Yahya Aldholmi* and Anne Pycha

Segmental contributions to word recognition in Arabic sentences

https://doi.org/10.1515/psicl-2022-2010

Abstract: We examined the contributions of segment type (consonants vs. vowels) and segment ratio to word recognition in Arabic sentences, a language that has a nonconcatenative morphological system in which consonants indicate semantic information, while vowels indicate structural information. In two experiments (with a balanced vowel-to-consonant ratio in Experiment 1 and an imbalanced ratio in Experiment 2), we presented participants with spoken sentences in Modern Standard Arabic, in which either consonants or vowels had been replaced by silence, and asked them to report what they could understand. The results indicate that consonants play a much greater role than vowels, both for balanced and also imbalanced sentences. The results also show greater word recognition for stimuli that contained a higher ratio of consonants to vowels. These results support and supplement previous findings on the role of consonantal roots in word recognition in Semitic languages, but clearly differ from those previously reported for non-Semitic languages which highlight the role of vowels in word recognition at the sentence level. We interpret this within the framework of root-and-pattern morphology, and further argue that segmental effects on word recognition and speech processing are crucially modulated by morphological structure.

Keywords: nonconcatenative languages; root-pattern morphology; Semitic languages; speech processing; vowels versus consonants

1 Introduction

The contributions of vowels versus consonants to word recognition have been explored in a small set of languages, including English (Cole et al. 1996), Spanish, Dutch (Cutler et al. 2000), and Mandarin Chinese (Chen et al. 2013, 2015). Such studies

*Corresponding author: Yahya Aldholmi, Department of Linguistics, College of Language Sciences, King Saud University, P. O. Box 87907, Riyadh 11652, Saudi Arabia, E-mail: yaldholmi@ksu.edu.sa. https://orcid.org/0000-0003-2373-1335
Anne Pycha, Department of Linguistics, College of Letters & Science, University of Wisconsin-Milwaukee, Milwaukee, WI 53211, USA
sparked a debate about which segment type contributes more to word recognition for isolated words and in complete sentences, and presented asymmetric findings supporting vowel privilege in some cases and consonant privilege in others. Many of the discrepancies in previous findings may be attributed to stimulus design choices. Studies such as Cole et al. (1996) used complete English sentences to examine segmental contributions to word recognition in fluent speech (i.e., sentences), and found that vowels play a larger role than consonants do. Others, such as Owren and Cardillo (2006), used isolated English words and found the opposite result. More recent studies (e.g., Fogerty et al. 2012) have compared isolated words and complete sentences. In general, these studies have shown that consonants contribute more than vowels to the recognition of words in isolation (although note that Chinese is an exception), whereas vowels contribute more than consonants to word recognition in sentences.

One key factor that differentiates word recognition in isolation from word recognition in sentences is the suprasegmental and contextual cues (also referred to as contextual benefits or sentential information) available from the rest of the sentence (e.g., Cutler and Foss 1977; Grant and Seitz 2000; Shields et al. 1974; Tajima and Port 2003), which help listeners to recognize words even in challenging listening conditions (e.g., Van Engen et al. 2014). Acoustic factors such as fundamental frequency have not been found to be strong enough to alter the contributions of vowels at the sentence level (Fogerty and Humes 2012). Fogerty et al. (2012) examined segmental contributions to the recognition of words in isolation and in sentences and compared segmental contributions in both types of stimuli (i.e., words vs. sentences). They devoted a proportion of their study to analyze the contextual benefits that were provided to vowels versus consonants in sentences as compared to words in isolation. Their analysis revealed that consonants did not gain contextual benefits at the sentence level while vowels by comparison did. Hence, they concluded that “spectral differences alone do not explain the observed greater vowel contributions in sentences [and that] vowels provide contextual cues for the higher-order processing of … sentences” (p. 1676).

Fogerty et al.’s (2012) conclusion could be interpreted as an indication that segmental contributions are not only tied to the acoustic properties of vowels versus consonants; rather, one segment type may contribute richer information than another depending on its context within a sentence. This allows listeners to use such non-acoustic information for top-down processing, which, along with bottom-up processing, enhances word recognition. This has also been found in Mandarin Chinese. Chen et al. (2013) examined segmental contributions to word recognition in Mandarin sentences and found that in addition to the advantage that vowels gain in English sentences, Mandarin, as a tonal language, provides listeners with
additional tonal information, effectively lowering the number of lexical candidates to yield better word recognition in sentences. This means that although vocalic and consonantal information is universally similar, different languages assign more information to one class of segments, as has been found for vowels in Mandarin.

It follows that a language that employs one of the two classes of segment types in a particular way – for example, for morphological reasons, as in Arabic – is likely to behave differently. This forms the foundation of possible opposition between concatenative languages such as English and nonconcatenative languages such as Arabic. Nonconcatenative languages assign a substantial role to consonants in their lexical architectures – namely, consonants carry lexical and semantic information while vowels convey structural information (e.g., Bentin and Feldman 1990; Deutsch and Frost 2003; Feldman 2000; Holes 2004; Ravid 2003; Shimron 2003). The languages in which segmental contributions to word recognition both at the word level and at the sentence level have been investigated are mainly concatenative systems in which lexical items are built by combining discrete stems with discrete affixes. Such languages do not assign fundamentally different roles to vowels and consonants in the morphology. For example, in an English word such as “edit” /edit/, the two consonants /d, t/ are morphologically equivalent to the two vowels /ɛ, ɪ/ in terms of how the word is built. Languages such as Arabic, on the other hand, use a nonconcatenative morphological system in which a discontinuous sequence of consonants (i.e., a root) signals lexical information, whereas vowels (i.e., a pattern) are intercalated to signal morphosyntactic information, similar to the function of affixes in concatenative languages (see Holes 2004; McCarthy 1981; Watson 2007, among many others). For example, the Modern Standard Arabic (MSA) consonantal root Ꙡ-ʃ-ʕ “related to LOVE” can be used to generate different lexical items by interdigitating this consonantal root with a string of vowels, such as /ʕaʃiʕ/ ‘loved’, /ʕuʃiʕ/ ‘was loved’, /ʕaʔaʃiʕ/ ‘lover’, and /ʕaʃiʕ/ ‘beloved’. Other Semitic languages behave similarly.

Speakers of Arabic use consonantal roots to recognize, access, and retrieve words (e.g., Abu-Rabia 1997; Bentin and Feldman 1990; Deutsch and Frost 2003; Deutsch et al. 2000; Feldman 2000; Holes 2004; Ravid 2003; Shimron 2003). These studies have concluded that the consonantal root is a fundamental unit in word recognition for other Semitic languages, but such studies almost exclusively used individual real or pseudo words (but not complete sentences) for their stimuli. The difference in the stimuli (i.e., the use of words vs. sentences) presented to the participants has been shown to be a factor in segmental contributions to the recognition of words in isolation and in sentences in non-Semitic languages such as English and Dutch. Within the root-and-pattern framework where lexical information is carried
by consonants, as in Arabic, consonants should play a greater role in word recognition in sentences than vowels do. In other words, the additional advantage that vowels gain in sentences is not expected to render vowels more important for word recognition in sentences; rather, the critical role that consonants play in Arabic should make consonants contribute more than vowels, even in sentences. To the best of our knowledge, no study has yet assessed the relative contributions of consonants versus vowels to word recognition in Arabic sentences as an exemplar of Semitic/nonconcatenative languages. Thus, we investigate the following:

(1) What are the segmental contributions to word recognition in Arabic sentences?

