2011

Cross Linguistic Differences in the Immediate Serial Recall of Consonants versus Vowels

Elizabeth M. Kissling
University of Richmond, ekisslin@richmond.edu

This is a pre-publication author manuscript of the final, published article.

Cross Linguistic Differences in the Immediate Serial Recall of Consonants versus Vowels

Short Title: L1 and ISR of Consonants versus Vowels

Elizabeth M. Kissling

Georgetown University

Georgetown University
Department of Linguistics
ICC 479
37th St. and O St., NW
Washington, DC 20057

telephone: (202) 687-5956
fax: (202) 687-6174
e-mail: emk47@georgetown.edu

ABSTRACT

The current study investigated native English and native Arabic speakers' phonological short term memory (PSTM) for sequences of consonants and vowels. PSTM was assessed in immediate serial recall tasks conducted in Arabic and English for both groups. Participants (n=39) heard series of 6 CV syllables and wrote down what they recalled. Native speakers of English recalled the vowel series better than consonant series in English and in Arabic, which was not true of native Arabic speakers. An analysis of variance showed that there was an interaction between first language (L1) and phoneme type. The results are discussed in light of current research on consonant and vowel processing.

INTRODUCTION

The differential processing of consonants and vowels is a phenomenon that has been investigated by researchers working in several areas, including infant word learning (e.g., Nazzi, 2005), visual and auditory word recognition (e.g., Carreiras, Vergara, & Perea, 2009; Toro, Nespor, Mehler, & Bonatti, 2008), neuropsychological disorders (Boatman, Hall, Goldstein, Lesser, & Gordon, 1997; Caramazza, Chialant, Capasso, & Miceli, 2000), and models of short term memory (e.g., Drewnowski, 1980). Work in all these areas indicates that consonants and vowels are processed differently, across the various languages and populations studied.

Several studies (Bonatti, Peña, Nespor, & Mehler, 2005; Carreiras, Duñabeitia, & Molinaro, 2009; Carreiras, Gillon-Dowens, Vergara, & Perea, 2009; Carreiras, Vergara, et al., 2009; Cutler, Sebastian-Galles, Soler-Vilageliu, & Van Ooijen, 2000; Havy, Bertoncini & Nazzi, 2011; Havy & Nazzi, 2009; Nazzi, 2005; Nazzi, Floccia, Moquet & Butler, 2009; Nazzi & New, 2007; Nespor, Peña, & Mehler, 2003; New, Araujo, & Nazzi, 2008; Van Ooijen, 1996) have suggested that there is a consonantal lexical processing bias, that is, that consonants give better cues to the lexicon than do vowels, and the evidence they provide comes from a variety of experimental tasks. While French-learning young infants (0 to 2 months) were able to detect phonetic differences between different syllables, their representations of the syllables did not seem to be structured in terms of phonetic segments, i.e., consonants and vowels (Bertoncini, Bijeljac-Babic, Jusczyk, & Kennedy, 1988; Jusczyk & Derrah, 1987), yet their representations change over the first year of life. By 16 to 20 months, French-learning infants taught names for novel objects were able to learn pairs of new words that contained contrasting consonantal phonemes, but they did not learn pairs that contained contrasting vocalic phonemes, even though they were able to distinguish both the consonantal and the vocalic phonemes tested (Havy &

Nazzi, 2009; Nazzi, 2005; Nazzi & New, 2007). This consonant bias was found independent of position of the phone in the speech signal (Nazzi, 2005; Nazzi & Bertoncini, 2009) and the consonant type (e.g., stop, fricative, or liquid) (Nazzi, 2005; Nazzi & New, 2007) for infants 20 months of age. Though by 30 months of age French-learning children were able to learn novel word pairs that differed only in vowels (Nazzi, et al., 2009), they demonstrated a consonant bias in word learning until at least 3 years of age (Havy, et al., 2011). English-learning children at 30 months of age demonstrated a consonant bias equal to that of their French counterparts when experimental conditions were held constant (Nazzi, et al., 2009).

In spoken word recognition experiments, adult native speakers of Spanish (Cutler, et al., 2000), Dutch (Cutler, et al., 2000) and English (Van Ooijen, 1996) were found to use consonants as lexical cues more than vowels. In these experiments, participants performed word reconstruction tasks in which they were presented with non-words and asked to change them into real words by changing a single sound. Participants were faster and more accurate when changing a single vowel sound rather than a consonant in order to make non-words into real words (e.g., “kebra” became “cobra”, not “zebra”). The consonantal lexical processing bias was found for all L1 groups tested, even though the phonemic repertoire and the relative distinctiveness of the phonemes within the repertoire differed across languages.

