of two or more repair strategies were tallied separately. The bottom row labelled ‘Total’ represents the overall number of repair strategy instances for that group/condition.
For the second-phoneme deletion blocks, there more instances of repair strategies for the novel-Bosnian words than the novel-BP words. The most frequent repair strategy employed for the novel-Bosnian words was ‘C1 change’ (e.g., ‘dlok’ as ‘glok’), which was also the only repair strategy used for the novel-BP words from the strategies proposed by Davidson and Shaw (e.g., ‘teko’ as ‘deko’). As in the first-phoneme deletion blocks, participants in the A conditions used more repair strategies for the novel-Bosnian words than the A+O conditions across the three language groups. Within the A condition, the BT group employed the most repair strategies ($n = 167$), followed by the BO group ($n = 127$) and the M group ($n = 108$). Within the A+O condition, the M and BT group had the same number of repair strategies ($n = 65$), while the BO group had fewer instances of repair strategies ($n = 54$). The most frequently used repair strategies for each novel-Bosnian word in the second-phoneme deletion blocks is illustrated in Table 10.

The number of incorrect responses for the novel-Bosnian words was much higher in the second-phoneme deletion blocks compared to the first-phoneme deletion blocks. Interestingly, this pattern was reversed for the novel-BP words as there were fewer incorrect responses in the second-phoneme deletion blocks. Of the 136 incorrect responses for the novel-BP words, participants deleted the second syllable of the novel-BP $24.26\%$ of the time rather than the second phoneme (33 trials; $A = 30$ trials, $A+O = 3$ trials; $BO = 18$ trials, $BT = 15$ trials). Interestingly, there were 17 instances (12.5%) where participants deleted the second consonant in the CVCV structure (e.g., ‘fafe’ as ‘fae’) as opposed to the second phoneme or the second syllable ($A = 17$ trials; $M = 9$ trials, $BT = 8$ trials).
Table 10

**Most Frequently Used Repair Strategy and Change for Each Novel-Bosnian Word During Second-Phoneme Deletion**

<table>
<thead>
<tr>
<th>Repair Strategy</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>dlok</td>
<td>C1 change (67.76%)</td>
</tr>
<tr>
<td></td>
<td>‘d’ as ‘g’ (64 out of 103 trials)</td>
</tr>
<tr>
<td>dnok</td>
<td>Vowel insertion (9.21%)</td>
</tr>
<tr>
<td>zvok</td>
<td>C1 change (32.24%)</td>
</tr>
<tr>
<td></td>
<td>‘z’ as ‘s’ (48 out of 49 trials)</td>
</tr>
<tr>
<td>dvok</td>
<td>C1 change (19.74%)</td>
</tr>
<tr>
<td></td>
<td>‘d’ as ‘s’ (14 out of 30 trials)</td>
</tr>
<tr>
<td>mrok</td>
<td>C1 change (10.53%)</td>
</tr>
<tr>
<td></td>
<td>Prosthesis (10.53%)</td>
</tr>
<tr>
<td></td>
<td>a/mr (8 out of 16 trials)</td>
</tr>
<tr>
<td></td>
<td>e/mr (8 out of 16 trials)</td>
</tr>
<tr>
<td>hvok</td>
<td>C1 change (25%)</td>
</tr>
<tr>
<td></td>
<td>‘h’ as ‘a’ (12 out of 38 trials)</td>
</tr>
<tr>
<td>mnok</td>
<td>C1 change (12.5%)</td>
</tr>
<tr>
<td></td>
<td>‘m’ as ‘v’ (13 out of 19 trials)</td>
</tr>
<tr>
<td>ptok</td>
<td>2 sounds perceived as 1 (40.13%)</td>
</tr>
<tr>
<td></td>
<td>‘pt’ as ‘d’ (37 out of 61 trials)</td>
</tr>
</tbody>
</table>

*Note. Words are listed from least to most difficult based on the total number of incorrect responses across all participants.*

*Note. Percentages represent the total number of instances of that repair strategy across the total number of trials for that word (each word had a total number of 152 trials across all participants).*