In addition, since the morphology of Arabic assigns lexical information mainly to consonants, it is especially important in Arabic to carefully control for segmental ratio in the construction of the stimuli and to examine its role in the contribution of vowels versus consonants. Hence, we also consider a second question:

(2) How does segmental ratio affect the intelligibility of Arabic sentences?

In order to answer these two questions, we conducted two experiments using MSA stimuli in which either the consonants or the vowels had been replaced by silence. The number of vowels and consonants were balanced in the first experiment and imbalanced in the second experiment. Our results show that consonants contribute more to the recognition of words in Arabic sentences than do vowels in both experiments. The results also show that the participants were relatively better able to recognize words in sentences that contained a higher ratio of consonants to vowels, which could be due to the fact that Arabic speakers heavily rely on consonants not only to recognize words in isolation (as previous studies have confirmed), but also in complete sentences, as the current study suggests.

2 Background

Segmental contributions to word recognition both for isolated words and in complete sentences have been conducted in non-Semitic languages such as English and Spanish. In Semitic languages, however, research has focused on a related but slightly different topic – namely, the role of roots versus patterns (which is roughly equal to consonants vs. vowels) in word recognition. These studies consistently suggest the importance of the consonantal root in mental representation, lexical access, and word recognition. Below, we review the major findings from these two lines of research prior to delving into the details of the current study.
2.1 Contributions of consonants versus vowels in non-Semitic languages

In their influential study, Cole et al. (1996) examined segmental contributions to English word recognition in fluent speech using sentences from the Texas Instruments/Massachusetts Institute of Technology (TIMIT) Corpus (Garofolo et al. 1993). The researchers classified English segments into three types: consonants, vowels, and weak sonorants.

Their results showed that speech intelligibility was higher when the participants were presented with vowel-only (VO) stimuli than when they were presented with consonant-only (CO) stimuli. This finding held regardless of whether white noise versus periodic tones were used to replace missing segments and whether vowel-consonant boundaries were shrunk or expanded, ruling out the possibility that vowels benefited from the coarticulatory information more than consonants did. On this basis, Cole and colleagues concluded that vowels’ contribution to word recognition in sentences is greater than that of consonants. Other similar studies that adopted the same type of stimuli (i.e., sentences) and examined the same topic (e.g., Burkle et al. 2004; Kewley-Port et al. 2007) are in agreement with Cole et al. For instance, in Kewley-Port et al. (2007), the performance of normal hearing listeners on word recognition at the sentence level was compared to that of hearing-impaired elderly listeners. The noise replacement paradigm used in Cole et al. (1996) was adopted with a small modification, specifically that the weak sonorants were classified as consonants rather than a separate category. The first experiment replicated the results of Cole et al. (1996) and showed an increase in performance in the VO condition compared to the CO condition. In the second experiment, although young normal hearing listeners performed better than the hearing-impaired listeners, both groups statistically performed better when only vowels were present in the stimuli, compared to when only consonants were present.

However, not all studies support Cole et al.’s findings regarding the greater contribution of vowels in English word recognition in general. Several studies have argued that consonants contribute to word recognition more than vowels do, although the experimental tasks used by these authors often differ from the noise replacement technique. van Ooijen (1996) showed that consonants are more useful than vowels in word recognition for English listeners. Her study used the “reconstruction task” in which participants were given auditory nonsense inputs and were instructed to rapidly substitute one segment to convert it into a meaningful English word. In the free-choice condition, either the consonant or the vowel could be changed, so that, for example, *kebra* can become *zebra* or *cobra*. In two other conditions, the subjects could only modify a consonant (in one condition) or a vowel (in
the other condition). The results showed more errors in vowel change compared to consonant change. The author concluded that listeners have a perceptual mechanism to deal with uncertainty and variability in English vowels but not consonants. She claimed that her findings would have implications for approaches to speech recognition/perception in both normal and noisy environments, because if the identity of a vowel is inherently more flexible (or perhaps under-defined) than that of consonants, consonants are more likely to be affected by distortion than vowels, which would affect word recognition and speech intelligibility. Four years later, in joint work with Cutler and other colleagues (Cutler et al. 2000), they addressed the issue of consonant-to-vowel ratio in a cross-linguistic comparison study using the above-described reconstruction task with Dutch and Castilian Spanish. The opposition between Dutch, which has 16 vowels and 19 consonants, and Spanish, which has 5 vowels and 20 consonants, produced a useful comparison between the two languages in terms of the ratio. In their first experiment, a group of native Dutch speakers were presented with disyllabic nonwords. Each nonword had the potential to become a real word if a vowel or consonant was altered; furthermore, in each word, an approximately similar number of vowels and consonants were potential targets for alteration. The findings were consistent with those of van Ooijen (1996); the proportion of correct responses involving vowels was higher than that of consonants, the reaction time was faster in the vowel condition, and the error rate was higher in the consonant condition. In their second experiment, a parallel design was used with Spanish, and the outcome was almost identical to the one obtained with Dutch. Thus, Cutler et al. (2000), along with confirming the previous results found in English by van Ooijen (1996), ruled out the possibility that consonant’s privilege could be due to the consonant-to-vowel ratio in the language inventory because the results from two languages with very different ratios, namely Spanish and Dutch, were parallel to one another. Owren and Cardillo (2006) investigated the contribution of vowels versus consonants to word recognition by adapting the methodologies of Cole et al. (1996) with two main modifications: the use of silence replacement in lieu of noise and the use of isolated words instead of sentences. In a same-different task with three conditions – intact speech, VO, and CO – the participants judged whether the stimuli they heard had the same or different meanings. The findings assert that consonants are more useful for listeners to recognize word meaning. This was inconsistent with the results found by Cole et al. (1996), although Owren and Cardillo (2006) used words in isolation while Cole et al. (1996) used words embedded in sentences.

One remaining issue is the fact that Dutch, English, and Spanish belong to the same Indo-European language family and share a similar concatenative morphological system that does not treat vowels and consonants differently. Other languages that demonstrate a special treatment of vowels versus consonants, such as Arabic...
and Hebrew, may display a divergent pattern. Using a different experimental technique, Bonatti et al. (2005) investigated the informational role of consonants and vowels using an artificially manipulated language. In a familiarization phase, French participants were informed that they would hear an artificial language and then were presented with a stream of speech with different fixed distributions of vowels and consonants. In the later test phase, they were presented with pairs of items and asked to indicate which series of segments sounded like a word in the artificial language. As specified in their experimental design, a word is a set of three consonants such as \(b_d_k\) concatenated with different patterns of vowels such as \(biduka\), \(bidoke\), \(byduka\), and \(bydoke\), whereas a nonword (called “part-word” in their study) is a series of consonants and vowels belonging to different words but combined together to form a word-like pattern. This design resembles lexical structure in Semitic languages where vocalic information determines grammatical properties and consonants individuate lexical items. In the second experiment, they reversed the role of vowels and consonants; vowels served as the root-like component. The findings from their experiments showed that the listeners were successful in using consonants, but not vowels, to extract words, which was interpreted as evidence that the consonant tier can play a key role in identifying words when processing speech. The preferential role for consonants in speech was not due to a numerical asymmetry in language consonant-to-vowel ratio leading to favoring consonants over vowels, because the French inventory has a relative balance between consonants (17) and vowels (16). This conclusion was also supported in a subsequent study (Toro et al. 2008). Taken together, these studies provide evidence that segmental contributions differ in words from sentences and in one language from another.

