

JOURNAL ARTICLE

# Templaticity Effects on Differential Processing of Consonants and Vowels

Maya L. Barzilai

Department of Linguistics, Georgetown University, Washington, DC, US  
mlb290@georgetown.edu

The aim of this study is to investigate the role of morphophonology and written representations in speech processing. Results are presented from an immediate serial recall (ISR) study, designed to determine the relative effects of L1 morphology and orthography on the recall of consonants versus vowels. Forty-five speakers of English, Amharic, and Arabic completed an ISR experiment testing the differential recall of these two segment types. English speakers remembered sequences of syllables in which the consonant is held constant and the vowel changes (e.g., “ma mi mu ...”) better than sequences in which the vowel is held constant and the consonant changes (e.g., “ka ma za ...”), whereas Arabic speakers remembered both types of sequences equally well, replicating findings from Kissling (2012). Crucially, Amharic speakers in this study performed similarly to Arabic speakers, remembering both sequence types with equal accuracy. Given that Amharic and Arabic share a templatic morphological system, this new result suggests that the morphophonology of a listener’s L1 impacts ISR. English and Amharic share similar orthographic systems; the mean accuracies of English and Amharic speakers were significantly different, and it therefore appears that orthography of the L1 does not affect recall accuracy. The results have implications for the role of the morphophonology of a given speaker’s L1 in ISR and in speech processing more generally.

**Keywords:** Templatic morphology; Writing systems; Semitic linguistics; Orthographic processing; Phonological processing; Consonants; Vowels

## 1. Introduction

Work in various linguistic subfields has shown that consonants have a privileged role in the grammar. Given that consonants tend to convey more lexical information than vowels (Nespor, Peña, & Mehler, 2003), it is not surprising that they have a stronger effect on priming and lexical decision tasks, and that vowels are in turn more easily manipulated or ignored in certain psycholinguistic tasks (Cutler, Sebastián-Gallés, Soler-Vilageliu, & Van Ooijen, 2000; Delle Luche et al., 2014; New, Araújo, & Nazzi, 2008; Toro, Nespor, Mehler, & Bonatti, 2008). However, it has also been shown that vowels, likely due to their relatively high perceptual salience (Crowder, 1971; Cutler et al., 2000), are more easily remembered in ISR tasks (Crowder, 1971; Drewnowski, 1980; Kissling, 2012). Therefore, while consonants are more important in some psycholinguistic tasks, vowels win out in others.

This paper hypothesizes that the morphological system of a speaker’s L1, as well as its orthography, can interact with the inherent, universal properties of consonants and vowels to impact ISR results. To this end, an ISR experiment was conducted with speakers of English, Amharic, and Arabic, with the aim of determining whether it is the morphology or orthography of a speaker’s L1 that is more likely to impact the processing of consonants versus vowels in ISR.

## 2. Background

Phonological patterns across the world's languages show that consonants and vowels comprise different entities in the grammar. Vowels tend to be louder than consonants are, and to have a longer duration than consonants do (Crowder, 1971; Cutler et al., 2000). However, there is evidence that it is not just this high sonority of vowels that distinguishes them from consonants in phonological systems. Neuropsychological data from speakers with aphasia show that consonants and vowels form distinct mental categories that can be damaged independently of one another. Caramazza, Chialant, Capasso, and Miceli (2000) show that when repeating words, one of two Italian speakers with aphasia was more likely to produce errors in repeating vowels whereas the other was more likely to confuse consonants. Furthermore, the consonantal category for the second speaker contained all consonants in the Italian inventory, regardless of their relative sonorities. This implies that the separation of consonants and vowels in the grammar is an abstract categorization rather than one based solely on phonetic properties. As expected given their psychological and abstract phonological differences, these two categories are not equivalent with respect to their roles in linguistic systems. Nespor et al. (2003) use typological data as well as results from various experimental studies to argue the CV hypothesis, stating the cross-linguistic tendency for consonants to be responsible for conveying lexical information, and for vowels conversely to encode morphosyntactic information.

Results from various psycholinguistic experiments further bolster the CV hypothesis, showing that it may best account for some aspects of phonological processing. Despite the high acoustic salience of vowels, words have been shown to be primed by non-words with the same consonants, but not by those with the same vowel, both in auditory (Delle Luche et al., 2014) and in visual priming experiments (New et al., 2008). While monosyllabic CVC words can be primed by non-words that share a rime ( $\_VC$ ) more so than they are by non-words that share an onset and nucleus ( $CV\_$ ), it has been shown that primes with only the same consonants as the target word ( $C\_C$ ) lead to even greater facilitation than rime primes in auditory lexical decision tasks (Turnbull & Peperkamp, 2018). Just as the presence of consonants can improve lexical decision results, their absence can also worsen them: Delaying the appearance of consonant graphemes slowed the reaction time for lexical decisions whereas delaying the appearance of vowel graphemes has no such effect in a study by Carreiras, Gillon-Dowens, Vergara, and Perea (2009). Similarly, speakers found it easier to change a non-word into an actual word by changing the vowels (i.e., *kebra* → *cobra*) than by changing the consonants (i.e., *kebra* → *zebra*) in a word reconstruction experiment (Cutler et al., 2000). The prominent role of consonants in phonological processing is evident even in experiments using entirely nonsense speech strings; Toro et al. (2008) show that participants use statistical dependencies across consonants, but not across vowels, to segment a nonsense CV speech stream into words.<sup>1</sup> The results of experimental studies such as these suggest that consonants' role of carrying lexical information, as put forward by Nespor et al. (2003), makes them more easily processed and less easily manipulated than vowels.

Results from ISR experiments, though, complicate this conclusion. When English speakers are tasked with remembering sequences of CV syllables, they tend to remember differences in vowels better than differences in consonants, whether the sequences are presented visually (Drewnowski, 1980) or auditorily (Crowder, 1971; Kissling, 2012). With ISR tasks, then, it seems that the effect of the high acoustic salience of vowels outweighs the effect of consonants carrying important lexical information. As a result,

---

<sup>1</sup> cf. Newport and Aslin (2004), in which non-adjacent vowel dependencies are acquired as easily as non-adjacent consonant dependencies.

vowels have the advantage over consonants in these tasks. If it is the case that ISR tasks do not involve accessing a phonological representation of the CV phonemes, but rather short-term memory simply of the acoustic properties of the stimulus, then these findings are not necessarily at odds with the intuition that “consonants have an overall privileged role over vowels at the phonological level” (Delle Luche et al., 2014). Rather, it could be argued that tasks requiring the accessing of abstract phonological categories make use of consonants and their ability to distinguish among lexical items, whereas tasks such as ISR instead favor vowels due to their high acoustic salience.