### 3.5 The Link Between Word Learning and Phoneme Deletion

To determine whether there was a link between word learning and phoneme deletion, a measure of phonological awareness, a correlation was conducted between the word learning and phoneme deletion accuracy scores for the 38 participants that completed both tasks. The word learning accuracy scores were out of a possible 240 as there were 240 trials in the test phase of word learning. The ‘lenient-correct’ scores from the first- and second-phoneme deletion blocks were calculated and summed to provide an overall score with a possible range.
of zero to 128. The scores for word learning and phoneme deletion were then converted into a percentage out of 100. Across all participants, the mean word learning accuracy percentage was 86.30% ($M = 207.13$, $SD = 23.96$) and the mean phoneme deletion accuracy percentage was 84.15% ($M = 107.71$, $SD = 21.52$). A bivariate correlation test confirmed a strong, positive relationship between word learning and phoneme deletion accuracy, $r (36) = .53$, $p = .001$, suggesting that participants with greater word learning accuracy generally had greater phoneme deletion accuracy scores. The correlation is illustrated in Figure 5.

![Fig. 5. Mean percentage accuracy scores for word learning and phoneme deletion.](image-url)
4. Discussion

The present study explored the effects of orthography on learning novel words, with a focus on complex onset consonant clusters. Of particular interest was whether knowledge of two orthographic systems, as opposed to one, enhances word learning and phoneme deletion accuracy, and whether the orthographic depth of these orthographic systems influences task performance. Furthermore, the present study investigated whether simultaneous presentation of words’ orthographic and auditory forms (A+O) compared to auditory input alone (A) facilitates task accuracy. Native AusE monolinguals and bilinguals with native-like auditory proficiency in AusE as well as a LOTE (categorised into ‘transparent’ or ‘opaque’ groups based on the orthographic depth of their LOTE) were assessed on novel word pairs in the form of non-MPs (e.g., dlok-fafe), easy-MPs (e.g., dlok-mrok) and difficult-MPs (e.g., dlok-dvok), during which half of the participants in each group received the words’ orthographic forms. Participants were then assessed on their ability to manipulate the sounds of the newly learned words in a phoneme deletion task.

Despite the literature suggesting that bilinguals have enhanced phonological awareness compared to monolinguals, the hypothesis that the bilingual groups would outperform the monolinguals in word learning and phoneme deletion was not supported as all three groups had comparable scores for both tasks. The second hypothesis that orthographic input would facilitate word learning and phoneme deletion accuracy was supported as the A+O condition significantly outperformed the A condition in both tasks. While the hypothesis that participants would learn non-MPs better than MPs during word learning was supported overall (as shown by the main effect of pair type), the interaction between group, condition and pair type revealed that the facilitative effects of orthographic input varied across groups and pair types (i.e., non-MPs, easy-MPs and difficult-MPs). The hypothesis that phoneme deletion accuracy would be significantly higher for the novel-BP words
compared to the novel-Bosnian words was only partially supported, as this was only the case for second-phoneme deletion. Listeners used more repair strategies for the novel-Bosnian words compared to the novel-BP words, most frequently perceiving the two consonants in the cluster as one when instructed to delete the first phoneme, and perceptually changing the first consonant of the novel words when instructed to delete the second phoneme. A strong correlation between word learning and phoneme deletion accuracy was found ($r = .53$, $p = .001$) such that those with higher word learning accuracy scores generally had higher phoneme deletion accuracy scores.