Fogerty et al. (2012) compared segmental contributions in isolated words and sentences and concluded that vowels contribute to word recognition and intelligibility more than consonants do, but only in sentences. Fogerty (2019) also investigated vocalic contributions to the intelligibility of sentences across eight American English dialects and found sentences preserving vocalic information to be more intelligible across all dialects. In addition to such studies on English and other Indo-European languages, two relatively recent studies investigated segmental contributions in Mandarin Chinese (Chen et al. 2013, 2015). Mandarin has tones that function phonemically, just as consonants and vowels do, but it does not show a morphologically special treatment of vowels in contrast to consonants. Chen et al. (2013) used sentences from the Mandarin speech perception (MSP) corpus (Fu et al. 2011) and adopted the noise replacement technique from Cole et al. (1996), while Chen et al. (2015) used isolated words from a large database. The findings from both studies confirm the advantage of vowels over consonants in word recognition both at the word level and the sentence level. The authors attribute this partial difference
between English and Mandarin to differences in the linguistic architecture of each language with respect to syllables and consonant clusters, a conclusion that supports the assumption that segmental contributions may not be universal.

Using a variety of experimental techniques and tools, results from these aforementioned studies suggest that, in languages such as English and Dutch, vowels are more important for word recognition in sentences, while consonants contribute more for word recognition in isolation. On the other hand, vowels contribute more than consonants for word recognition both in isolated words and in sentence context in other languages such as Chinese. However, all of these previous studies examined primarily concatenative languages, which have a similar morphological treatment of vowels and consonants. The difference in contribution of vowels versus consonants may turn out to be language-specific, depending on the roles assigned to these segments by the morphological system. Previous studies on word recognition in Semitic languages have shown that listeners decompose words into individual morphemes corresponding to roots (i.e., consonants) and patterns (i.e., vowels). Consonants in such languages may play a role similar to that played by consonants in the artificial language processed by French speakers discussed above. The next subsection will review the respective roles of vowels and consonants in word recognition specifically in Semitic languages.

2.2 Roles of consonants and vowels in Semitic languages

Semitic languages such as Arabic and Hebrew use a nonconcatenative morphological system (also referred to as non-linear, discontinuous, or root-and-pattern morphology) in which a sequence of consonants signals lexical (semantic) information, whereas vowels are intercalated to signal structural (morphosyntactic) information, similar to the function of affixes in concatenative languages (e.g., McCarthy 1981). An example paradigm is provided in Table 1, in which the MSA consonantal root ʕ-ʃ-q ‘related to LOVE’ is used to generate different lexical items by

<table>
<thead>
<tr>
<th>Lexical item</th>
<th>Meaning</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ʕaʃiq/</td>
<td>‘loved’</td>
<td>Perfective/past</td>
</tr>
<tr>
<td>/ʕaaʃiq/</td>
<td>‘lover’</td>
<td>Agent/doer/active participle</td>
</tr>
<tr>
<td>/ʕaʃiŋ/</td>
<td>‘beloved’</td>
<td>Adjective</td>
</tr>
<tr>
<td>/ʕaʃuqq/</td>
<td>‘who is in deep love’</td>
<td>Adjective</td>
</tr>
<tr>
<td>/ʕuʃʃaaq/</td>
<td>‘who are in deep love’</td>
<td>Plural of active participle</td>
</tr>
<tr>
<td>/ʕaʃaq/</td>
<td>‘love/loving’</td>
<td>Base form/gerund/nominal</td>
</tr>
</tbody>
</table>
interdigitating this consonantal root with a string of vowels (often referred to as a pattern). The consonantal skeleton typically consists of three consonants (although a sequence of two or four consonants is also possible) and forms a discontinuous and unpronounceable morpheme on one tier to supply the core semantic information around which all related forms are clustered. A separate pattern composed mainly of vowels acts as another morpheme on a separate tier to supply morphosyntactic information such as ontological categories for nouns and aspectual properties for verbs (see Goral and Obler 2003; Holes 2004; McCarthy 1981; Watson 2007, among many others). The patterns, which usually consist of vowels, are not wholly represented in Arabic orthography except for some teaching and learning purposes; this is also the case in Hebrew (Feldman et al. 1995).

The meaning of a Semitic word is collectively constructed via the root and pattern combination, sometimes in an idiosyncratic fashion (Bentin and Feldman 1990; Feldman et al. 1995). As the partial derivational paradigm above illustrates, the semantic information associated with LOVE is shared across the different forms, but each has a different structure reflected by a change in the quantity and/or quality of vowels. The number of consonants and vowels in the various forms is generally fixed and predictable. For example, a basic (not causative or reciprocal) active perfective form will have three consonants and two vowels, as in /ʕaʃiq/ ‘loved’, and the consonant-to-vowel ratio will remain 3:2 for all other roots conjugated in the active perfective, as in /karih/ ‘hated’, /fariḥ/ ‘got happy’, /ḥazin/ ‘got sad’, /saxitˤ/ ‘got furious’, /nadim/ ‘regretted’, and so on. A passive form of these examples will change only the quality but not the quantity of the first vowel, as in /ʕuʃiq/ ‘was loved’ and /kuriḥ/ ‘was hated’, whereas the agentive form will feature a change in the quantity but not the quality of the first vowel (3C-3V ratio) as in /ʕaaʃiq/ ‘lover’ and /kaarih/ ‘hater’.

Several phonological, morphological, and psycholinguistic studies have examined whether the root and/or pattern is used as a unit to represent, access, and recognize lexical items in Semitic languages, especially Hebrew. The findings from such studies consistently report that Semitic language speakers heavily rely on the root to represent, access, and recognize items from their mental lexicon. Bentin and Feldman (1990) sought evidence from Hebrew for the contribution of morphological (and semantic) relatedness effects in repetition priming, known as the “morphological repetition effect”. Using nonwords and Hebrew words that are morphologically related (i.e., sharing a root), semantically related (i.e., sharing meaning but from different roots), or morphologically and semantically related (i.e., sharing both a root and similar meaning), the experimenters tested for facilitation in a lexical decision task in which the stimuli were presented visually. The reaction time as well as the error rate confirmed the results found in other languages; morphological facilitation was maintained over long lags while semantic facilitation was not. However, the root
in Semitic languages differs from the stem in other languages such as English, since the Semitic root is an abstract discontinuous unit, such as /sfr/ ‘related to reporting’, while the English stem is usually a word that can stand on its own, such as ‘pay’. The authors concluded that the root and pattern are unlikely to have separate lexical units in the lexicon, and suggested a mechanism of ‘extraction’ rather than decomposition.