Perhaps unsurprisingly, certain characteristics of a speaker’s native language impact their performance on tasks that require phonological categorization or manipulation. In fact, though the ratio of consonants to vowels in the phoneme inventory of a language has been shown not to impact the effect of segmental priming (Cutler et al., 2000; Delle Luche et al., 2014), there is evidence that the phonotactics of a speaker’s native language govern speech perception and segmentation. Dupoux, Kakehi, Hirose, Pallier, and Mehler (1999) show that Japanese speakers perceive an epenthetic vowel between consecutive consonants in VCCV stimuli, implying that Japanese phonotactics impacted speech perception such that a vowel with no acoustic correlates in the stimulus was in fact perceived. Similarly, El Aissati, McQueen, and Cutler (2012) show that speakers of Tarifiyt Berber segmented nonce speech streams into vowelless non-words, which are phonotactically permissible in Tarifiyt but are otherwise shown to be dispreferred by speakers of other languages performing similar tasks. In other words, the phonological properties of consonants and vowels, and the way that they combine in a speaker’s language, impact that speaker’s performance on psycholinguistic tasks that require the use of phonological knowledge. However, results from Kissling (2012) reveal that even in ISR tasks, which arguably seem to call only on processing at the phonetic level, certain properties of a speaker’s native language can impact recall accuracy. Kissling (2012) shows that the result that vowels are recalled more accurately than consonants does not hold for speakers of all languages. In that study, native speakers of English and Arabic were presented with auditory stimulus sequences of six CV syllables. In each sequence, the syllables either had the same consonant and different vowels (e.g., “ki ka ki ku ku ka”) or the same vowel and different consonants (e.g., “ma za ka za ka ma”). After hearing each stimulus, participants had 12 seconds to record the six syllables they had heard on an answer sheet containing six blanks. Results showed that while English speakers remembered the sequences with different vowels better, replicating previous findings (Crowder, 1971; Drewnowski, 1980), Arabic speakers scored similarly on the two types of sequences.

There are two possible explanations for the surprising results in Kissling (2012). One explanation is that better consonant recall by Arabic speakers is a result of the morphological system of Arabic, a Semitic language exhibiting templatic morphology (Ryding, 2005). In this morphophonological system, consonants are exclusively responsible for conveying lexical information, whereas in non-templatic languages consonants are merely more likely to carry lexical content (Nespor et al., 2003; Toro et al., 2008). Because of this patterning, consonants have been argued to be more linguistically salient than vowels in Arabic and other languages exhibiting templatic morphology (Katz & Frost, 1992). Therefore, it is possible that native Arabic speakers attend to and remember consonants as well as they do vowels as a result of the morphological templaticity in Arabic. If this theory holds, it would argue against the notion that ISR tasks require only low-level phonetic processing and are not impacted by phonological categories, as entertained above.

Another possibility is that the results in Kissling (2012) surface due to an orthography effect. Before discussing the details of Arabic’s orthographic system, it is worth mentioning that the term ‘orthography’ here refers to the rate at which different segment

types (consonants versus vowels) are encoded in the writing system and, conversely, to what type of segment each grapheme corresponds. Not discussed here are issues such as grapheme-to-phoneme regularities and other similar questions often discussed as properties of a given orthographic system. Therefore, though Spanish and English orthographies, for example, differ in several writing conventions, they would be treated here as having comparable orthographies as they share a common alphabet and most graphemes correspond to the same segment type in both writing systems. This notion of orthographic similarity is used here throughout.

Arabic's abjad orthography encodes only consonants and long vowels consistently.<sup>2</sup> For instance, the words *kataba* ('wrote') and *kutiba* ('was written') are orthographically identical, with only the <k>, <t>, and <b> graphemes; readers must infer the short vowels based on the morphosyntactic context. Therefore, it is possible that this creates a perceptual bias towards consonants, which in turn explains the boosted consonant recall among Arabic speakers. Under this explanation, the results in Kissling (2012) would not be explained by whether or not phonological information is accessed in ISR, but rather by an additional level of processing or representation introduced by orthographical knowledge.

Research showing the impact of orthography on phonological representations supports the possibility of an orthography effect in ISR tasks. For instance, it has been shown that literate adults perform phoneme manipulation tasks more successfully than illiterate adults with otherwise comparable linguistic experience, upbringings, and environments (Morais, Bertelson, Cary, & Alegria, 1986; Morais, Cary, Alegria, & Bertelson, 1979). Detey and Nespoulous (2008) show that Japanese learners of French were better able to correctly syllabify French words with consonant clusters that are illicit in Japanese if the words were presented either only orthographically or both orthographically and auditorily. The interaction of orthography and phonology is apparent in loanword adaptation as well; English words written with a double consonant (e.g., *splatter*) are more likely to be borrowed into Italian with a phonemic geminate than those without (Hamann & Colombo, 2017). Importantly, different orthographic systems can have different impacts on speech perception and phonological representations in general. For instance, Frost, Katz, and Bentin (1987) show that three different orthographic systems with distinct levels of 'orthographic depth' have different impacts on lexical access. They conclude that speakers of languages with transparent ('orthographically shallow') orthographic systems can derive phonological representations directly from the corresponding orthographic representations, but that speakers of languages for which the grapheme-to-phoneme correspondence is not one-to-one ('orthographically deep') cannot. Given the varied findings on the effects of orthographic knowledge on phonological representation, it is reasonable to posit that an orthographic effect may best explain the findings in Kissling (2012). Both this orthographic effect and a morphological effect are plausible explanations for the results found among Arabic speakers, and the current literature has no basis on which to claim which one is more likely to be responsible.