4.1 Interpretation of Results

4.1.1 Word Learning

The similar word learning performance between the monolingual and bilingual groups is in contrast with previous studies that found a bilingual advantage for word learning (Escudero, Mulak, Fu et al., 2016; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009b; Kaushanskaya & Rechtzigel, 2012). A possible explanation for this finding may lie in the relationship between the novel words used in the present study and their corresponding visual referents. Previous studies have found a bilingual advantage for learning novel words when there is semantic information associated with them. For example, Kaushanskaya and Marian (2009b) and Kaushanskaya and Rechtzigel (2012) found that bilinguals would outperform monolinguals when the novel words were presented with translations that referred to concrete, high-frequency words in English, while the latter study found that monolinguals and bilinguals would perform similarly when the novel words had abstract English referents. As the novel words and visual referents in the present study were unfamiliar to all participants, the absence of semantic information might have resulted in the similar scores between the three language groups.
The complexity of the novel-Bosnian words could have also led to the similar scores across the three groups. Escudero, Mulak, Fu et al. (2016) found that bilinguals were able to match novel words to their corresponding novel visual referents significantly better than monolinguals, however participants were assessed on eight monosyllabic words with CVC structures that adhered to English phonotactics. In the present study, not only did participants have the more difficult task of learning 16 novel words (double the number of words in Escudero, Mulak, Fu et al., 2016), but the eight novel-Bosnian words formed monosyllabic MPs with consonant clusters that were non-native to all listeners. As there are fewer phonemic cues in monosyllabic words, and as consonant clusters are difficult linguistic structures for learners to master (Colantoni et al., 2015), all participants may have experienced similar complexities when distinguishing between the novel-Bosnian words.

The similar scores between monolinguals and bilinguals challenge the ODH which states that language learning is directly influenced by the orthographic distance between the learners’ L1/L2 orthographic systems and the language the learner is acquiring (Hamada & Koda, 2008; Ijalba & Obler, 2015). As the participants in the BO group had knowledge of two orthographically opaque languages, it is surprising that their performance was comparable to that of the BT and M groups. Therefore, an alternative explanation for the similar scores across the groups could be that native-like comprehension in two languages may not be sufficient in order to clarify the effects of orthography on word learning in bilingual populations. As a result, future research should recruit and assess bilinguals who have native proficiency in reading and writing in both of their languages, as these skills are directly influenced by orthographic processing. Further research should aim to match the bilinguals on other variables that have potentially confounding effects on the current results including a) type of bilingualism, as simultaneous bilinguals have been found to outperform monolinguals in word learning (e.g., Escudero, Mulak, Fu et al., 2016), while studies
comparing mixed-bilingual groups and monolinguals (as in the present study) have found no bilingual advantage for word learning (e.g., Escudero, Mulak & Vlach, 2016), b) age of L2 acquisition, as earlier L2 exposure leads to better word learning performance (e.g., Kaushanskaya & Rechtzigel, 2012), and c) socio-economic status (SES), as students from lower SES homes perform worse on reading and spelling tasks compared to those from higher SES homes (e.g., McDowell, Lonigan & Goldstein, 2007; Lundberg, Larsman & Strid, 2012).

The overall better performance of the A+O condition compared to the A condition supports previous studies which have also found that presenting learners with a words’ auditory and written form can facilitate word learning (Erdener & Burnham, 2005; Escudero et al., 2008; Escudero & Wanrooij, 2010; Steele, 2005). However, the fact that orthographic input only facilitated the monolinguals’ performance for difficult-MPs suggests that orthographic input may be redundant for word pairs with less phonological overlap (e.g., non-MPs and easy-MPs) that are already perceived with high accuracy (Escudero, 2015). The fact that orthographic input facilitated the bilinguals’ performance for non-MPs and easy-MPs, but not for difficult-MPs, might suggest that for learners with two existing lexicons, orthographic input might enhance differences between words that are already relatively easy to perceive (Escudero, 2015).

All participants, regardless of group and condition, performed significantly better for non-MPs compared to difficult-MPs. This finding is in line with previous studies that have found higher learning percentages for word pairs that have greater phonological differences from one another compared to word pairs that differ in fine phonetic detail (e.g., Escudero, Mulak, Fu et al., 2016; Escudero, Mulak & Vlach, 2016; Zjakic et al., 2016). The similar BT scores for the non-MPs and easy-MPs in the A+O condition could be a result of (i) the BT participants’ experience with a LOTE with similar orthographic depth to the stimuli
Orthographic Influences on Word Learning

(Kaushanskaya & Marian, 2009b), (ii) the easy-MPs differing in their two consonants as consonants play a vital role in lexical recognition (Nespor et al., 2003), and/or (iii) orthographic input enhancing the phonological differences between the words in the easy-MPs as they may have already been relatively easy for the listeners to perceive auditorily. An unexpected finding was the similar performance for non-MPs and easy-MPs for the M group in the A condition, particularly as their mean accuracy score for easy-MPs was the highest of the three language groups in the A condition. The monolinguals may have shown increased sensitivity to the phonetic properties of the easy-MPs compared to the bilinguals who would have also had to inhibit competitor words or sounds in their two lexicons.