However, in a later study, Feldman and colleagues (Feldman et al. 1995) compared decomposition in English and Hebrew. The necessity for this study came from the realization that, although both English and Hebrew can be categorized as inflected languages, they differ in the manner by which they generate different lexical items from the stem (in English) and root (in Hebrew). In their study, the researchers used the so-called “segment-shifting task” in which Hebrew speakers were presented with written Hebrew source words and asked to detach the pattern and apply it to a pseudoroot (a sequence of three letters that are not used in Hebrew for any meaning). The findings support the hypothesis that Hebrew lexical organization is based on nonconcatenative morphemic units, i.e., roots and patterns, and that morphological decomposition as found in English is also available in Hebrew, irrespective of the difference in the stem or root-affix appending process. Feldman (2000) investigated the morphological effects relative to semantic and orthographic effects in English. While the study did not include any Semitic languages, it complements the study by Bentin and Feldman (1990) and provides a preview of another study on Hebrew, which will be introduced shortly. Deutsch et al. (2000) contrasted morphological relatedness with orthographic relatedness to address the factors that affect word identification in Hebrew. They use a technique known as “parafoveal preview benefit” to test whether it would lead to the same or different findings that previous techniques have led to. The authors used three types of stimuli: identical, morphologically related, and orthographically related Hebrew prime and target words. The participants were instructed to name the target as rapidly as they could, and the reaction time and error rate were calculated to examine preview benefit effects. It was found that morphological facilitation was greater than the facilitation in the orthographic condition.

The above studies on Hebrew, mainly using masked and cross-modal priming, support the root-and-pattern framework that distinguishes between the roles of consonants and vowels in word formation (as well as word representation and recognition) in Semitic languages. As for Arabic, Boudelaa and colleagues (e.g., Boudelaa and Marslen-Wilson 2000, 2001, 2004a, 2004b, 2005, 2011, 2013; Boudelaa et al. 2010) have conducted a sizable number of studies on the status of Arabic roots/consonants and patterns/vowels mainly in spoken and occasionally in orthographic word recognition. These studies individually and collectively support a unified conclusion that Arabic morphology is best accounted for by the root-and-pattern based
approach, which has been challenged every now and then by stem-based approaches, at least for some Arabic forms or dialects (e.g., Benmamoun 1999, 2003; Ratcliffe 1998). One of the most recent studies that sought to compare the applicability of the stem-based (concatenative) and root-based (nonconcatenative) approaches in Arabic is Boudelaa and Marslen-Wilson (2015) who examined the processing of different Arabic nouns and verbs to determine the status of roots and patterns in mental representation. Their findings from five experiments illustrate that root and pattern morphemes are two cognitive, abstract units, which clearly supports the historically accepted root-and-pattern framework in the literature. Other studies (mainly by Abu-Rabia and colleagues) using reading tasks and orthographic stimuli have also provided some support to the role of consonants in written Arabic word recognition for both native and nonnative speakers (e.g., Abu-Rabia 2002, 2012), as well as findings that challenge the theory (Abu-Rabia and Jasmin 2004). However, discussion on written word recognition will expand this background beyond the scope of the current study. In the next section, we will delve into the details of our first experiment.

3 Experiment 1

In this experiment, we investigated segmental contributions at the sentence level in Arabic, using the replacement paradigm that has been used in many previous studies that tackled the same topic (e.g., Chen et al. 2015; Cole et al. 1996; Fogerty et al. 2012; Kewley-Port et al. 2007) and balancing the ratio of consonant-to-vowel in each sentence. Given the importance of consonants in Arabic word recognition within the root-and-pattern framework for Semitic languages, and given the results of previous studies on segmental contribution that have emphasized the role of vowels in word recognition at the sentence level in non-Semitic languages, the segmental contribution at the sentence level in Arabic would be worthy of pursuit.

3.1 Methodology

3.1.1 Sentence construction

We constructed 48 sentences each of which contained exactly six words representing different parts of speech. This decision was based on Kewley-Port et al. (2007), whose sentences ranged from 6 to 10 words with an average of 8.01 words per sentence. Since Arabic enjoys a rich morphology that allows case, mood, agreement, (in)definite pronouns, and clitics to all be part of what can be counted as one word, a 6-word sentence would likely be close to the average number of English words used in Kewley-Port et al. (2007). In addition, each sentence had a 1:1 segmental ratio, with
23 consonants and 23 vowels. Across the sentences, all consonants and vowels in Arabic's phonemic inventory were represented. The sentences covered a wide range of topics familiar to an Arabic speaker such as politics, sports, and life events. The example below demonstrates one sentence from the stimuli.

(1) ʃaarak-a jahjaan hamiiidan
participate.PRF-3PSM Yahya-NOM Hamid-ACC
fi muʔtamar-i l-lisaa[n]ijjaat
in conference-GEN the-linguistics
‘Yahya participated with Hamid in the linguistics conference.’

3.1.2 Judgements

We submitted all sentences to 10 native speakers of Arabic, who judged, on a scale from 1 to 7, every sentence according to how natural it sounded to them, and again based on how likely it would be heard in their daily lives. The sentences were sent to five judges in written form once, and later to five different judges to rate the sentences in spoken form. Any sentence that was overall rated less than 6 was subject to be modified and re-judged.

3.1.3 Recording

The sentences were recorded by a male native speaker of Arabic with a linguistics background. Although previous studies used stimuli produced by a variety of different speakers, we determined that having only one speaker would rule out any possibility of inter-speaker variation effects. The speaker was asked to record the sentences on Praat (Boersma and Weenink 2022) in a careful manner at an intermediate speech rate. He was also asked to listen to the first recorded sentence and take it as a model in order to ensure that each sentence lasted for a similar length of time.

3.1.4 Stimuli preparation

Using Praat (Boersma and Weenink 2022), we annotated each sentence and marked the boundaries of vowels and consonants segment by segment. We then produced
two versions of each sentence, one in which vowels were replaced by silence (CO) and the other in which consonants were replaced by silence (VO), a technique which has been used by several other researchers (e.g., Owren and Cardillo 2006). In determining the boundaries of segments, we made a sequence of decisions and consulted several acoustics sources such as Bird et al. (2019), Wright and Nichols (2016), and Reetz and Jongman (2020). In marking the boundaries of all stops, closure and burst were included in the stop duration.