Like Arabic, Amharic is a Semitic language with templatic morphology (Leslau, 1995). However, unlike Arabic, Amharic is written with the Ge'ez script, called Fidel, an abugida in which each grapheme encodes one consonant and a following vowel (Leslau, 1995). Therefore, Amharic is an ideal test case to employ in order to disentangle the potential morphological and orthographic effects posited in previous research. To this end, the

<sup>2</sup> Short vowels are optionally encoded in Arabic writing by diacritics above or below the consonant graphemes. However, the inclusion of these short vowel diacritics is relatively rare in most writing contexts, and therefore it is assumed here that short vowels are effectively absent from orthographic representations. It is worth noting that gemination is also represented in Arabic orthography by an optional diacritic over the consonant grapheme. If it is assumed that geminate consonants are two consecutive identical consonants, then some consonants, and not only short vowels, are also often absent from Arabic orthography.

present study builds on the work in Kissling (2012), testing the recall of consonants versus vowels among speakers of Arabic and Amharic. A third experimental group is comprised of native speakers of English, for the purpose of a comparison against a non-templatic language with an alphabetic orthography. Though Fidel differs from the English alphabet in meaningful ways, as discussed below in Section 5.3, the two orthographies are equivalent for the purposes of this study because they both encode consonants and vowels with equal regularity. By comparing results from these three groups of speakers, this study aims to determine whether segmental processing is affected by the morphology or the orthography of a speaker's native language. If Amharic speakers exhibit the same recall behaviors as Arabic speakers, this would support the hypothesis that morphology has a greater impact on segmental processing in ISR; conversely, similar recall between Amharic and English speakers would support the hypothesis that orthography is responsible for the language differences. The results from this study will provide greater insight into how the morphophonological and orthographic properties of the L1 can interact to influence the way sounds are processed and stored in short-term memory.

### **3. Methods**

#### **3.1. Participants**

Native speakers of English (7 female, 15 male), Amharic (8 female, 1 male), and Arabic (6 female, 8 male) participated in the present study. All participants were over the age of 18, and were students, employees, or faculty members at an American university who reported being literate in their L1. Though each participant was at least moderately proficient in English and used it in their daily life, the Amharic and Arabic speakers were all dominant in respective L1s. Participants were recruited via e-mail and were not compensated for their participation. The difference in participant sample sizes for the three languages reflects the imbalance in speaker populations available at the university.

#### **3.2. Materials**

The stimuli in the experiment were sequences of 6 CV syllables, comprised of the segments /m k z i u a/. All of these segments are phonemic in each of the languages tested. Each of the 9 possible CV syllables generated from this inventory was recorded once by a female native speaker of American English. The syllables were then concatenated into sequences that were either consonant-variable with a constant vowel (e.g., "mi ki zi zi mi ki") or vowel-variable with a constant consonant (e.g., "ma mi mu ma mu mi"). Syllables appeared one to three times in a given sequence; this rule was an effort to prevent speakers from inferring a pattern in which a sequence consisted of exactly two occurrences each of three syllables. The stimulus sequences were those employed in the methodology of Kissling (2012).

#### **3.3. Procedure**

All portions of the experiment were conducted in a quiet room on the university's campus, and all stimuli and responses were recorded on an H4N Zoom Recorder. The first portion of the experiment was a training phase comprised of two stimulus sequences, neither of which was repeated during the testing phase of the experiment. The researcher was present in the room during the training phase to ensure that the participant understood the procedure. After the training period, the researcher left the room and the participant was instructed to press a key on the computer to begin the testing period. There were 23 sequences in total, and all sequences were randomized for each participant in the testing period.

Stimulus sequences were presented auditorily via PsychoPy (Peirce, 2007) on a MacBook Air laptop. Each sequence was approximately 7 seconds in duration, and was played while the computer screen was gray. Approximately 1500 ms after the end of the stimulus, the screen turned blue; the participants were instructed to repeat the sequence they had just

heard, to the best of their ability, once the screen was blue. After 8 seconds of response time, the screen turned gray again and the next sequence played automatically. This procedure continued until all stimuli were tested one time.

### 3.4. Scoring

The recordings of the stimuli and responses were transcribed and coded auditorily by a native speaker of American English. The coder recorded the correct syllables for each sequence, followed by the syllables produced by each participant. For each syllable in a stimulus, a score of 1 was awarded if the syllable was correctly repeated and 0 if the participant produced an incorrect syllable or no corresponding syllable at all. Though speakers occasionally produced fewer than six syllables in their response, there were no cases in which the identity of either the consonant or the vowel in a produced syllable was difficult to hear or identify and therefore not recorded.

Each stimulus sequence also received an overall accuracy score. Three coding systems were used to determine each stimulus score, as detailed below: the ‘aligned’ scoring system, the ‘strict’ system, and the ‘partially aligned’ system.

In the aligned scoring system, each sequence could receive a total of six possible points, one for each syllable in the sequence. For responses in which the participant did not produce exactly six syllables, the syllables produced were aligned to best fit the stimulus sequence, when possible. An example of such a response is shown in (1).

- (1) Example stimulus and response with < 6 syllables

**stimulus sequence:** ka ku ki ki ka ku  
**response:** ka ku ka ku

- (2) Aligned scoring for response in (1)

<b>stimulus sequence</b>	ka	ku	ki	ki	ka	ku	
<b>raw score</b>	ka	ku	ka	ku	–	–	2 points
<b>aligned score</b>	ka	ku	–	–	ka	ku	4 points

Example (1) shows the syllable alignment for a response with only four syllables. The raw response yields a total of two out of six points, as only the first two syllables were correctly reproduced. However, aligning the final two syllables of the raw response to those of the stimulus sequence results in a total of four out of six points, two for the first two syllables and two for the final two syllables produced.

For responses longer than six syllables in length, the final syllables produced were moved leftward in cases in which this would improve the score. An example of a response more than six syllables long is shown in (3).

- (3) Example stimulus and response with > 6 syllables

**stimulus sequence:** ka ku ki ki ka ku  
**response:** ka ku ki ku ku ka ku

- (4) Aligned scoring for response in (3)

<b>stimulus sequence</b>	ka	ku	ki	ki	ka	ku	
<b>raw score</b>	ka	ku	ki	ku	ku	ka	(ku) 3 points
<b>aligned score</b>	ka	ku	ki	ku	ka	ku	5 points

Example (4) shows the response syllable alignment for a response with seven syllables. The raw response has three correct syllables, followed by three incorrect syllables (in italics) and a seventh extraneous syllable (in parentheses). Moving the sixth and seventh syllable leftwards by one slot each to replace the fifth and sixth syllable produced, respectively, makes for a response in which the first three syllables, as well as the fifth and sixth, are correct. In this case, alignment brings the score from three to five points.