Despite the fact that the difficult-MPs were phonologically less distinct from one another compared to the easy-MPs, the bilinguals in the A condition had similar accuracy scores for both MP types, suggesting that they may have been more sensitive to the phonetic properties of the difficult-MPs in the auditory input compared to the monolinguals. The fact that the bilinguals had to inhibit the activation of competitor words while encoding fine phonetic detail between the MPs could also explain the similar scores between the two MP types. On the other hand, the bilinguals’ significantly better performance for easy-MPs compared to difficult-MPs in the A+O condition might suggest that orthographic input facilitates learning of similar word pairs but in this case, only when the words differed in both of their consonants in the cluster (e.g., mrok-dlok as opposed to mrok-mnok). The monolinguals’ significantly higher easy-MP scores compared to difficult-MPs in the A condition, and similar scores between the easy- and difficult-MPs in the A+O condition lends further support to the facilitative effect of orthographic input on monolinguals’ discrimination of difficult word pairs.
4.1.2 Phoneme Deletion

The similar phoneme deletion performance between the monolinguals and the two bilingual groups is in contrast with previous studies that found bilinguals with orthographically similar L1s and L2s perform significantly better on phonological awareness tasks compared to monolinguals and bilinguals with orthographically dissimilar L1s and L2s (e.g., Bialystok et al., 2005: Hebrew-English and Spanish-English bilinguals outperformed English monolinguals and Chinese-English bilinguals; Bialystok et al., 2003: Spanish-English bilinguals outperformed English monolinguals and Chinese-English bilinguals; Wade-Woolley, 1999: Russian-English bilinguals outperformed Japanese-English bilinguals; Wang et al., 2003; Korean-English bilinguals outperformed Chinese-English bilinguals). As Spanish-English bilinguals have previously been found to perform significantly better on phonological awareness tasks compared to monolinguals despite the fact that the two languages have different orthographic transparencies (i.e., English being orthographically opaque and Spanish being orthographically transparent), the comparable phoneme deletion results between the BT and M groups could have been a result of the BT participants having minimal experience with their LOTE orthography. As a result, the BO group should have performed worse than the BT group due to (i) their experience with two orthographically opaque languages, and/or (ii) the BO participants’ LOTE being orthographically different to English (e.g., Arabic and Urdu use consonantal writing systems, and Cantonese uses a logographic writing system) which has previously led to lower performance on phonological awareness tasks. Therefore, the similar performance between the BO and BT groups suggests that participants may have utilised their phonological awareness skills in English, which would have facilitated their phoneme deletion accuracy as English shares the same alphabetic script as the novel words in the study. Thus, like for the word learning results discussed above, further research is necessary on bilinguals with higher
reading and writing proficiencies in their LOTE in order to clarify the effects of orthography on phonological awareness skills in bilingual populations.

The overall facilitative effect of orthographic input during the phoneme deletion task shown by the low frequency of repair strategies used, is in contrast with previous studies that have found orthographic input to lead to greater use of repair strategies for consonant clusters compared to auditory input alone (e.g., Davidson, 2007; Detey & Nespoulous, 2008; Young-Scholten, 1998; Young-Scholten et al., 1999). Davidson (2007) found that English listeners would add an epenthetic vowel between non-native fricative-initial consonant clusters less than 15% of the time when presented with orthographic input, which was significantly less than the epenthesis rate of 67% found for Japanese listeners when parsing consonant clusters in novel French words (Detey & Nespoulous, 2008). As the presentation of orthographic input during word learning significantly benefitted participants' phoneme deletion accuracy across the three language groups in the present study, it could be the case that orthographic input has varying effects on listeners depending on their L1 orthographic writing system. As Japanese has a syllabic writing system with a dominant CV syllable, the Japanese listeners in Detey & Nespoulous (2008) may have been inclined to insert epenthetic vowels between the consonant clusters even when presented with orthographic input. On the other hand, the facilitative effects of orthographic input in the present study may have been more evident for English listeners as English allows consonant clusters in onset position and uses an alphabetic script (the same as the novel-Bosnian and novel-BP words in the stimuli).