For voiced stops, the voice bar appearing in the spectrogram was used as evidence for the stop-vowel distinction and was included in the stop duration. For voiceless stops, the absence of F1 was taken as evidence for the short aspiration noise that follows the release burst and precedes the following vowel, and was included in the consonant duration. The onset of voicing was marked as the beginning of the following vowels. One problem arose with sentences starting with a target word that had an initial stop, such as /taabaʕ/ ‘followed up’. It was difficult to decide how long the closure should be, and there were two possible solutions. One was to look at the same word in final position where it was preceded by a vowel and to use the same closure duration for the same word in initial position. The other solution was to calculate the exact segment duration in all sentences and use the average closure length. The latter solution was preferred since a final position may not mirror the initial position, as the former is subject to lengthening. Among stops, the two Arabic nasal stops, /n/ and /m/, were sometimes hard to distinguish, especially when occurring pre- or post-vocalically, but we applied some techniques to identify nasals boundaries. Nasals were identified by the relative lack of aperiodicity, simple waveform pattern, and low frequency spectral components. When a nasal was preceded by a vowel, the sharply-down-pointing formants on the spectrogram were used as an indication of the nasal onset, and the sudden increase in intensity at the beginning of the following vowel was taken as a marker for the ending of the vowel.

Fricatives were relatively easy to identify and mark. The beginning and ending of the fricative noise or turbulence were identified as the duration determinants of the fricative. Vowels that occurred after fricatives were marked at the point where the higher formants started to become apparent on the spectrogram, while vowels that occurred before fricatives were marked by a drop in intensity and loss of energy, especially in higher frequencies. For the approximants, the decision on where to mark the boundaries was based on the abrupt change of intensity in F2 and F3 noticeable in the spectrogram, coupled with the help of the dramatic change of the overall amplitude in the waveform. To identify the liquids and glides boundaries as well as the preceding or following vowels, we relied on the liquid/glide-vowel transition of formants where F2 and F3 start to show some signs of movement and the stabilization points where the formants start to maintain their shape. We made the
decision to include the transitions into the vowels rather than the consonants; this was dictated by the hypothesis. The reason is because, as Wright and Nichols (2016) pointed out, the experimenters should make the decision that introduces as little bias as possible to the research. The Arabic system may support the superiority of consonants, and therefore, it was necessary to avoid any possible factor that, albeit minimally, could potentially affect the results. One important consideration was how to deal with the Arabic /r/. In previous studies, the English /ɹ/ was problematic and researchers had to make some decisions on whether to consider it as a separate consonant or as part of vowels (r-colored). The Arabic /r/, however, is a trill that is, in most cases, clearly visible on the waveform and spectrogram as a triple of the same segment and therefore, we considered it as a separate consonant.

3.1.5 Task

On each trial, participants were presented with the target stimulus twice. After listening to both repetitions, their task was to say out loud what they had heard. Their responses were recorded using a high-quality microphone and digital recorder. The spoken responses, rather than written responses, were used to avoid potential difficulties with the Arabic orthographical system, which does not explicitly represent most vowels. The participants were encouraged to report whatever they could understand regardless of whether it was a complete and meaningful sentence or not.

3.1.6 Participants

Twenty native Arabic speakers (15 males and 5 females) volunteered to participate in the study. The participants were from different Arabic-speaking countries (Saudi Arabia, Jordan, Iraq, Syria, and Morocco) and spoke MSA as well as different dialects, such as Najdi, Hijazi, Southern Saudi, Jordanian, Syrian, Iraqi, and Moroccan Arabic. All were students (mostly graduate, age range = 26–40, age $M = 31$) at the University of Wisconsin-Milwaukee, or other universities in the area. They all reported Arabic as their sole native language; in addition, all spoke English as a second language. Some participants were linguistics students, but none of them had prior exposure to the stimuli.

3.1.7 Procedure

Each participant was met individually in a quiet laboratory environment, where they were seated in front of a desktop computer equipped with high-quality headphones,
adjacent to the microphone and recording equipment. The experimenters instructed each participant on the task, which he or she then completed at his or her own pace. The experiment began with six familiarization trials. Order of trial presentation was randomized for each participant. If participants mistakenly recorded their response after the first repetition, they were instructed to listen to the second repetition and re-record their response; only four trials of this type occurred.

3.1.8 Data scoring and analysis

Following previous authors (e.g., Fogerty et al. 2012; Kewley-Port et al. 2007), we first assigned each correct word a score of 1, and each incorrect word a score of 0. Then, to calculate overall word recognition per sentence, we divided the number of correct words by six (i.e., the total number of words in each sentence), which produced a percent correct. GEEs were utilized to run a repeated-measures regression analysis using the row number of words as the count dependent variable.

3.2 Results

As shown in Table 2, the results show that word recognition per sentence was higher in the CO condition (75.06%) than in the VO condition (53.11%). Figure 1 shows the mean, median, range, and distribution of recognized words. We see that, on average, participants identified 4.5 words in the CO condition, compared to 3.19 words in the VO condition.

GEEs were used to run a repeated-measures regression analysis, using the number of correctly-identified words as the count dependent variable. As shown in Table 3, the model showed that CO versus VO predicted word recognition, Wald $\chi^2(1) = 18.448, p = 0.001$. The individual parameters table (Table 4) shows that speech in the CO category was significantly predicted to have a log count 1.507 times ($\approx 60.11\%$) higher than in the VO category, $b = 0.410$, $SE = 0.0955$, 95% CI = [0.223, 0.597], Wald $\chi^2(1) = 18.448, p = 0.001$, Exp(B) = 1.507 ($\approx 61.11\%$), 95% CI = [1.250, 1.818].

Table 2: Word recognition (%) per sentence broken down by CO versus VO in Experiment 1.

<table>
<thead>
<tr>
<th>CO vs. VO</th>
<th>Responses</th>
<th>Overall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO</td>
<td>240</td>
<td>75.06%</td>
</tr>
<tr>
<td>VO</td>
<td>240</td>
<td>53.11%</td>
</tr>
</tbody>
</table>
3.3 Interim discussion

Experiment 1 was designed to examine segmental contributions to word recognition in Arabic sentences. It was predicted (and confirmed) that, due to substantial differences between nonconcatenative and concatenative languages with regard to their morphological architecture, segments contribute differently to word recognition in Arabic compared to in other concatenative languages such as English. While previous studies (e.g., Cole et al. 1996; Fogerty et al. 2012) have shown that vowels clearly contribute to word recognition at the sentence level more than consonants do, our study clearly shows an opposite finding. In particular, the results show that despite the fact that each stimulus sentence contained equal numbers of consonants and vowels, word recognition was higher when participants were presented with only consonantal information than when presented with only vocalic information. This difference between our results and those of similar studies on concatenative languages is dramatic. For example, Cole et al. (1996) reported no cases of complete word recognition in fluent speech when only consonants were preserved in English.

![Figure 1: Word recognition (# of correctly-identified words per sentence) broken down by CO versus VO in Experiment 1.](image)

<table>
<thead>
<tr>
<th>Source</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>1239.104</td>
<td>1</td>
<td>0.001</td>
</tr>
<tr>
<td>CO vs. VO</td>
<td>18.448</td>
<td>1</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Dependent variable: Word recognition (# of words per sentence). Model: (Intercept), CO versus VO.
sentences. Similarly, in other studies of sentences, Fogerty et al. (2012) reported approximately 18% word recognition when only consonants were preserved, and Kewley-Port et al. (2007) concluded that consonantal information is only responsible for approximately 30% of word recognition in non-noisy environments.