The same processing bias has been found in visual word recognition experiments using masked priming techniques and measuring both behavioral (New, et al., 2008) and ERP responses (Carreiras, Duñabeitia, et al., 2009; Carreiras, Gillon-Dowens, et al., 2009; Carreiras, Vergara, et al., 2009). New et al. (2008) submitted French university students to a visual masked-priming lexical-decision task and found priming effects for identical primes (e.g., for the word *joli*, *joli*) and consonant-related primes (*jalu*), but not for vowel-related primes (*vobi*) or unrelated primes (*vabu*), suggesting that consonants constrain lexical access in a way that vowels do not. Carreiras and colleagues (Carreiras, Duñabeitia, et al, 2009; Carreiras, Gillon-Dowens, et al., 2009; Carreiras, Vergara, et al., 2009) conducted a series of visual masked-priming lexical-decision task experiments with Spanish university students, utilizing both the transposed-letter and delayed letter masked prime paradigms, and found priming effects for identical and consonant-related primes, even transposed consonant primes, but not for vowel-related or unrelated primes. They also reported parallels between identical primes and consonant-related primes but not vowel-related primes in terms of ERP data (N250 and N400), suggesting that consonants and vowels are processed at successive temporal stages during visual word recognition (Carreiras, Duñabeitia, et al., 2009). Carreiras and colleagues argue that computational models of visual word recognition should be revised to account for the status of letters as either consonants or vowels. Indeed, New et al. (2008) also conclude that their findings challenge current models of word reading (p. 1226) because they do not treat vowels and consonants as categorically different (though see Berent and Perfetti's (1995) two cycles model of reading for an exception).

The consonantal lexical processing bias has also been found in auditory speech processing experiments in which university students, both native speakers of French (Bonatti, et al., 2005) and Italian (Toro, et al., 2008), listened to continuous streams that contained nonsense syllables as well as "words" in an artificial language. When asked to isolate words from the stream, participants were able to learn transitional probabilities and locate words when the context was that of fixed consonants with variable vowels, but they were unable to do so if the consonants were varied and the vowels were fixed. Conversely, they used vowels more when detecting structural regularities or "rules" in the stream. These results support the Nespor et al.

(2003) "CV hypothesis," which states that consonants provide cues primarily about the lexicon, whereas vowels provide cues primarily about morphosyntax.

In sum, the research on consonant-vowel differences indicates that a consonantal lexical processing bias exists across languages and age groups. One might be led to hypothesize then that consonants are universally processed more rapidly and/or more deeply than vowels and expect an advantage for consonants to obtain in short term memory (STM) experiments. However, the few STM studies that have directly compared consonants to vowels (Cole, 1973; Crowder, 1971; Drewnowski, 1980) have found just the opposite to be true. In these studies, native English speakers recalled vowels better than consonants during immediate serial recall tasks. In one of a series of immediate serial recall experiments, participants (Yale university students, n=32) read 60 stimulus sentences consisting of either 7-syllable consonant-variable series made of the syllables "Bah, Dah, Gah" or 7-syllable vowel-variable series made of the syllables "Bee, Boo, Bih" (Crowder, 1971). The syllables appeared for 2 seconds, after which time participants wrote down what they recalled. Lower error rates obtained for all positions in the vowel-variable series as compared to the consonants. Drewnowski (1980) also performed a series of experiments that presented participants visually with series of English-like syllables, some series differing only in consonants and other series differing only in vowels, and found that participants recalled the order of vowels better than consonants, particularly in later positions in the series. The syllables were formed with 3 stops ([b, d, g]), 3 fricatives ([f, s, z]), and 6 vowel phonemes, both short ([æ, ɜ, ɪ]) and long ([e, ɔɪ, ɑɪ], the vowels being spelled as AH, EH, IH, EY, OY, IY, respectively. Syllables were presented at a rate of 2 items per second on a memory drum. The experiments reported in Drewnowski varied in terms of modality of presentation (visual and/or auditory) and recall (vocalized or silent), series length (3- or 6-syllables) and phones included (short and/or long vowels). In general, Drewnowski concluded that vowels were remembered better than consonants, particularly towards the end of a series. Cole (1973) found identical results in immediate serial recall tasks with both English CV and VC syllables when they were presented auditorily, though not when they were presented visually. Cole concluded that the limited capacity acoustic storage component of STM preserved vowels longer than consonants.

These STM researchers (Cole, 1973; Crowder, 1971; Drewnowski, 1980) reached the perhaps surprising conclusion that vowels are remembered better than consonants, at least by native English speakers, but they stopped short of sufficiently explaining the mechanism underlying this differential ability in STM. Crowder suggested that vowels may be remembered better because their acoustic cues are longer in duration than those of consonants, and the increased duration of the vowels may result in vowels forming better traces in STM. Even though the presentation of stimuli in Crowder was visual, participants presumably made use of the subvocal articulatory rehearsal mechanism believed to preserve memory traces in the phonological loop of the working memory system (Baddeley, 1986, 2000, 2003; Baddeley & Hitch, 1974), such that they processed the stimuli as an acoustic signal containing cues for consonantal and vocalic phonemes. Drewnowski did not provide a explanation for the STM advantage of vowels, though the fact that long vowels were recalled better than short vowels and fricative consonants were recalled better than stops does lend some credence to Crowder’s suggestion that duration of stimulus has an effect on performance in immediate serial recall tasks and may underlie the STM advantage for vowels over consonants.