Participants’ higher first-phoneme deletion scores for the novel-Bosnian words compared to the novel-BP words was surprising as previous studies investigating phonological awareness in children have found that consonants in clusters are harder to identify than consonants that are onsets on their own (Treiman, 1985; 1991). Furthermore, as the consonant clusters in the present study were non-native to all listeners, they should have
elicited more instances of prosthesis (e.g., ‘mrok’ – ‘emrok’) which would have led to a greater number of incorrect responses (Davidson & Shaw, 2012). The striking finding that participants would delete the first syllable of the novel-BP words (rather than the first sound) 71% of the time suggests a preference for CV as an onset syllable compared to VC, lending further support to the notion that cross-linguistically, the unmarked CV syllable is the most preferred linguistic structure (Ventura, Kolinsky, Brito-Mendes & Morais, 2001).

Alternatively, as participants received both oral and written instructions for the task, perhaps another explanation for the high instance of syllable deletion is that the term ‘sound’ may have been too ambiguous and confusing for participants. The fact that the A+O condition had significantly higher accuracy scores for first-phoneme deletion of the novel-BP words compared to the A condition further reinforces the facilitative effects of orthographic input.

As it is widely known that listeners will employ perceptual modifications to bring non-native consonant clusters in line with their L1 phonotactics (Colantoni et al., 2015), it is not surprising that there were a greater number of repair strategies used for the novel-Bosnian words compared to the novel-BP words. Participants would most often perceive the two consonants in the cluster as one sound (e.g., ‘ptok’ as ‘tok’) which is in line with Colantoni et al. (2015) who state that consonant deletion is one of the main types of repair strategies employed by listeners. This repair strategy was most frequently used for each novel-Bosnian word apart from ‘mrok’ for which prosthesis was most frequently used. The low number of incorrect responses for ‘dlok’ and ‘dnok’ during the phoneme deletion task can be explained by the sonority sequencing principle that states that most often, sonority increases from the onset and coda towards the nucleus (Colantoni et al., 2015). Therefore, although the clusters violated English phonotactics, listeners may have found them easier to perceive compared to words with onsets that allow sonority plateaux (i.e., the two consonants in the cluster have equal sonority, e.g., /mn/ in ‘mnok’ are both nasals which may have created perceptual
difficulties for listeners). During the first-phoneme deletion blocks, the word ‘mrok’ received
the lowest number of correct responses supporting the notion that consonants within clusters
are more difficult to perceive when they are closer together in the sonority scale (Colantoni et
al., 2015; Moreton, 2002). The most unexpected finding was that the stop-stop cluster in the
word ‘ptok’ received the second highest number of correct responses.

Participants’ most frequently used repair strategy during the second-phoneme deletion
blocks was C1 change, however, this varied across the individual novel-Bosnian words. The
/dl/ cluster in ‘dlok’ was most frequently perceptually modified to /gl/, which is a common
onset of English words. This finding is in line with Halle et al. (1998) who also found that
French speakers would perceptually change /dl/ to /gl/. Listeners would insert an epenthetic
vowel between the /dn/ cluster in ‘dnok’, which is in contrast to Davidson and Shaw’s (2012)
finding that listeners would confuse stop-nasal sequences with a change or deletion of the
first consonant. The remaining six novel-Bosnian words elicited a larger number of incorrect
responses, providing further evidence that consonant clusters are harder to perceive when
they are closer together on the sonority scale or when they are positioned in the same place
on the sonority scale. The /zv/ and /dv/ in ‘zvok’ and ‘dvok’ were both frequently confused
with /sv/. The /z – s/ modification could be explained by the fact that both sounds are denti-
alveolar fricatives. Alternatively, listeners may have perceived the /v/ as the /f/ sound in
words such as ‘sphere’ or ‘sphinx’, and therefore confused the /z/ for an /s/. On the other
hand, the /d – s/ modification is puzzling. One possible explanation could be that participants
who perceived ‘zvok’ as ‘svok’ may have been primed to instantly perceive /s/ when
presented with ‘-vok’ regardless of whether the word was ‘zvok’ or ‘dvok’. Other unusual
findings were the frequent changes of /mr/ to /vr/ in ‘mrok’ as well as /mn/ to /vn/ in ‘mnok’
considering that /vr/ and /vn/ are not acceptable onsets in English words. On the other hand,
the frequent change from /hv/ to /av/ in the word ‘hvok’ is evidence of L1-based CLI as the
sequence /av/ is common in word-initial position in English. Furthermore, as /h/ is a non-
sonorous consonant, listeners may have easily confused it with a vowel. The change of /pt/ to
/d/ in the word ‘ptok’ challenges Davidson and Shaw (2012) who found that stop-stop
sequences are frequently confused with vowel insertion. The differences between Davidson
and Shaw’s findings and the findings of the present study could be a result of the specific
clusters used as well as the experimental paradigm employed in the current study. As
Davidson and Shaw (2012) used a discrimination task for their study, future avenues for the
present study might be to assess participants on a phoneme deletion task and then an
identification task to examine whether their perceptual modifications vary across tasks for
specific clusters. Furthermore, as English was not the L1 for all participants, further research
should investigate whether certain perceptual modifications observed in present study (e.g.,
the frequent change of the onset /mr/ to /vr/, which also violates English phonotactics) are
native to the participants’ LOTE.