It is not surprising that the pattern in the current study is different from that reported in the literature given the unique morphological system of Arabic. All previous studies used stimuli from languages that do not assign different morphological roles to consonants versus vowels, such as English, Dutch, and Spanish. Thus, these studies were obliged to focus primarily on the acoustic information conveyed by consonants versus vowels, and not their morphological properties. Hence, it seems reasonable that their conclusions about segmental contributions were overwhelmingly based in phonetics. Arabic, as well as other languages with nonconcatenative systems, clusters its lexicon around the consonantal root used to organize, access, and recognize lexical items (see e.g., Abu-Rabia 1997; Bentin and Feldman 1990; Deutsch et al. 2000; Deutsch and Frost 2003; Feldman et al. 1995; Holes 2004; Ravid 2003; Shimron 2003 among many others). This suggests a consonantal privilege that allows listeners to recognize and process words while relying primarily on the information available in the root. The current experiment, with its high word recognition in the CO condition, provides further support for this notion.

Our sentence stimuli were crucially balanced so as to contain 23 consonants and 23 vowels each. This is an important methodological point that controls for an important factor especially in Arabic. For example, Cole et al. (1996) only assessed the total number of vowels and consonants in the whole set of stimuli, rather than on a sentence-by-sentence basis as we did here. In Experiment 2, we explicitly manipulated segmental ratio at the sentence level to test whether consonantal information will retain or instead strengthen its superior influence as the consonantal roots will be less interrupted by the silences.

### Table 4: Summary of the GEE parameters in Experiment 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std. error</th>
<th>95% Wald CI</th>
<th>Hypothesis test</th>
<th>Exp(B)</th>
<th>95% Wald CI for Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Intercept)</td>
<td>1.136</td>
<td>0.0719</td>
<td>0.995</td>
<td>1.277</td>
<td>249.918</td>
<td>1 0.000 3.114 2.705 3.586</td>
</tr>
<tr>
<td>[CO vs. VO = 1.00]</td>
<td>0.410</td>
<td>0.0955</td>
<td>0.223</td>
<td>0.597</td>
<td>18.448</td>
<td>1 0.000 1.507 1.250 1.818</td>
</tr>
<tr>
<td>[CO vs. VO = 0.00]</td>
<td>0a</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>1</td>
</tr>
<tr>
<td>(Negative binomial)</td>
<td>4.152</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent variable: Word recognition (# of words per sentence). Model: (Intercept), CO versus VO.
4 Experiment 2

Experiment 2 tests a different segmental ratio in complete sentences. Specifically, whereas the vowel-to-consonant ratio in Experiment 1 was 23:23, the ratio in Experiment 2 is 13:23 (i.e., 0.57:1); that is, the number of consonants were kept the same while the number of vowels were reduced by 10 segments. If consonants contribute more than vowels to word recognition in Arabic sentences, word recognition will improve in the CO condition even with the same number of consonants (23 Cs). The time interval between the two experiments was roughly 6 months.

4.1 Methodology

4.1.1 Sentence construction

The stimuli consisted of 24 sentences. The average number of words in each sentence was kept approximately the same as in Experiment 2, \( M = 5.87 \) words, although the total number of segments differed (13 Vs vs. 23 Cs). All other construction criteria and decisions regarding experimental design were followed as in Experiment 2.

Ideally, we would have also included a condition with relatively fewer consonants and relatively many vowels, but such sentences can be constructed only with great difficulty by using a disproportionate number of long vowels, adversely affecting the naturalness of most sentences and ultimately achieving only a small and possibly irrelevant difference in segment count. The example below demonstrates one sentence from the stimuli.

(2) lam ja-kun baʔdʕ-u sʕʕ-sʕʕahb-i
not.PRF 3PSM-be some-NOM the-companions-GEN

muktariθ-an bi-lʔamr
concerned-ACC with-the-matter

‘Some of the companions were not concerned about the matter.’

4.1.2 Judgements

We submitted all sentences in written and spoken forms to the same 10 native speakers of Arabic from the previous experiment to judge sentences for naturalness and likelihood, using the same criteria as in Experiment 1.
4.1.3 Recording and stimuli preparation

We followed the same steps taken in Experiment 1. The speaker of the new stimuli was the same speaker who recorded the stimuli in Experiment 1.

4.1.4 Task

The participants performed the same tasks performed in the previous experiment.

4.1.5 Participants

Twenty native Arabic speakers participated in the study. The participants come from different Arab countries and speak different dialects. Ten of them have participated in the previous experiment, but participants were as close as possible to the continuing participants in terms of age (age ranges from 23 to 42, \( M = 33 \)), the foreign languages they speak, and education level. They had no exposure to the stimuli during any phase of the sentence construction, sentence rating, or stimuli recording.

4.1.6 Procedure

We followed the same procedure as in Experiment 1. The two experiments were separated by six months and a few days.

4.1.7 Data scoring and analysis

We scored responses according to word recognition per sentence. The same scoring steps in Experiment 1 were also applied to the data in Experiment 2.

4.2 Results

Descriptive results show that word recognition was much better in the CO condition (90.41%) than in the VO condition (28%), as shown in Table 5. As shown in Figure 2, the mean number of recognized words per sentence was 5.34 words in the CO condition and 1.65 words in the VO condition. The boxplot in Figure 2 shows almost opposite patterns for the two conditions. The boxplot shows that the number of recognized words in the CO condition ranged between four and six words (as represented by the lower and upper whiskers), with 6 as the median. In contrast, the number of recognized words in the VO condition ranged between zero and three words, with 1
as the median. The figure also shows more variability in CO compared to VO. The dot-plot in Figure 3 shows a similar pattern.

Following the same procedure as in Experiment 1, GEEs were used to perform a repeated-measures negative binomial regression analysis. The model showed that CO versus VO statistically significantly predicted word recognition, Wald $\chi^2(1) = 95.277$, $p = 0.001$ (see Table 6 below).

Since there are no other independent factors in the model, the results here are sufficient to show that CO versus VO is a significant explanatory factor, but Table 7 provides the individual parameters. Specifically, speech in the CO category was significantly predicted to have a higher log count at 3.260 times ($\approx 72.52\%$) more than that in the VO category, $b = 1.182$, $SE = 0.1211$, 95% CI = [0.9451, 419], Wald $\chi^2(1) = 95.277$, $p = 0.001$, Exp(B) $= 3.260$ ($\approx 72.52\%$), 95% CI = [2.572, 4.134].

Table 8 compares the results from both experiments. It shows that word recognition was higher in the CO conditions than in the VO conditions in both experiments, higher in the CO condition from Experiment 2 than in the CO condition from Experiment 1, and higher in the VO condition from Experiment 1 than in the VO condition from Experiment 2.
4.3 Interim discussion

Compared to the 23:23 segmental ratio used in Experiment 1, the 13:23 segmental ratio used in Experiment 2 improved word recognition in the CO by approximately 15%. At
the same time, word recognition deteriorated in the VO condition by approximately 20%. Again, these results are completely different from those reported for other languages such as English. In Experiment 2, the disparity between participants’ performance in the CO condition and the VO condition, 90.41–28.00% = 62.41, is large. We can confidently assert that the absence of vocalic information in the CO condition had only a marginal impact on word recognition in sentences. The results from this experiment make for an interesting comparison with those of Cole et al. (1996), who reported a mean rate of 86.40% correct for unmodified (i.e., fully intact) English sentences. By contrast, we found a mean rate of 90.41% correct for CO Arabic sentences. Furthermore, in Experiment 2, word recognition in the VO condition (28.00% correct) would appear to be the lowest value reported for a VO condition in the literature so far.