The current study attempts to replicate the results of Crowder (1971), Drewnowski (1980) and Cole (1973) with English and also expand the investigation of consonants and vowels in STM to include a language other than English. Specifically, the current study tests the ability of native speakers of English to recall sequences of consonants and vowels in their first language (L1), English, and their second language (L2), Arabic, as well as the ability of native speakers of Arabic to recall sequences of consonants and vowels in their L1 (Arabic) and L2 (English). Such a comparison will shed light on what might underlie the relative advantage found for vowels in STM and thus contribute to the growing body of research in consonant and vowel processing. Arabic was chosen for the current study because the relative status of consonants and vowels within its linguistic system and script differs markedly from English (and the other Indo-European languages investigated in the consonant and vowel processing literature thus far). Namely, the phonemic inventory of Arabic vowels is restricted to three short ([a, i, u]) and three long ([a:, i:, u:]) vowels and 28 consonants (Ryding, 2005), as compared to English's relatively expansive inventory of 17 vowels and 24 consonants (Nazzi, 2005); in written Arabic texts only long vowels are indicated orthographically, whereas all short vowels are invisible (Ryding); and Arabic orthography is transparent in that the grapheme-phoneme correspondences are highly regular, as compared to the more irregular grapheme-phoneme correspondences of English (Keiko & Zehler, 2008).

The research question motivating the current study was as follows: Do all participants recall sequences of vowels better than sequences of consonants, regardless of their L1 and the language of the stimuli? Following prior STM research (Cole, 1973; Crowder, 1971; Drewnowski, 1980), it was hypothesized that vowels would be remembered better than consonants. Furthermore, given that researchers have thus far not found evidence of a cross-linguistic difference in consonant and vowel processing, and some have gone as far as to say that consonant and vowel processing differences may be universal (e.g., Nespor, et al., 2003), it was hypothesized that this relative advantage in STM recall for vowels over consonants would obtain for all participants, regardless of L1 and language of stimuli.

METHOD

PARTICIPANTS

Participants were all students or staff members of a small, private university in the Eastern United States, recruited via e-mails and fliers. They were not compensated for their participation. There were 19 native speakers of Arabic and 20 native speakers of English. No participant began learning a second language before the age of 5. The group of native English speakers consisted of 10 females and 10 males whose ages ranged from 18-35 years (mean = 22.00, SD = 3.52) and who had studied Arabic for an average of 2.54 years (range of 3 months – 6 years, SD = 1.65). The group of native speakers of Arabic consisted of 10 females and 9 males whose ages ranged from 18-52 years (mean = 27.63, SD = 8.49) and who had studied English as an L2 for an average of 12.05 years (range of 4 months – 46 years, SD = 12.03). The groups were imbalanced with regards to L2 exposure, but this imbalance was difficult to correct given the geographic limitations of recruitment for the study. The native speakers of English reported also having studied Farsi (1 participant), French (9 participants), German (4 participants), Latin (1 participant), Italian (2 participants), Russian (1 participant), and Spanish (8 participants); the average amount of exposure they had to these additional languages was 4.3 years. The native speakers of Arabic reported also having studied French (6 participants), German (1 participant), Hebrew (3 participants), and Spanish (1 participant); the average amount of exposure they had to these additional languages was 5 years.

MATERIALS

A female native speaker of Arabic and a female native speaker of English (the researcher) were recorded vocalizing the following nine syllables in their native languages: [mu], [ku], [zu], [mi], [ki], [zi], [ma], [ka], and [za]. The inventory of sounds was limited to these nine syllables for several reasons. An attempt was made to choose sounds that were equivalent in both languages and unproblematic for second language learners. The vowel inventory of Arabic is quite restricted and capitalizes on distinctions of duration between long and short vowels. However, in Arabic, isolated syllables of the type CV are restricted to long vowels. Therefore, with a total of three vowels that could be used ([i], [a], [u]), it was decided that just three consonants should be used as well. The consonants used ([k], [m], [z]) were chosen because they represent a variety of articulatory places and manners, their English and Arabic varieties are virtually equivalent, and producing them vocally and orthographically is also unproblematic for even beginning L2 learners. Recordings were made with the audio editing software Soundtrack Pro with a 16 bit/48 kHz sampling rate.

A subsequent analysis of the natural language stimuli revealed that there was variation in duration across phonemes within each language and across the two languages. Table 1 displays the duration of each syllable stimulus. The difference between consonant durations in English (mean = 58.59 ms., SD = 11.67 ms.) and in Arabic (mean = 92.50 ms., SD = 33.7 ms.) was significant ($t_{(8.5)} = 2.68$, $p = .026$). However, the difference between vowel durations in English (mean = 144.44 ms., SD = 25.55 ms.) and in Arabic (mean = 210.00 ms., SD = 100.33 ms.) was not significant ($t_{(10.3)} = 2.00$, $p = .073$). As is clear from Table 1, the native Arabic speaker produced relatively lengthy [z] and [a] phones, tending to fricate [z] and diphthongize [a] more than other phones, which accounts for the difference between the Arabic- and English-language stimuli. It was decided not eliminate these differences in duration through editing of the natural language stimuli.