### 4.1.3 The Link Between Word Learning and Phoneme Deletion

Studies have previously found strong correlations between phonological awareness
and reading (e.g., Hogan, Catts & Little, 2005; Treiman, 1991), as well as spelling (e.g.,
Bruck & Treiman, 1990; Treiman, 1991), particularly for learners of alphabetic languages
such as English. The present study extends on this concept by demonstrating that learners
who can distinguish between novel words that differ only in fine phonetic detail can also
成功fully manipulate the sounds of these newly learned words, even if the words contain
consonant clusters that are non-native to them.

More interestingly, the results show similar patterns for the effects of orthography on
the acquisition and perception of non-native consonant clusters across both experimental
tasks. Firstly, participants, regardless of whether they were monolingual or bilingual, seemed
to utilise their knowledge of English to complete both tasks. As previously mentioned, the
bilingual participants had low reading and writing proficiency in their LOTE which may have made it an easier task for the BO group, in particular, to suppress any knowledge of their LOTE orthographic system in order to successfully learn the novel words. Secondly, orthographic input facilitated task performance compared to auditory input alone for both word learning and phoneme deletion. Finally, the novel-Bosnian words in CCVC form presented difficulties for participants across both tasks as highlighted by the low accuracy scores for the difficult-MPs compared to the non-MPs during word learning, and the increased number of repair strategies employed by the participants to bring the non-native structures in line with their native phonotactics during phoneme deletion.

Overall, the present study not only highlights the dynamic system of language learning, but also the complex interaction between orthography and perception. The results can inform existing models that describe the interaction between orthography and perception (e.g., Hamann & Colombo, 2017), extending these models to account for cases when complex consonant clusters are involved and when learners are exposed to multiple writing systems. Furthermore, investigating the effects of orthography on language learning can further enhance our understanding of L1 orthographic influences on all areas of L2 phonology such as perception, production and the acquisition of L2 sounds and words.

### 4.2 Implications

Based on the findings of the present study, potential practical implications for L2 learning can be inferred. Firstly, investigating orthographic effects on adult learners is particularly important as their language input is typically in written form and is often re-interpreted according to their L1 orthographic features, unlike infant L2 learners who are predominantly exposed to auditory input (Bassetti, 2008).