Comparing results across experiments, the degree of word recognition in the CO condition in Experiment 2 is interesting given that the amount of consonantal information in the CO condition was exactly the same as that in the CO condition from Experiment 1 (i.e., 23 segments). This could be due to comparably few occurrences of silence replacements in the CO condition in Experiment 2 as compared to Experiment 1, making the speech stream less interrupted and hence more intelligible. Thus, with fewer silence replacements participants were better able to recognize lexical boundaries and assign consonants to their correct roots. It is important to note that, even though Experiments 1 and 2 recruited ten of the same participants, the improved performance in Experiment 2 would not be a matter of practice effect, as the interval between the two experiments was over six months.

The imbalanced segmental ratio used in Experiment 2, rather than that in Experiment 1, would be a close match to the ratio used in previous studies on English as the morphology of English allows for more consonant clusters than that of Arabic does. Nevertheless, as we have pointed out, our results differ dramatically from what such studies have reported. This difference between Arabic and other languages such as English in terms of segmental contribution to word recognition in sentences (i.e., consonant superiority) is similar to that found between Chinese (Chen et al. 2015) and other languages such as English with regard to segmental contributions to the

Table 8: A comparison between the results in Experiments 1 and 2.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Intelligibility</th>
<th>Mean percent correct</th>
<th>Mean # of recognized words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td></td>
<td>75.06</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td>Mean # of recognized words</td>
<td>53.11</td>
<td>3.19</td>
</tr>
<tr>
<td>Experiment 2</td>
<td></td>
<td>90.41</td>
<td>5.34</td>
</tr>
<tr>
<td></td>
<td>Mean percent correct</td>
<td>28.00</td>
<td>1.65</td>
</tr>
</tbody>
</table>
recognition of words. In languages such as English, vowels contribute to word recognition in complete sentences more than consonants, while consonants contribute more when words are presented in isolation. Studies on Arabic (including the current one) have shown that consonants contribute more than vowels both in isolated words and complete sentences, while studies on Chinese have shown exactly the reverse.

The current findings add to the literature on word recognition in Arabic using experimental techniques such as priming (e.g., Boudelaa and colleagues) and show a robust consonantal privilege in the recognition of Arabic words. To the best of our knowledge, the silence-replacement method implemented at the sentence level in the current study has not been utilized to examine root-and-pattern morphology in previous studies. The use of this technique here does not only provide a new way to test the root-and-pattern framework, but also allows us to examine the impact of the interruption that occurs to the consonantal root as part of word formation. When consonants were less interrupted by the silences in the current experiment compared to the case in Experiment 1, word recognition improved. This may indicate that Arabic speakers process speech by “extracting” the consonantal root, which has been alluded to in previous studies (e.g., Bentin and Feldman 1990). This mechanism of extraction is different from the “decomposition” mechanism found in stem-affix languages (e.g., Caramazza et al. 1988; Taft 1983; Taft and Forster 1975).

5 General discussion and conclusion

Our findings illustrate that consonantal information is the primary source for lexical information in Arabic and provide new evidence that consonants in Arabic make stronger contributions to word recognition both in isolation and in sentences. Although this partially concurs with what has been reported in studies on word recognition in isolation on non-Semitic languages, such as van Ooijen (1996), Cutler et al. (2000), and Owren and Cardillo (2006), the interpretation presented in the current study is different. We attribute this consonantal privilege to the role consonants play in the root-and-pattern morphological system of Arabic, and claim that a lack of consonantal information hampers bottom-up processing. As many previous studies have shown, lexical items in Semitic languages are mainly built from a string of consonants that get interdigitated with another string of vowels. The current findings are in line with these previous results, and add to them by specifically evaluating segmental contributions to word recognition at the sentence level. Indeed, our results go one step further by showing that consonants are paramount not just for word recognition in isolation, but also in complete sentences.

The difference between the results from early research on the same topic and the present experiments is twofold. First, word recognition in the CO sentences was
considerably higher than that in the VO sentences, which clearly contrasts with what has been reported in other non-Semitic languages such as English (e.g., Fogerty et al. 2012). Second, in comparison to previous studies (e.g., Fogerty and Humes 2012; Fogerty et al. 2012), the role of suprasegmental and contextual information in the VO sentences was extremely limited in the current study. In fact, sometimes the listener’s ability to recognize the pattern in the VO led to an incorrect answer. That is, in taking a closer look at results from individual trials, there were instances in which the participants successfully recognized the word structure (correct vowels) but failed to recognize the intended words (incorrect consonants). For instance, a verb-subject (VS) sentence in Experiment 2 starting with the two words /taḥaddaθa/ ‘talked’ and /muhalliluun/ ‘analysts’ (‘analysts talked’) was reported as /takallama/ ‘spoke’/muhaddiθuun/ ‘talkers’ (‘talkers spoke’) in the VO condition by several participants. Similarly, in the VO condition, some participants provided the response /tˤaaʔir-at-un kabiir-at-un/ ‘big plane’ for the intended two words /ðˤaahir-at-un xatir-at-un/ ‘dangerous phenomenon’. Although the intended words were missed and we scored such responses as incorrect (following the no-error-tolerance scoring standard in previous studies such as Kewley-Port et al. (2007) and Fogerty et al. (2012), the vowels are nevertheless identical in both the presented sentence and the reported sentence, and on this basis, the participants were able to identify the first item as being verbal and the second as being nominal. We may speculate that there were some other instances in which such structural information helped the participants to correctly recognize the intended lexical item, which would suggest that structural information enhanced word recognition in the VO condition at the sentence level in Arabic. However, as stated above, this was minimal.