In the strict scoring system, each response received one point if and only if the response was identical to the stimulus sequence. Otherwise, the response received zero points. If more than six syllables were produced, the response was awarded a point if the first six syllables were identical to the six stimulus syllables; otherwise, the response received no points under the strict scoring system.

The aligned scoring system described above was used in attempt to be as faithful as possible to the study in Kissling (2012), in which the answer sheet containing six slots per stimulus sequence allowed for participants to choose in which slot to write each syllable of their response sequences. However, when aligning spoken responses to the corresponding stimulus sequence in the present study, it was impossible to determine with certainty the alignment intended by the participant. For instance, in the example response in (1), aligning the final two syllables produced to the fifth and sixth slots results in a higher score, but it is possible that the speaker simply misremembered the third and fourth syllables and forgot the final two syllables entirely. In other words, the aligned scoring system may have awarded scores that overrepresent the recall accuracy of the speakers. A larger segmental inventory would have led to more possible syllables, and could have mitigated some of this uncertainty, but the phonemic inventories of the languages tested did not allow for this; Arabic contains only three phonemic vowels, and the subphonemic details of many other consonants across the three languages may have been enough to unnecessarily increase an L1 bias in the recall results, as discussed below.

In the end, the manipulation imposed by the aligned scoring system resulted in scores that were on the whole too uniformly high for significant effects to surface. This pattern is shown below in **Table 1**. The presence of a ceiling effect here is corroborated by a comparison to the mean scores in Kissling (2012). The mean accuracies in that study ranged from 52.7% to 64.85%. In **Table 1**, the lowest mean score, 3.778 (out of a possible 6 points), is equivalent to 63%; the highest, 4.762, equivalent to 79%. The difference between the aligned scores in the present study and those in Kissling (2012) further suggests that the aligned scoring method had the unintended consequence of obscuring any effects of language groups and stimulus type.

The strict scoring method resulted in scores that did not obscure the effects of language group and stimulus type, and that were more similar to those in Kissling (2012), as presented in **Table 2**. Given that the strict scoring system is used here, all mean scores are out of a possible mean of one point. However, a scoring system that involves no alignment obscures a different type of effect relevant to recall studies, namely the recency effect, in which the final syllables of the stimuli are advantaged over those in the medial positions of the sequence (Crowder, 1971; Frankish, 1996). The final scoring system, the partially

**Table 1:** Mean aligned score (SE) by L1 and sequence type, aligned scoring.

	Consonant-Variable	Vowel-Variable
English	4.332 (0.097)	4.762 (0.097)
Amharic	3.778 (0.167)	4.02 (0.169)
Arabic	4.292 (0.122)	4.442 (0.124)

**Table 2:** Mean sequence score (SE) by L1 and sequence type, strict scoring.

	Consonant-Variable	Vowel-Variable
English	0.344 (0.029)	0.482 (0.031)
Amharic	0.222 (0.04)	0.222 (0.042)
Arabic	0.351 (0.037)	0.331 (0.038)

**Table 3:** Mean sequence score (SE) by L1 and sequence type, partially aligned scoring.

	Consonant-Variable	Vowel-Variable
English	4.228 (0.040)	4.704 (0.041)
Amharic	3.648 (0.068)	3.869 (0.072)
Arabic	4.238 (0.050)	4.312 (0.052)

aligned system, accounts for the possible presence of a recency effect by aligning the final syllable produced to the final syllable of the stimulus, but avoids artificially inflating the scores by limiting alignment to only the final syllable. In this system, as with the fully aligned system, scores are out of a possible 6 total points, one per stimulus syllable. The examples in (5) and (6) demonstrate how the example responses in (1) and (3), respectively, are scored under the aligned scoring system.

(5) Partially aligned scoring for response in (1)

stimulus sequence	ka	ku	ki	ki	ka	ku	
raw score	ka	ku	ka	ku	–	–	2 points
aligned score	ka	ku	ka	–	–	ku	3 points

(6) Partially aligned scoring for response in (3)

stimulus sequence	ka	ku	ki	ki	ka	ku	
raw score	ka	ku	ki	<i>ku</i>	<i>ku</i>	<i>ka</i>	(ku) 3 points
aligned score	ka	ku	ki	<i>ku</i>	<i>ku</i>	ku	4 points

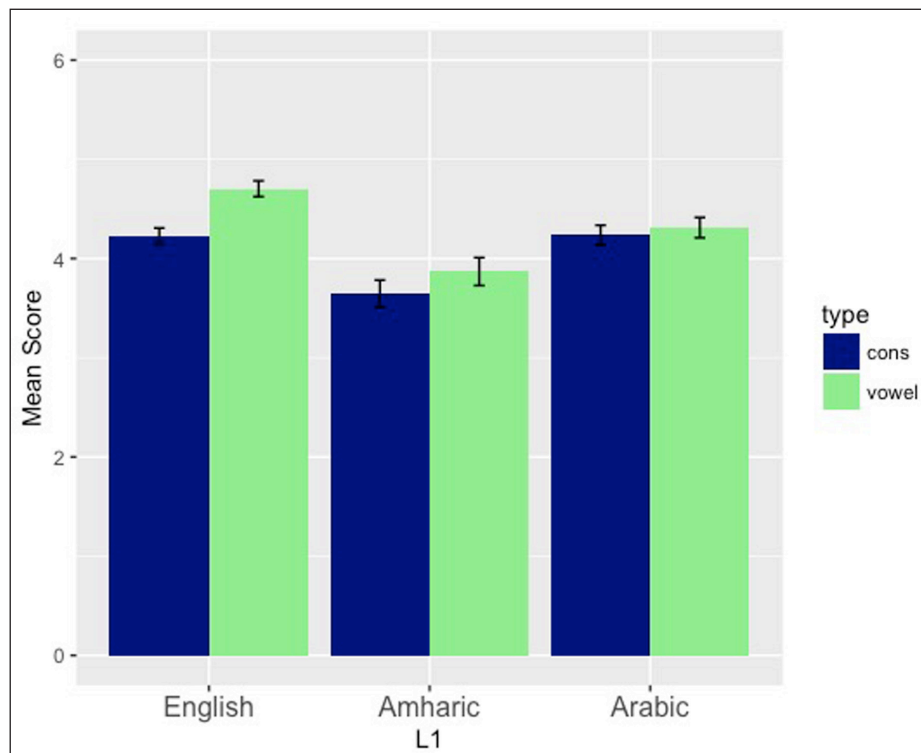
In (5), only the final syllable produced is moved rightward to align with the final syllable of the stimulus, bringing the score from 2 to 3 out of a possible 6 points. Similarly, in (6), only the final syllable produced is moved leftward to align with the final syllable of the stimulus, bringing the score of this production from 3 to 4. Given that the partially aligned scoring system captures a possible recency effect while avoiding ceiling effects, only results from this method are reported on and analyzed below.