The facilitative effect of orthographic input on word learning and phoneme deletion performance suggests that teachers should be incorporating the spelling of target language
words into vocabulary instruction during the early stages of L2 acquisition, particularly for words containing complex sound structures such as consonant clusters (Ehri, 2014). Additionally, the relationship found between word learning and phoneme deletion accuracy in the current study suggests that the implementation of phonological awareness training in instructed contexts could enhance L2 word learning abilities, as well as L2 reading and writing development. Finally, the large number of repair strategies used across participants in the phoneme deletion task because of their L1 orthography suggests that teaching learners the orthographic similarities and differences between their L1 and L2 could potentially ease their learning process (Escudero & Wanrooij, 2010).
References


doi: 10.1179/136132808805297214


doi: 10.1016/j.esp.2017.03.001


doi: 10.1017/S1366728916000249


Spencer, K. (2007). Predicting children's word-spelling difficulty for common English words from measures of orthographic transparency, phonemic and graphemic length and
doi: 10.1348/000712606X123002


doi: 10.1016/0022-0965(85)90034-7


doi: 10.3389/fpsyg.2015.01000


APPENDIX 1

Participant Language Background Questionnaire

<table>
<thead>
<tr>
<th>Participant information – To be filled out by the researcher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project code: _______________  Today’s date: _______________  Participant code: _______________</td>
</tr>
</tbody>
</table>

1. Demographic information

Date of birth _______________.
Which hand do you use to write with (circle the one that best describes you):
- always right
- usually right
- ambidexterous
- usually left
- always left

Do/did you have a hearing impairment or reading difficulties (e.g., difficulties learning to read), or language development or speaking difficulties (e.g., delayed language development, stuttering, lisping, etc.)?
- Yes
- No

If yes, circle any/all that apply:
- hearing
- reading
- speaking
- language

Please tell us every place you have lived for at least three months or more, starting with your infancy and childhood. If you lived in a city, please also indicate which area or town/suburb:

<table>
<thead>
<tr>
<th>Country</th>
<th>City/Town/Region</th>
<th>Between which ages?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.1 Parent/Caregivers’ information

Parent/Primary caregiver (from birth to 24 months): _______________.

Native language(s)/accent(s): _______________.

Other language(s)/accent(s): _______________.

During your infancy, their language(s) and/or accent(s) of communication:

- with you: _______________.
- in the home: _______________.
- with other family members: _______________.
- with people outside the home: _______________.

Place of birth: _______________.

Current residence: _______________.

Please list all places where this caregiver has lived, when they lived there, and for how long, in chronological order:

What is the highest level of education they have completed?

Less than high school  high school  TAFE/college degree  bachelor’s degree  master’s degree  PhD

Parent/Other primary caregiver (from birth to 24 months):

Native language(s)/accent(s):

Other language(s)/accent(s):

During your infancy, their language(s) and/or accent(s) of communication:

- with you: 
- in the home: 
- with other family members: 
- with people outside the home: 

Place of birth: Current residence: 

Please list all places where this caregiver has lived, when they lived there, and for how long, in chronological order:

What is the highest level of education they have completed?

Less than high school  high school  TAFE/college degree  bachelor’s degree  master’s degree  PhD

Other caregiver(s) (from birth to 24 months):

Native language(s)/accent(s):

Other language(s)/accent(s):

During your infancy, their language(s) and/or accent(s) of communication:

- with you: 
- in the home: 
- with other family members: 
- with people outside the home: 

Place of birth: Current residence: 

Please list all places where this caregiver has lived, when they lived there, and for how long, in chronological order:

What is the highest level of education they have completed?

Less than high school  high school  TAFE/college degree  bachelor’s degree  master’s degree  PhD

Please list these caregivers in order of amount of time you spent with them. Also indicate approximately how much time you spent with each of them during a typical week (from birth to 24 months).

1. __________________________ Approx. time spent per week: __________________________
2. __________________________ Approx. time spent per week: __________________________
3. __________________________ Approx. time spent per week: __________________________
2.2 Languages/Accents in your environment during infancy (0 to 24 months of age inclusively)

Please list all the languages that you were exposed to during your infancy and, for each language, the approximate percentage of the time that you heard it on a weekly basis. Note: This should add up to 100%.