The recordings of the 9 English and 9 Arabic CV syllables were arranged into sequences of 6 syllables each using Soundtrack Pro. One unique token of each syllable was used for all sequences, so that all tokens of a particular syllable sounded exactly the same across sequences, and there was no recognizable prosody to the sequences. A total of 56 sequences were generated, divided into four blocks: Arabic consonants, Arabic vowels, English consonants, and English vowels. The syllables of the sequences in the consonant blocks varied only in terms of their consonant, while the vowel remained the same, e.g., [ma za ka za ka za]. The sequences in the vowel blocks varied only in terms of their vowels, e.g., [ki ka ki ku ku ka]. Syllables were distributed randomly but evenly throughout the blocks. The complete set of stimuli used is reproduced in Appendices A and B.

METHOD

The stimuli and conditions used were based on the designs of Crowder (1971) and Drewnowski (1980), adapted for use with bilingual participants and multiple language stimuli. Crowder and Drewnowski presented English-like stimuli visually and/or auditorily to their native English-speaking participants. It was assumed that when presented visually with English-like syllables the participants' processing of the phones presented would precede in a consistent, uniform and predictable manner because the orthographic-phonological relationships would be fully automatized. However, in the present study, it could not be assumed that the L2 learners would necessarily have behaved so consistently. For instance, an L2 English speaker reading the syllable *koo* multiple times might variably process it as [ku], [ku:], [ko], [ko:], or some other variant. Such variation could have become a confound. For that reason, all stimuli were presented auditorily. Crowder (1971, p. 593) concluded that for natural language stimuli, optimal performance is supported by vocalized input and silent output. To optimize their performance, the participants in this study remained silent while they reported their recollections in writing.

Participants were seated at Macintosh computer stations and listened to the stimuli through noise-reducing headphones at a volume that was comfortable for them. The stimuli were presented in four blocks: English consonants, English vowels, Arabic consonants, and Arabic vowels. The order of the four blocks was counterbalanced across participants. Each block contained 2 practice sequences followed by three sets of 4 sequences each (see Appendices A and B), with an opportunity to pause between sets. Each sequence consisted of 6 syllables. Stimuli were presented at an approximate rate of one syllable per second, but due to variations in syllable duration (as noted in Table 1), as well as minor inaccuracies in the sound editing, the presentation of the stimuli was very nearly, but not exactly, equal across blocks. The duration, in seconds, of the 6-syllable sequences in each block were: English consonant mean = 6.98, SD = 0.23; Arabic consonant mean = 6.79, SD = 0.14; English vowel mean = 6.88, SD = 0.22; Arabic vowel mean = 6.88, SD = 0.24.

After hearing the sixth and last syllable of a sequence, participants were allotted 12 seconds to respond, a response time window which was found to be appropriate during pilot testing. Participants wrote down the syllables that they recalled from the sequence on an answer sheet. All participants wrote in Arabic script during the Arabic blocks and in Roman script during the English blocks. The answer sheet accommodated differences in writing conventions, i.e. right-to-left in Arabic and left-to-right in English. Each test block was preceded by two practice sequences, which were checked immediately by the researcher to assure that participants understood the instructions. Thus a total of 48 sequences per participant were used in the subsequent analysis: 12 vowel-alternating sequences in English, 12 consonant-alternating sequences in English, 12 vowel-alternating sequences in Arabic, and 12 consonant-alternating sequences in Arabic, as shown in Appendices A and B. Each sequence consisted of 6 syllables. Therefore this set of stimuli included 144 total tokens of test consonants (i.e., consonants in consonant-alternating sequences) and 144 total tokens of test vowels (i.e., vowels in vowel-alternating sequences), half of which were in English (72) and half were in Arabic (72).

SCORING

Participants' responses were analyzed to determine the total number of correct syllables reported on each block. Response sheets provided six blanks for each sequence of six syllables, and any missing, incorrect, or misplaced syllable was counted as one error. For example, if a test string was [ma ka za ka za ma] and a participant's response was _za_ _ma_ _ka_ _ka_ _za_ _ma_, then the first three syllables were misplaced, constituting three errors, and the score for that string was counted as 3 (out of 6). Alternative spellings were not counted as errors, e.g. _zu,_ _zo,_ and _zoo_ were all accepted as possible spellings of [zu]. Errors were also analyzed in terms of error type. All errors were transpositions, incorrect syllables, or missing syllables. For instance, if a test string was [ma ka za ka za ma] and a participant's response was _za_ _ma_ _ka_ _ka_ _za_ _ma_, then the errors in the first three syllables were coded as three transposition errors. If a participant's response was _ma_ _za_ _ka_ _ka_ _za_ _za_, then the errors in the second and third syllables were coded as two transposition errors, and the error in the sixth syllable was coded as an incorrect syllable. If a participant left a syllable position blank, it was coded as one missing syllable error.