One theoretical implication for our study is the observation that a phonetics-only approach to word recognition is insufficient. The magnitude and type of acoustic information carried by consonants versus vowels may not alone be responsible for word recognition; rather, the way a language structurally employs each segment type – such as the semantic information in the Arabic root (consonants) versus the morphosyntactic information in the Arabic pattern (vowels) – is also crucial. This means that morphology intersects with acoustic cues to modulate the contribution of each segment to word recognition. However, theories such as the Cohort Model (Marslen-Wilson and Welsh 1978) and Shortlist (Norris 1994) treated vowels and consonants as interchangeable units. In these models, an initial segment activates some lexical candidates (a “cohort”) and another eliminates some competitors. However, if vowels versus consonants provide different amounts of information about the identity of a word, this raises a question as to whether vowels versus consonants also play different roles in activating cohort candidates. If consonants are robust while vowels are prone to contextual effects, then we need to come up with a more detailed approach that incorporates such differential contributions.
Another implication for the current study is of methodological nature. Despite the fact that the role of the root in mental representations and lexical access in Arabic has long been recognized by researchers, many previous studies on Semitic languages focused on the recognition of isolated words rather than entire sentences. That is, evidence for the root as a morphemic unit has been drawn from many morphological experiments using partial or complete words. In the results of the current study, consonantal information as present in the root turned out to be crucial not only at the word level but also at the sentence level, and the replacement paradigm borrowed from studies on non-Semitic languages can be used for future research to investigate the root-and-pattern approach further. In contrast, the morphological pattern represented by vowels was not sufficient for word recognition at the word level, nor was it as useful as expected for word recognition at the sentence level. This is plausible because a given Arabic root can generate only a limited set of possible words, but a given pattern can be applied to a very large number of words. For instance, the root “ʔʃq” can generate different words that are related to LOVE, but the pattern CaaCiC, which is used to form the active participle/doer of something, can be used to form thousands of words. Thus, it is possible that participants could use a given pattern to falsely recognize an unintended word. For instance, in the VO condition, some participants provided the response /tˤaahir-at-un kabir-at-un/ ‘big plane’ for the intended two words /ʔaahir-at-un xatiir-at-un/ ‘dangerous phenomenon’. In nonconcatenative languages, vowels provide information about context, position within the sentence, intonation, and other supra-segmental information that can potentially assist in determining structure (e.g., noun vs. verb, singular vs. plural, passive vs. active, etc.).

The manipulation of vowel-to-consonant ratio also affected our results. Reducing the number of vowels in the target stimulus resulted in better word recognition than when equal numbers of vowels and consonants were presented. Although these effects were overridden by the strong role played by vowels and consonants themselves, the results nevertheless provide strong supporting evidence for the importance of consonants in Arabic word recognition. One remaining issue related to the ratio used in this experiment is the concern that silence- or noise-replacement research has raised about coarticulatory information that may allow vowels to provide some information about the consonants but not the opposite. For example, Cole et al. (1996) took care to eliminate any coarticulation effects in one of their experiments, and a reduction/expansion of 10 ms was implemented, and Owren and Cardillo (2006) increased the magnitude of reduction to 50% of the entire segment, but that did not change the role of vowels versus segments. Manipulating the boundaries to include or eliminate coarticulatory information has been found to have an imperceptible or no effect on the contribution of vowels versus consonants in word recognition in isolation or in sentences (e.g., Fogerty and Kewley-Port 2009).
The manipulation of vowel-to-consonant ratio in the current experiment and the previous one did not change the overall role of consonants in word recognition. In other words, the vocalic information was roughly 50% (23 Vs vs. 23 Cs) in the first experiment, but that figure would increase if we assumed that the vowels also carried some coarticulation information from their adjacent consonants. Notwithstanding, word recognition was poor in the VO condition, which, again, indicates the strong role of consonantal root as a lexical unit in word processing of Arabic.

The results from both experiments are directly relevant to theories of speech processing in general (i.e., not only in Semitic languages). Research has established two preeminent modes of speech processing: bottom-up and top-down. In bottom-up processing, computation of segmental information is the fundamental operation that occurs first. This approach has been supported by many scholars in speech science and psycholinguistics (see e.g., Cutler et al. 1987). Some studies even go beyond arguing for bottom-up processing to stress that speech processing is monodirectional in a strictly bottom-up manner, such that top-down processing could hinder word recognition and speech intelligibility (e.g., Norris et al. 2000). Such a thesis states that the flow of information from segments to words (or even sentences) is always necessary for word recognition, but backward feedback from words to segments is unnecessary or probably implausible. The results from the current study are in line with this emphasis on the role of segments for word recognition, but since Arabic differentiates vowels and consonants, it goes beyond this claim. That is, the acoustic information that comes from segments is crucial for word recognition, but root-and-pattern morphology imbalances this segmental contribution to the advantage of consonants.

In a top-down model, on the other hand, higher levels of information (i.e., contextual information) is processed to recognize words and comprehend speech, even with limited segmental information (see e.g., Samuel 1997). For example, if the listener believes she has recognized the word “thrift”, this higher-level information influences her perception of lower-level phonemes, biasing her to believe she has perceived the segments /θ/, /ɹ/, /ɪ/, /f/, and /t/. Full sentences provide the canonical example of top-down information because they can provide sufficiently strong contexts for a listener to essentially “restore” entire words without access to bottom-up information from speech. For example, provided with the sentential context “The touch of soft fur and the sound of a meow made me realize that a ____ had entered the room”, English listeners will invariably hear cat. In studies such as Fogerty et al. (2012), this contextual information was illustrated in the high rate of English word recognition at the sentence level, and was even claimed to be responsible for the different roles that consonants and vowels played in word recognition in two types of stimuli, words versus sentences. In their study, it was
suggested that the effect of top-down processing is strong. This is not the case in our current study, arguably due to the fact that Arabic, as in other Semitic languages, is better accounted for by the root-and-pattern framework.

Our study also contributes to research on word recognition and speech processing in its use of full sentences. Models of word recognition differ primarily in terms of how they implement top-down versus bottom-up approaches. The dispute between the two approaches is centered around irreconcilable evidence from different experimental tasks and techniques such as phonemic categorization and phonemic restoration. A number of previous studies attempted to evaluate the amount of bottom-up versus top-down information needed for speech intelligibility (see e.g., Samuel 2001), using words and sentences. That is, such studies tried to examine how much intelligibility occurs as a function of segmental information versus non-segmental information, such as sentential context. In general, the conclusion from this series of studies was that the top-down effect is real but fragile, and the necessity of an interactive account for top-down and bottom-up modelling may be questionable. Nevertheless, other studies using the replacement paradigm such as Fogerty et al. (2012) and Chen et al. (2013) suggest that the effect of top-down processing on speech processing is evident. Here, we further show that the discussion on bottom-up and top-down speech processing is not only relevant to the perception of segmental versus non-segmental information, but also to the morphological architecture of a given language. As such, models of the lexicon should take this into consideration.

To summarize, the current study investigated the role of segmental contributions to word recognition at the sentence level in Arabic, using a replacement paradigm used in similar previous studies. The pattern of segmental contributions found in Arabic consistently presents consonants as the primary contributor to word recognition at the sentence level. This pattern is quite different from what has been reported for English and other non-Semitic languages, but supports previous findings that have reported consonants as an important segment for Arabic word recognition at the word level. We attribute this difference to the root-and-pattern morphology that favors consonants in Arabic. Future work on word recognition and speech intelligibility is called to focus on word recognition at the sentence level to further examine the root-and-pattern approach to Arabic morphology.2

**Research funding:** Researchers Supporting Project number (RSPD2023R580), King Saud University, Riyadh, Saudi Arabia.

---

2 This paper is based on Aldholmi’s (2018) dissertation research conducted while he was a PhD student at the University of Wisconsin-Milwaukee, USA.
References


Cole, Ronald A., Yonghong Yan, Brian Mak, Mark Fanty & Troy Bailey. 1996. The contribution of consonants versus vowels to word recognition in fluent speech. In *International conference on acoustics, speech, and signal processing (ICASSP’96)*. Atlanta, GA: IEEE.