<table>
<thead>
<tr>
<th>Language(s)/Accent(s)</th>
<th>Percentage of the time that you heard this language/accent on a weekly basis DURING INFANCY:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Current Language/Accent Proficiency

Please evaluate your current level of proficiency for all languages/accents that you have ever been exposed to:

<table>
<thead>
<tr>
<th>Language/Accent: ________________________________</th>
<th>Oral comprehension:</th>
<th>Oral production:</th>
<th>Writing proficiency:</th>
<th>Reading proficiency:</th>
<th>Pronunciation:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>very low</td>
<td>low</td>
<td>intermediate</td>
<td>advanced</td>
<td>near-native</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
If you have been exposed to more than the above 5 languages/accents, please list the others here and comment on your general proficiency for each:

<table>
<thead>
<tr>
<th>Language/Accent:</th>
<th>______________________________</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral comprehension:</td>
<td>very low low intermediate advanced near-native native</td>
</tr>
<tr>
<td>Oral production:</td>
<td>very low low intermediate advanced near-native native</td>
</tr>
<tr>
<td>Writing proficiency:</td>
<td>very low low intermediate advanced near-native native</td>
</tr>
<tr>
<td>Reading proficiency:</td>
<td>very low low intermediate advanced near-native native</td>
</tr>
<tr>
<td>Pronunciation:</td>
<td>very accented accented slightly accented near-native native</td>
</tr>
</tbody>
</table>
## 4. Evolution of Language/Accent Use

1. In the first row, fill in the ages, grades, or calendar years corresponding to the education level specified on top of each column. **Please ask the researcher if you need help.**

2. In each cell, use percentages to indicate your usage of/exposure to Australian English and other languages/accents (combined) for the corresponding context and age.

Note: If your language/accent use changed within these age groups, or if the age group is inaccurate with respect to education level, please specify by writing it inside the box or explain in the comments sections below. If you have any questions, please speak to your experimenter.

<table>
<thead>
<tr>
<th>Ages/grades/calendar years</th>
<th>Age 2 to kindergarten</th>
<th>Kindergarten</th>
<th>Early elementary school (Year 1-3)</th>
<th>Late elementary school (Year 4-6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>School:</strong> i.e. language of instruction. If language immersion, please specify in comments below which grades and number of hours/week</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
</tr>
<tr>
<td><strong>At home:</strong> interactions with immediate and extended family, significant other (if you lived with them) and roommates</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
</tr>
<tr>
<td><strong>Friends:</strong> interactions with friends and significant other (if you did not live with them)</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
</tr>
<tr>
<td><strong>Media use:</strong> social media, leisurely reading, television, cinema, radio, internet, music, etc.</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
</tr>
<tr>
<td><strong>Extracurricular activities:</strong> sports, hobbies, work (if part-time), etc.</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
</tr>
<tr>
<td><strong>Daily activities in the community:</strong> grocery store, shopping mall, restaurants, gas station, etc.</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
</tr>
<tr>
<td><strong>Other:</strong></td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
<td>AusEng: other</td>
</tr>
<tr>
<td>High school</td>
<td>College</td>
<td>University</td>
<td>Other: ________</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---------</td>
<td>------------</td>
<td>----------------</td>
<td></td>
</tr>
</tbody>
</table>

### Ages/grades/calendar years

### School: i.e. language of instruction. If language immersion, please specify in comments below which grades and number of hours/week
- **AusEng:**
- **other:**

### At home: interactions with immediate and extended family, significant other (if you lived with them) and roommates
- **AusEng:**
- **other:**

### Friends: interactions with friends and significant other (if you did not live with them)
- **AusEng:**
- **other:**

### Media use: social media, leisurely reading, television, cinema, radio, internet, music, etc.
- **AusEng:**
- **other:**

### Extracurricular activities: sports, hobbies, work (if less than 20h/week), etc.
- **AusEng:**
- **other:**

### Daily activities in the community: grocery store, shopping mall, restaurants, gas station, etc.
- **AusEng:**
- **other:**

### Work: (if over 20h/week)
- **AusEng:**
- **other:**

### Other:
- **AusEng:**
- **other:**
4. Current Language/Accent Use

Please use the following table to indicate your usual daily activities and the language(s)/accent(s) that you hear and use when involved in these activities. For instance, your items may include: School – Australian English, Grandparents visit – Spanish, Work – British English, etc

<table>
<thead>
<tr>
<th>Time</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
<th>Sunday</th>
</tr>
</thead>
<tbody>
<tr>
<td>8am – 10am</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10am – 12pm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12m – 2pm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2pm – 4pm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4pm – 6pm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6pm – 8pm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8pm – 10pm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Thank you!