RESULTS AND DISCUSSION

Descriptive statistics of the participants' mean scores are reported in Table 2 and depicted in Figure 1. A repeated measures analysis of variance (RM ANOVA) was run on these data with one between-groups factor (participants' L1) and two within-groups factors (phoneme type and language of the stimuli). Each factor had two levels; participants' L1 was either English or Arabic, phoneme type was either consonant or vowel, and language of the stimuli was either English or Arabic. Note that the data do meet the assumptions of the RM ANOVA: an examination of box plots and the results of the Shapiro-Wilk Test for goodness-of-fit indicate that the data are normally distributed, and the results of the Levene's Test indicate that variances are homogeneous. The two-way interaction of Phoneme Type*L1 was statistically significant ($F_{1,37} = 17.61$, $p = .000$, $\eta_p^2 = .32$, power = .98). Thus there was a significant group difference in consonant and vowel recall with a large effect size, suggesting that participants' differential processing of consonants and vowels was dependent on participants' L1. The two-way interaction of Language of Stimuli*L1 was also statistically significant ($F_{1,37} = 13.68$, $p = .001$, $\eta_p^2 = .27$, power = .95), suggesting that participants' ability to recall Arabic-language or English-language sequences accurately also differed depending on their L1. There were statistically significant main effects of phoneme type ($F_{1,37} = 9.71$, $p = .004$, $\eta_p^2 = .21$, power = .86) and language of stimuli ($F_{1,37} = 5.98$, $p = .019$, $\eta_p^2 = .14$, power = .66), but the main effect for L1 was not statistically significant ($F_{1,37} = 2.59$, $p = .116$, $\eta_p^2 = .07$, power = .35). No other interactions were statistically significant.

Recall that the groups were not evenly matched on at least two potentially important factors: age and exposure to L2. The English group's ages ranged from 18 – 35 years (mean = 22.00, SD = 3.52), and they had studied Arabic for an average of 2.54 years (range = 3 months – 6 years, SD = 1.65). The Arabic group's ages ranged from 18 – 52 years (mean = 27.63, SD =

8.49), and they had studied English as an L2 for an average of 12.05 years (range = 4 months – 46 years, SD = 12.03). In order to control for these potentially important group differences, a repeated measures analysis of covariance (RM ANCOVA) was run on these data with one between-groups factor (participants' L1), two within-groups factors (phoneme type and language of the stimuli), and two covariates (age and exposure to L2). The tests of Between-Subjects Effects indicated that age was a statistically significant covariate ($F_{1,35} = 7.61$, $p = .009$, $\eta_p^2 = .18$, power = .76), as was exposure to L2 ($F_{1,35} = 5.99$, $p = .019$, $\eta_p^2 = .15$, power = .66). However, controlling for group differences in age and in L2 exposure by including them as covariates in the RM ANCOVA only strengthened the overall pattern of results found in the RM ANOVA as reported above. Indeed, the two-way interaction of Phoneme Type*L1 was again found to be statistically significant ($F_{1,35} = 29.69$, $p = .000$, $\eta_p^2 = .46$, power = 1.0) and found to have an even greater effect size once age and exposure to L2 were added as covariates, as was the two-way interaction of Language of Stimuli*L1 ($F_{1,35} = 14.67$, $p = .001$, $\eta_p^2 = .30$, power = .96).

As can be seen in Table 2 and Figure 1, native English speakers recalled vowels better than consonants, particularly on the English-language blocks. Paired-samples *t*-tests confirmed that the English group's differences in vowel/consonant scores were statistically significant on both the English-language blocks ($t = 4.68$, $p = .000$, df = 19, 95% CI = 11.51, 4.39) and the Arabic-language blocks ($t = 2.30$, $p = .033$, df = 19, 95% CI = 8.67, 0.40). The English group's superior performance on the English-language blocks (mean = 121.75, SD = 15.12) as opposed to the Arabic-language blocks (mean = 109.95, SD = 13.33) was also statistically significant ($t = 5.69$, $p = .000$, df = 19, 95% CI = 7.46, 16.14). Native speakers of Arabic, on the other hand, demonstrated no statistically significant variation in performance across blocks. The Arabic group's mean scores across vowel/consonant blocks in both languages and across Arabic-language/English-language blocks overall were not statistically different.

At first blush it may appear that the native Arabic speaking group performed worse than the native English speaking group overall. A comparison of the groups' aggregated scores from all four stimuli blocks shows that the English overall mean is 231.70 whereas the Arabic overall mean is 216.26 (out of a possible 288 points). Recall, however, that the RM ANOVA showed no main effect for L1 group. An independent samples *t*-test comparing the overall group scores further confirmed that the difference was not statistically significant ($t = 1.61$, $p = .116$, df = 37). Also apparent from Table 2 is that the Arabic group displayed larger standard deviations on three out of the four stimuli blocks. Given the Arabic group's larger standard deviations and apparently (though not statistically significant) lower overall scores, the data were reexamined to find potential outliers. Indeed, there were only two participants whose overall scores (143 and 152, respectively) fell outside of two standard deviations (SD = 30.67) away from the sample mean (223.98), and both of those participants were in the native Arabic-speaking group. The RM ANOVA as described above was run again on the data after the two potential outliers were removed, but the results were the same and supported exactly the same conclusions that have already been drawn from the whole data set. It was decided not to eliminate these two participants' scores.