#### 4. Results

Table 3 shows the mean scores for participants in all three L1 groups, for consonant-variable and vowel-variable sequences.

Figure 1 shows the mean stimulus scores, out of a possible 6 points, for each language group and sequence type, revealing similar scores across sequence types for Amharic and





**Figure 1:** Stimulus scores by language and sequence type.

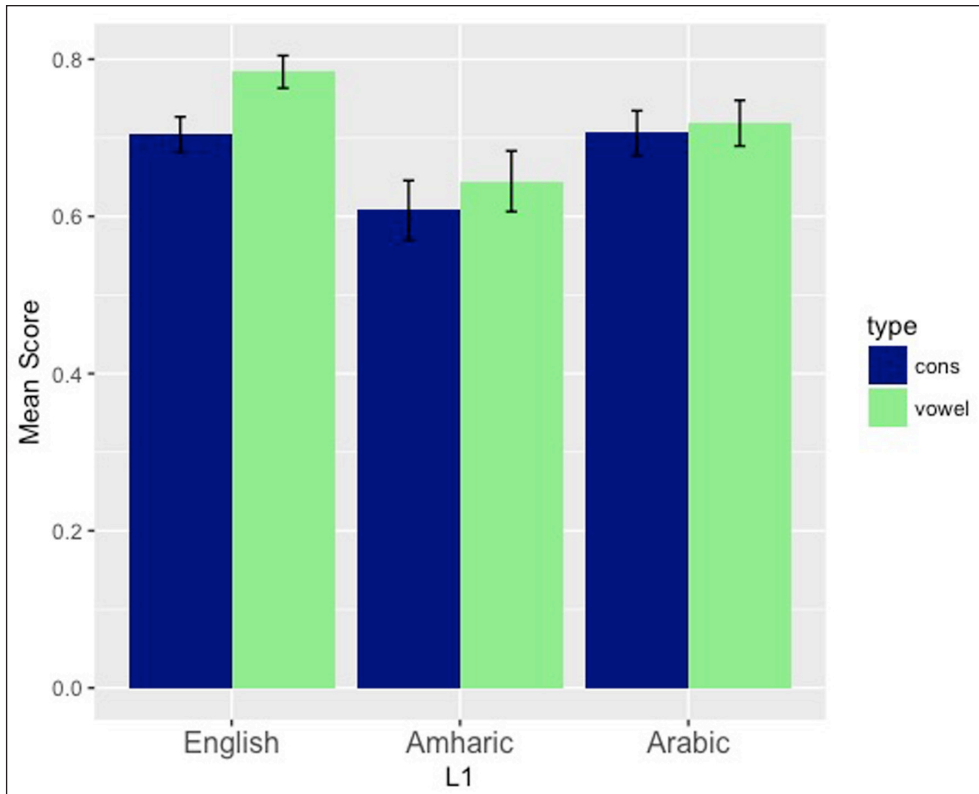
Arabic speakers. English speakers, on the other hand, scored higher remembering vowel-variable sequences than they did remembering consonant-variable sequences.

**Figure 2** shows the mean syllable score for each language group and sequence type, where each syllable received a score of 1 or 0. The same pattern emerges here: Whereas English speakers are better at remembering syllables in vowel-variable sequences than syllables in consonant-variable sequences, speakers of Amharic and Arabic show similar accuracies for syllables in both sequence types.

A mixed-effects logistic regression model was fit using the `glmer` function in the `lme4` R package (Bates, Mächler, Bolker, & Walker, 2015) to predict syllable accuracy (**Table 4**). The interaction between L1 and sequence type was a predictor, as was syllable position in the sequence, coded as initial, medial, or final. These three levels are motivated by the robust finding that ISR results tend to be ‘bowl-shaped,’ with initial and final portions of the stimuli being easier to recall than the medial portions (Crowder, 1971; Frankish, 1996).

The results in **Table 4** show that syllables in vowel-variable sequences remembered by English speakers were significantly different from those in consonant-variable sequences remembered by Arabic speakers. There is no other significant effect in the interaction between type and L1. In other words, whereas English speakers score significantly higher on syllables in vowel-variable sequences than on syllables in consonant-variable sequences, there is no such significant difference between stimulus type in the other two languages.

**Table 4** also reveals significant effects of syllable position. Initial syllables are significantly more likely to be remembered than those in medial position ( $p < 0.0001$ ; mean initial score = 0.9, mean medial score = 0.684). This reveals a clear primacy effect, as is apparent in **Figure 3**. There is also a significant difference between medial syllables and final syllables ( $p = 0.0001$ ); however, with a mean final score of 0.625, these final syllables have a significantly lower score than medial syllables, whereas recency effects would predict that final syllables would have higher recall scores than