Recall that participants' errors were further subdivided into three types: transposition, incorrect syllable, or missing syllable. Descriptive statistics of the participants' error types are reported in Table 3. At first blush, it appears that the L1 groups produced roughly equivalent amounts of error types across the stimuli blocks. For both groups, transposition errors constituted the majority of the errors on all four stimuli blocks. This result is perhaps not surprising given that for any given series there were just three possible syllables that could be used, which made the series relatively predictable. Incorrect and missing syllable errors were much less frequent than transposition errors, across all blocks. Independent-samples *t*-tests were used to compare the numbers of transpositions made by the L1-English and L1-Arabic groups on each of the four stimuli blocks, and the only difference that was statistically significant pertained to the English Vowel block, in which English speakers made fewer transposition errors (mean = 4.45) than Arabic speakers (mean = 12.95) (t = 8.50, p = .000, df = 37, 95% CI = 12.66, 4.34). A series of paired-samples *t*-tests run on each pair wise combination of all four stimuli blocks for both L1-groups determined that while there were *no* statistically significant differences across blocks in terms of transposition errors for the Arabic speakers, *all* comparisons reached statistical significance for the English speakers. That is, English speakers made more transposition errors on the English consonant block compared to the English vowel block (t = 4.67, p = .000, df = 19, 95% CI = 3.70, 9.70), more transposition errors on the Arabic consonant block compared to the Arabic vowel block (t = 2.53, p = .020, df = 19, 95% CI = 0.80, 8.50), fewer transposition errors on the English consonant block compared to the Arabic consonant block (t = -2.52, p = .021,df = 19, 95% CI = -7.97, -0.73), and fewer transposition errors on the English vowel block compared to the Arabic vowel block (t = -5.42, p = .000, df = 19, 95% CI = -8.87, -3.93). Of course, since transposition errors accounted for the majority of errors made on all blocks, for both groups, it is not surprising that there were fewer transposition errors on the blocks with higher accuracy overall. Therefore, the entire series of *t*-tests as described above was run again, this time with transpositions errors as a percentage of the overall errors made on each block (see column 3 on Table 3), rather than raw scores, in order to control for overall block accuracy when comparing error types across blocks. No differences were found to be significant, across groups or within groups and across blocks. Therefore it was concluded that all participants tended to make transposition errors more often than other types of errors, but the relative contribution of transposition errors to overall error rates errors did not vary significantly either between blocks and within L1 groups or between L1 groups.

CONCLUSION

The present study investigated whether all participants would recall sequences of vowels better than sequences of consonants, regardless of their L1 and the language of the stimuli, as prior research would predict. To the contrary, the data clearly demonstrate that there is a group difference, or an interaction between L1 and short term memory for consonants versus vowels. The effect size found ($\eta_p^2 = .32$) for the interaction was large, indicating that approximately one-third of the variance on scores was accounted for by the L1 of the participants. This result suggests that the differential processing of consonants and vowels, at least the level of processing tapped by the immediate serial recall task used, is language-dependent and not universal. To the author's knowledge, the research to date (e.g., Cutler et al., 2000; Nespor et al., 2003; Toro et al., 2008) has indicated that the consonant/vowel processing differences do not vary cross-linguistically. Such research has focused on lexical processing and has found a consonantal lexical processing bias in all languages and populations studied thus far. The current study does not challenge those findings, but rather suggests that there is at least one other level of processing – the STM processing level tapped by immediate serial recall tasks – in which the relative processing of consonants versus vowels is, in fact, language-dependent.

The current study replicated previous findings (Cole, 1973; Crowder, 1971; Drewnowski, 1980) with its native English-speaking participants, who remembered vowels better than consonants in the English stimuli blocks. The current study also extended those findings by testing the participants in their second language, Arabic, and found that, again, as predicted, they remembered vowels better than consonants. However, the native Arabic speaking participants remembered consonants and vowels equally well, in both languages. This result calls into question the suggestion (Crowder, 1971) that vowels are remembered better because they are longer. If the underlying cause of differences in STM performance was differences in phoneme duration, then vowels should have been remembered better by all participants and in all languages, since the vowels were longer than the consonants in all stimuli blocks (see Table 1). An anonymous reviewer suggested that vowels might be better recalled because they appear more frequently in natural language or because they have to be retrieved from a shorter set than consonants. These suggestions, while plausible, are not fully supported by the present data either, as they would not explain the L1 group differences obtained. It may be the case that differences in STM for consonants and vowels can only be sufficiently accounted for by considering numerous factors, some of which are relatively language-independent, such as relative phoneme duration, and some of which are language-dependent, such as the ratio of consonants to vowels and the relative status of the phonemes in written script. Additionally, individual differences other than L1 may have to be considered. In the present data, for instance, age and L2 experience were found to be statistically significant covariates.

More research is needed to isolate all the factors that may account for differences in STM for consonants versus vowels. The current study included only CV syllables because it attempted to replicate previous research, but in future studies, stimuli should also include a variety of syllable types, such as CVC syllables, which would allow inclusion of Arabic short vowels and permit exploration of to what extent vowel duration may be a significant factor. Other languages, including Semitic languages and languages of other language families, should be studied as well. There may be continuum of relative consonant/vowel STM ability that is predictable in large part based on L1. The current study compared only two languages. Future studies should also include native English speakers with more experience in Arabic or another L2. The current data sample was limited in that 6 years was the most L2 experience reported by any native English speaking participant. It would be important to note whether or not highly experienced Arabic L2 learners continued to demonstrate better recall of vowels than consonants in their L1 and L2. Also, measuring L2 experience in terms of number of years studying or using L2 may be an overly coarse rubric. Age of onset of L2 learning might be more influential, though it was not included in the present study design. Another individual difference measure that was not collected here, yet may have further illuminated the results, is a general test of phonological short term memory, such as nonword repetition in L1 and L2. Future studies should attempt to address all of these limitations of the current study.

APPENDIX A

English Consonants
Set 1 (practice):
1. za ka ma ka za ma
2. ka za ma za ma ka

Set 2:
1. mee kee zee zee mee kee
2. mee zee kee zee zee mee
3. kee mee mee zee mee kee
4. zee kee mee kee zee mee

Set 3:
1. koo zoo moo zoo koo koo
2. zoo moo koo koo zoo moo
3. koo zoo moo moo zoo koo
4. moo zoo zoo moo koo zoo

Set 4:
1. ma za ka ma za ma
2. ma ka za za ma ka
3. za ka ma ka za ma
4. ka za ka za ma ka

English Vowels
Set 1 (practice):
1. za zee zoo zoo zee za
2. zee zoo za zoo zee za

Set 2:
1. mee ma moo mee mee ma
2. mee ma ma ma moo mee
3. moo mee ma ma mee moo
4. moo moo ma mee moo ma

Set 3:
1. ka kee koo koo kee ka
2. koo kee ka kee ka koo
3. ka koo kee kee ka ka
4. koo koo ka koo kee ka

Set 4:
1. zee za zoo zee zoo za
2. zoo za zee za za zoo
3. za zoo zee zee za zoo
4. zee zoo zee zoo zee za

APPENDIX B

Arabic Consonants

مجموعة 1

1 ما زا كا ما كا زا
2 كا ما زا ما زا كا

مجموعة 2

1 مو كو زو مو زو كو
2 كو زو كو كو مو زو
3 كو زو مو مو زو مو
4 زو كو مو كو زو مو

مجموعة 3

1 كي مي زي زي كي مي
2 مي زي كي كي زي مي
3 مي مي زي مي زي كي
4 زي زي كي مي كي زي

مجموعة 4

1 كا زا كا كا زا ما
2 كا ما زا زا كا ما
3 ما زا كا زا كا زا
4 كا ما زا ما زا كا

Arabic Vowels

مجموعة 1

1 زا زي زو زو زي زا
2 زا زي زو زا زو زي

مجموعة 2

1 كا كا كو كو كي كا
2 كو كا كي كا كي كو
3 كو كا كا كي كو كا
4 كا كي كو كا كو كي

مجموعة 3

1 ما مي مو مي ما مي
2 مي مو ما مو ما مي
3 مو مي ما مو مي مو
4 ما مو مي ما مي مو

مجموعة 4

1 زا زو زي زو زا زي
2 زو زي زا زو زا زو
3 زي زا زي زي زو زا
4 زا زي زو زا زو زي

ACKNOWLEDGEMENTS

Special thanks to Drs. Robert DeKeyser, Elizabeth Zsiga, Rob Podesva and two anonymous reviewers for their constructive comments on initial drafts of this paper. The author also gratefully thanks Dr. Rusan Chen and Luke Amoroso for their expertise and assistance with statistical analyses and Huda Al-Mufti for her assistance with creation of the Arabic-language stimuli.

REFERENCES

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 1364-1366.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829–839.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In Bower, G. (Ed.), *The psychology of learning and motivation,* Vol. 8. New York: Academic Press.

Berent, I., Perfetti, C.A. (1995). A rose is a REEZ: the two cycles model of phonology assembly in reading English. *Psychological Review, 102*(1), 146-184.

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., & Kennedy, L.J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General, 117*(1), 21-33.

Boatman, D., Hall, C., Goldstein, M. H., Lesser, R. & Gordon, B. (1997). Neuroperceptual differences in consonant and vowel discrimination: as revealed by direct cortical electrical interference. *Cortex, 33,* 83-98.

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science*, *16*(6), 451-459.