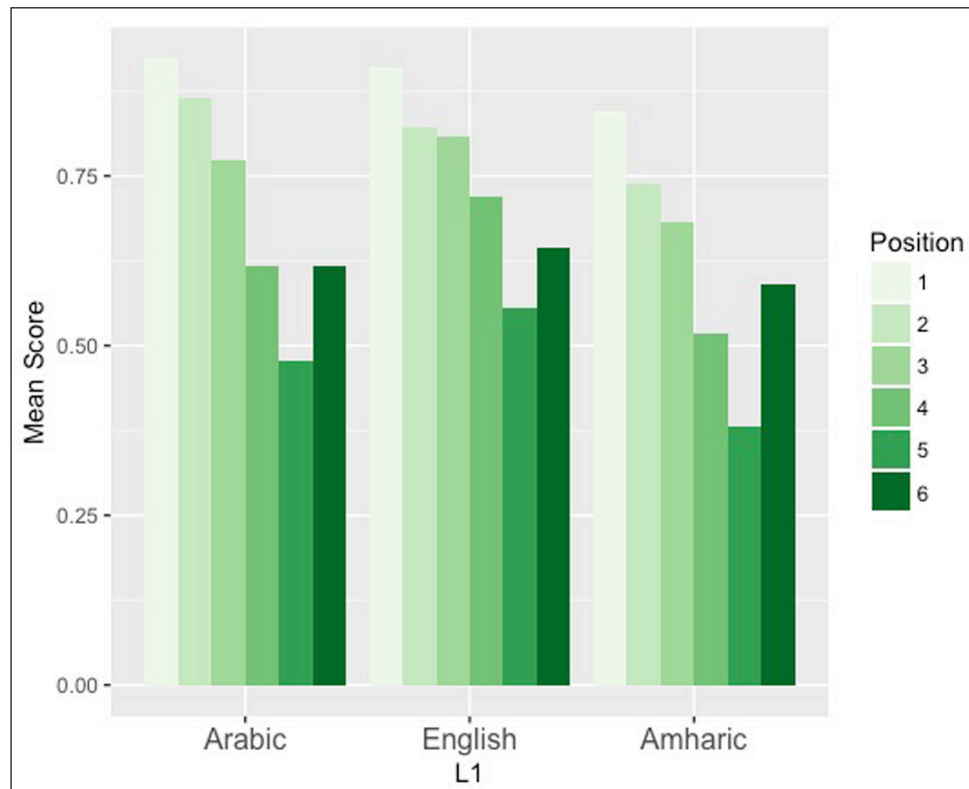


**Figure 2:** Syllable scores by language and sequence type.

**Table 4:** Mixed-effects logistic regression model: syllable accuracy. Consonant as reference level for stimulus type; Arabic as reference level for L1; medial as reference level for syllable position.

Fixed Effects	Estimate	SE	z-value	p-value
(Intercept)	0.8141	0.2393	3.402	<0.001 ***
<b>Type</b>				
Vowel	0.0672	0.2129	0.316	0.7523
<b>L1</b>				
Amharic	-0.4167	0.3224	-1.292	0.1962
English	0.0292	0.2582	0.113	0.9099
<b>Type * L1</b>				
Vowel:Amharic	0.1212	0.1672	0.727	0.4673
Vowel:English	0.4245	0.1389	3.056	0.0022 **
<b>Syllable Position</b>				
Initial	1.5822	0.1123	14.09	<0.001 ***
Final	-0.2966	0.07667	-3.869	0.0001 ***

medial syllables. Though **Figure 3** shows that the average score of the final syllable is higher than that of the penultimate syllable for all three L1 groups, and also higher than that of the antepenultimate syllable for Amharic speakers, they are not higher than the average score of all medial syllables. Therefore, though there is a strong primacy effect here, these results fail to show the presence of a recency effect as previously observed with ISR.



**Figure 3:** Syllable scores by language and position in sequence.

## 5. Discussion

### 5.1. Templaticity Effects

Results from this study show that English speakers remember vowel-variable sequences better than they do consonant-variable sequences, replicating findings in Drewnowski (1980), Crowder (1971), and Kissling (2012). Furthermore, the finding that Arabic speakers score the same on consonant-variable sequences as they do on vowel-variable sequences is also in line with the results in Kissling (2012). The novel piece of the present study, however, is the inclusion of Amharic speakers, used to tease potential templaticity effects apart from orthography effects. As seen here, Amharic speakers score equally well in this ISR task on consonant-variable sequences as they do on vowel-variable sequences, mirroring the behavior of the Arabic speakers. Given the similar morphological systems in Arabic and Amharic, and the significant difference in accuracy of English and Amharic speakers, the results suggest that it is morphological templaticity, not orthography, that has the greater effect on processing of consonants versus vowels. The templatic morphology of both Arabic and Amharic, in which the lexical root is comprised of three consonants, facilitates short-term memory of consonants for accuracy that is higher than that observed in speakers of non-templatic languages. This effect on consonants in templatic languages effectively counteracts any positive effect of vowel salience on ISR.

### 5.2. Task Effects

Though this study is in some ways a partial replication of that in Kissling (2012), there are important task differences that may reduce the comparability between the two. Whereas Kissling (2012) provided participants with 12 seconds to record their responses in writing, the current study asked participants to repeat the sequences they had heard aloud during an eight-second window. Despite several results in the present study that confirm those in Kissling (2012), it is possible that the 1500 ms lag time between the end of the stimulus

and the response time amounts to a processing burden that is not equivalent to that of reproducing the stimulus in writing. It may also be the case that there is a difference between using a phonological and orthographic buffer in completing this recall task. Therefore, it is possible that the present study does not replicate Kissling (2012) but rather shows that the same effects hold in a different type of speech sound processing. In order to determine the relationship between the results of this study and those of previous studies, it is necessary to investigate the processing demands on the participants in different types of tasks, and the extent to which they are comparable.

Another effect of the task differences between the present study and that in Kissling (2012) is that the latter experiment provided participants with six blanks in which to record their response to each stimulus sequence, as discussed above. Though neither experiment included explicit instructions about the number of syllables in each sequence, participants in Kissling (2012), after seeing the response sheet, presumably learned to expect six syllables in each stimulus. This methodology could have improved overall scores; it is possible that participants created memorization strategies with the knowledge of the stimulus length that in turn improved ISR overall. However, in the present study, participants had no such overt clue that each stimulus should be expected to have six syllables. Therefore, while some participants may have inferred this pattern over the course of the experiment and used it to create effective strategies, others may not have had the benefit of anticipating a particular stimulus length at all. In fact, the recorded responses impressionistically suggest that while some participants consistently produced responses with two prosodic units of three syllables each, there were some participants who each produced responses that varied from three to nine syllables in length, with intonation that revealed no such prosodic organization. Therefore, the ways in which the responses were recorded could have led to meaningful differences between the present study and that in Kissling (2012) with respect to the difficulty of the short-term memory task.

**5.3. Future Research**

**Consonants versus vowels in Amharic orthography** This study operates under the assumption that the Fidel abugida is relatively similar to English orthography, in that it explicitly encodes both consonantal and vocalic information, whereas most vowels in the Arabic orthography are only optionally written. However, each Fidel grapheme is comprised of a main shape that denotes the consonant and a relatively smaller component denoting the vowel. For example, as shown in the table in (7), the grapheme corresponding to the syllable /te/ is comprised of a large cross shape, encoding the /t/ onset, with a small circle in the lower right-hand quadrant encoding the /e/ nucleus. The relative sizes of the consonant and vowel components of this grapheme are representative of all of the <CV> graphemes in the Fidel syllabary.

(7) Examples of Fidel graphemes

	<Ci>	<Ca>	<Ce>
<tV>	፲	፳	፴
<sV>	ሰ	ሱ	ሲ
<bV>	ቦ	ባ	ቤ

Therefore, though both segments in a given CV grapheme are necessarily encoded in this system, it could be argued that the vowels are less visually prominent, making Fidel more similar to the Arabic script than assumed in this study. It is possible, then, that the

similarities between the Amharic and Arabic speakers in the performance on the ISR task in this study are additionally influenced by a similar orthographic system. A more fine-grained investigation into the exact visual and spatial properties of an orthographic system that impact the short-term memory employed in ISR would be necessary to substantiate this claim. The similar templatic morphological systems of both languages nonetheless suggest that morphophonology is at least partially, if not entirely, responsible for the results obtained in this study.

**Stimulus Language** A limitation of this study is that the stimuli were recorded by a native speaker of English; the English speakers were tested on their recall of syllables spoken by a speaker of their native language whereas Amharic and Arabic speakers listened to stimuli produced by a speaker of their L2. It has been shown that there is a first-language advantage in serial recall tasks, such that listeners more easily remember stimuli produced in their native language than in others, including languages in which they are relatively proficient (see discussion in Thorn, Gathercole, & Frankish, 2002). Based on this body of findings, it is not surprising that the overall means of the English speakers in this experiment were higher than those of the Amharic and Arabic speakers.<sup>3</sup> Forthcoming work<sup>4</sup> addresses this issue by conducting a similar recall experiment in which the stimulus language is the L2 of all participants. It is anticipated that this will eliminate any main effect of L1.<sup>5</sup>

However, this effect of L1 cannot explain the patterns found here. It is the difference between consonantal and vocalic recall for each L1 group that is most informative for the questions addressed in this study, not the difference in overall means across L1 groups. Though the language of the stimuli impacts recall, there is no reason to suspect that this effect impacts consonantal and vocalic recall to different degrees. Therefore, the results hold that the Arabic and Amharic speakers remembered the sequence types equally well.

**Participant Sampling** Due to the nature of participant sampling at an American university, there are potential confounds across the language groups in this study. For instance, the English speakers were more likely to be students at the university whereas the Amharic speakers were more often university staff members. Not only does this create different levels of familiarity and comfort with participation in an academic study across language groups, but it also raises the deeper issue of degree of literacy among the participants. Especially given that the university itself is a largely English-speaking environment, it is feasible that the English-speaking students are exposed to more written English than, for example, the Amharic speakers are to written Amharic. If this was the case, it weakens the impact of the comparison across orthography types, though there was no orthography effect observed in the results. Future iterations of this study should include speakers of all three languages with demonstrably comparable exposure to the orthography of their respective native languages.

Though orthography was not shown to have an effect on the ISR results in this study, it is possible that the orthographic system with which a speaker is familiar can impact segmental processing in some psycholinguistic tasks. Versions of this study should be conducted to examine the possible effects of different types of orthographic systems,

---

<sup>3</sup> Though the mixed-effects models do not show a main effect of L1, it is possible that this is due to the imbalance in speakers from each language group, and more data from Amharic and Arabic speakers would have revealed a statistically significant effect of L1 in this task.

<sup>4</sup> This work is ongoing as part of the author's forthcoming doctoral dissertation, entitled *The Relative Effects of Phonetic and Phonological Salience in Speech Sound Processing*.

<sup>5</sup> Another possible solution would be to have each participant hear stimuli in their own L1, such that three L1 groups would require three sets of stimuli. Though this would eliminate the confound of a first-language advantage, it would likely introduce other experimental confounds, as each participant would no longer be hearing identical sets of stimuli.

including purely syllabic systems (unlike Fidel, an abugida which is not quite a canonical syllabary) as well as logographic alphabets which do not encode explicit segmental information at all. Furthermore, given that processing effects such as priming and recall have been attested in the visual as well as auditory modalities (e.g., Drewnowski, 1980; New et al., 2008), conducting an ISR study similar to the present one but with visually-presented stimuli could reveal an impact of orthography on processing.

To bolster claims made here about the interaction between acoustic salience and language-specific morphophonological salience, future work will investigate other test cases in which phonetic and phonological prominence make opposite predictions. This work will add depth to this line of research, ultimately providing evidence for whether acoustic salience should in fact be considered a universal property of speech perception, or whether a given listener's L1 can impact speech perception such that acoustic salience does not have a measurable effect.

## 6. Conclusion

The present study set out to determine whether it is the morphology of a language or its orthography which impacts the way in which consonants and vowels are processed and stored in short-term memory. This question is motivated by the fact that whereas some experimental results show a categorical effect of consonants having a privileged role in phonological processing, other studies reveal the presence of language-specific influences on processing tasks such as ISR that appear to interact with the more general effects of consonants versus vowels on processing. The study compared Amharic speakers to Arabic speakers and English speakers; Amharic shares a morphological system with Arabic but its orthography is more similar to that of English, and therefore the comparison across all three groups allowed for the disentanglement of morphological effects from orthographic effects. Results show that Amharic speakers remember consonants as well as they do vowels, mirroring the results of Arabic speakers, and implying that the templatic morphology of the two languages is responsible for the observed effect. These results contribute to a broader understanding of how the morphophonological properties of a speaker's L1 impact their processing of different types of speech sounds.

## Acknowledgements

I am grateful to the participants in this study, as well as to audiences at the 9<sup>th</sup> World Congress of African Linguistics in Rabat, Morocco and the Linguistics Society of America annual meeting in New York City, and to Georgetown University's PhonLab for their feedback on earlier versions of this work.

## Competing Interests

The author has no competing interests to declare.

## References

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Caramazza, A., Chialant, D., Capasso, R., & Miceli, G. (2000). Separable processing of consonants and vowels. *Nature*, 403(6768), 428–430. DOI: <https://doi.org/10.1038/35000206>
- Carreiras, M., Gillon-Dowens, M., Vergara, M., & Perea, M. (2009). Are vowels and consonants processed differently? Event-related potential evidence with a delayed letter paradigm. *Journal of Cognitive Neuroscience*, 21(2), 275–288. DOI: <https://doi.org/10.1162/jocn.2008.21023>


- Crowder, R. G. (1971). The sound of vowels and consonants in immediate memory. *Journal of Verbal Learning and Verbal Behavior*, 10(6), 587–596. DOI: [https://doi.org/10.1016/S0022-5371\(71\)80063-4](https://doi.org/10.1016/S0022-5371(71)80063-4)
- Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & Van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory & cognition*, 28(5), 746–755. DOI: <https://doi.org/10.3758/BF03198409>
- Delle Luche, C., Poltrock, S., Goslin, J., New, B., Floccia, C., & Nazzi, T. (2014). Differential processing of consonants and vowels in the auditory modality: A cross-linguistic study. *Journal of Memory and Language*, 72, 1–15. DOI: <https://doi.org/10.1016/j.jml.2013.12.001>
- Detey, S., & Nespoulous, J.-L. (2008). Can orthography influence second language syllabic segmentation?: Japanese epenthetic vowels and French consonantal clusters. *Lingua*, 118(1), 66–81. DOI: <https://doi.org/10.1016/j.lingua.2007.04.003>
- Drewnowski, A. (1980). Memory functions for vowels and consonants: A reinterpretation of acoustic similarity effects. *Journal of Verbal Learning and Verbal Behavior*, 19(2), 176–193. DOI: [https://doi.org/10.1016/S0022-5371\(80\)90162-0](https://doi.org/10.1016/S0022-5371(80)90162-0)
- Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25(6), 1568. DOI: <https://doi.org/10.1037//0096-1523.25.6.1568>
- El Aissati, A., McQueen, J. M., & Cutler, A. (2012). Finding words in a language that allows words without vowels. *Cognition*, 124(1), 79–84. DOI: <https://doi.org/10.1016/j.cognition.2012.03.006>
- Frankish, C. (1996). Auditory short-term memory and the perception of speech. *Models of short-term memory*, 179–207.
- Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, 13(1), 104. DOI: <https://doi.org/10.1037//0096-1523.13.1.104>
- Hamann, S., & Colombo, I. E. (2017). A formal account of the interaction of orthography and perception. *Natural Language & Linguistic Theory*, 35(3), 683–714. DOI: <https://doi.org/10.1007/s11049-017-9362-3>
- Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In *Advances in psychology* (Vol. 94, pp. 67–84). Elsevier. DOI: [https://doi.org/10.1016/S0166-4115\(08\)62789-2](https://doi.org/10.1016/S0166-4115(08)62789-2)
- Kissling, E. M. (2012). Cross-linguistic differences in the immediate serial recall of consonants versus vowels. *Applied Psycholinguistics*, 33(3), 605–621. DOI: <https://doi.org/10.1017/S014271641100049X>
- Leslau, W. (1995). *Reference grammar of Amharic*. Otto Harrassowitz Verlag.
- Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition*, 24(1), 45–64. DOI: [https://doi.org/10.1016/0010-0277\(86\)90004-1](https://doi.org/10.1016/0010-0277(86)90004-1)
- Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, 7(4), 323–331. DOI: [https://doi.org/10.1016/0010-0277\(79\)90020-9](https://doi.org/10.1016/0010-0277(79)90020-9)
- Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e linguaggio*, 2(2), 203–230.
- New, B., Araújo, V., & Nazzi, T. (2008). Differential processing of consonants and vowels in lexical access through reading. *Psychological Science*, 19(12), 1223–1227. DOI: <https://doi.org/10.1111/j.1467-9280.2008.02228.x>

- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of nonadjacent dependencies. *Cognitive psychology*, *48*(2), 127–162. DOI: [https://doi.org/10.1016/S0010-0285\(03\)00128-2](https://doi.org/10.1016/S0010-0285(03)00128-2)
- Peirce, J. W. (2007). Psychopy – psychophysics software in python. *Journal of neuroscience methods*, *162*(1–2), 8–13. DOI: <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Ryding, K. C. (2005). *A reference grammar of Modern Standard Arabic*. Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511486975>
- Thorn, A. S., Gathercole, S. E., & Frankish, C. R. (2002). Language familiarity effects in short-term memory: The role of output delay and long-term knowledge. *The Quarterly Journal of Experimental Psychology: Section A*, *55*(4), 1363–1383. DOI: <https://doi.org/10.1080/02724980244000198>
- Toro, J. M., Nespors, M., Mehler, J., & Bonatti, L. L. (2008). Finding words and rules in a speech stream: Functional differences between vowels and consonants. *Psychological Science*, *19*(2), 137–144. DOI: <https://doi.org/10.1111/j.1467-9280.2008.02059.x>
- Turnbull, R., & Peperkamp, S. (2018). The asymmetric contribution of consonants and vowels to phonological similarity. *The Mental Lexicon*, *12*(3), 404–430. DOI: <https://doi.org/10.1075/ml.17010.tur>

**How to cite this article:** Barzilai, M. L. 2019 Templaticity Effects on Differential Processing of Consonants and Vowels. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 10(1):17, pp.1–16. DOI: <https://doi.org/10.5334/labphon.180>

**Submitted:** 14 November 2018    **Accepted:** 23 September 2019    **Published:** 17 October 2019

**Copyright:** © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Laboratory Phonology: Journal of the Association for Laboratory Phonology* is a peer-reviewed open access journal published by Ubiquity Press.

**OPEN ACCESS** 