Caramazza, A., Chialant, D., Capasso, R., & Miceli, G. (2000). Separable processing of consonants and vowels. *Nature, 403*(6768), 428–430.

Carreiras, M., Duñabeitia, J. A. & Molinaro, N. (2009). Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. *Cerebral Cortex, 19,* 2659 - 2670.

Carreiras, M., Gillon-Dowens, M., Vergara, M. & Perea, M. (2009). Are vowels and consonants processed differently? ERP evidence with a delayed letter paradigm. *Journal of Cognitive Neuroscience, 21*, 275-288.

Carreiras, M., Vergara, M., & Perea, M. (2009). ERP correlates of transposed-letter priming effects: The role of vowels vs. consonants. *Psychophysiology, 46*, 34-42.

Cole, R. A. (1973). Different memory functions for consonants and vowels. *Cognitive Psychology, 4*(1), 39-54.

Crowder, R. (1971). The Sounds of Vowels and Consonants in Immediate Memory. *Journal of Verbal Learning and Verbal Behavior, 10*, 587-596.

Cutler, A., Sebastian-Galles, N., Soler-Vilageliu, O., & Van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory and Cognition*, *28*, 746–755.

Drewnowski, A. (1980). Memory Functions for Vowels and Consonants: A Reinterpretation of Acoustic Similarity Effects. *Journal of Verbal Learning and Verbal Behavior, 19*, 176-193.

Havy, M., Bertoncini, J., & Nazzi, T. (2011). Word learning and phonetic processing in preschool-age children. *Journal of Experimental Child Psychology, 108*, 25-43.

Havy, M., & Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy, 14*, 439–456.

Jusczyk, P. W., & Derrah, C. (1987). A study of phonetic categorization in 2-month-olds. *Developmental Psychology, 23*, 648-654.

Keiko, K. & Zehler, A.M. (2008). *Learning to Read Across Languages: Cross-Linguistic Relationships in First- and Second-Language Literacy Development*. New York: Routledge.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: differences between consonants and vowels. *Cognition, 98*, 13–30.

Nazzi, T. & Bertoncini, J. (2009). A comparison of consonant specificity in onset and coda positions in early lexical acquisition. *Language and Speech, 52,* 463-480.

Nazzi, T., Floccia, C., Moquet, B., & Butler, J. (2009). Bias for consonantal information over vocalic information in 30-month-olds: Cross-linguistic evidence from French and English. *Journal of Experimental Child Psychology, 102,* 522-537.

Nazzi, T. & New, B. (2007). Beyond stop consonants: consonantal specificity in early lexical acquisition. *Cognitive Development*, *22,* 271–279.

Nespor, M., Peña, M. & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio, ii,* 221–247.

New, B., Araujo, V., & Nazzi, T. (2008). Consonants and Vowels in Reading. *Psychological Science*, *19*(12), 1223-1227.

Ryding, K.C. (2005). *A Reference Grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.

Toro, J., Nespor, M., Mehler, J., & Bonatti, L. (2008). Finding Words and Rules in a Speech Stream: Functional Differences Between Vowels and Consonants. *Psychological Science, 19*(2), 137-144.

Van Ooijen, B. (1996). Vowel mutability and lexical selection in English: Evidence from a word reconstruction task. *Memory and Cognition*, *24*, 573–583.

Table 1. Duration (in milliseconds) of syllable stimuli

English

Syllable	Consonant	Duration	Vowel	Duration	Syllable Duration
[mu]	[m]	50	[u]	160	210
[ku]	[k]	70	[u]	160	230
[zu]	[z]	60	[u]	90	150
[mi]	[m]	50	[i]	140	190
[ki]	[k]	60	[i]	160	220
[zi]	[z]	60	[i]	140	200
[ma]	[m]	40	[a]	170	210
[ka]	[k]	80	[a]	120	200
[za]	[z]	60	[a]	160	220
	mean	58.89	mean	144.44	203.33
	SD	11.67	SD	25.55	23.45

Arabic

Syllable	Consonant	Duration	Vowel	Duration	Syllable Duration
[mu]	[m]	80	[u]	180	260
[ku]	[k]	60	[u]	170	230
[zu]	[z]	150	[u]	110	260
[mi]	[m]	70	[i]	150	220
[ki]	[k]	90	[i]	170	260
[zi]	[z]	140	[i]	150	290
[ma]	[m]	70	[a]	310	380
[ka]	[k]	80	[a]	370	450
[za]	[z]	120	[a]	370	490
	mean	95.56	mean	220.00	315.56
	SD	32.83	SD	101.00	99.39

Table 2. Descriptive Statistics by Block and L1 Group

	L1 English (n=20)		L1 Arabic (n=19)	
	Mean	SD	Mean	SD
English Consonants	56.90	9.35	53.84	7.97
English Vowels	64.85	7.46	53.32	11.17
Arabic Consonants	52.70	7.80	55.68	9.20
Arabic Vowels	57.25	8.20	53.42	10.13

Figure 1. Mean Scores by Block and L1 Group

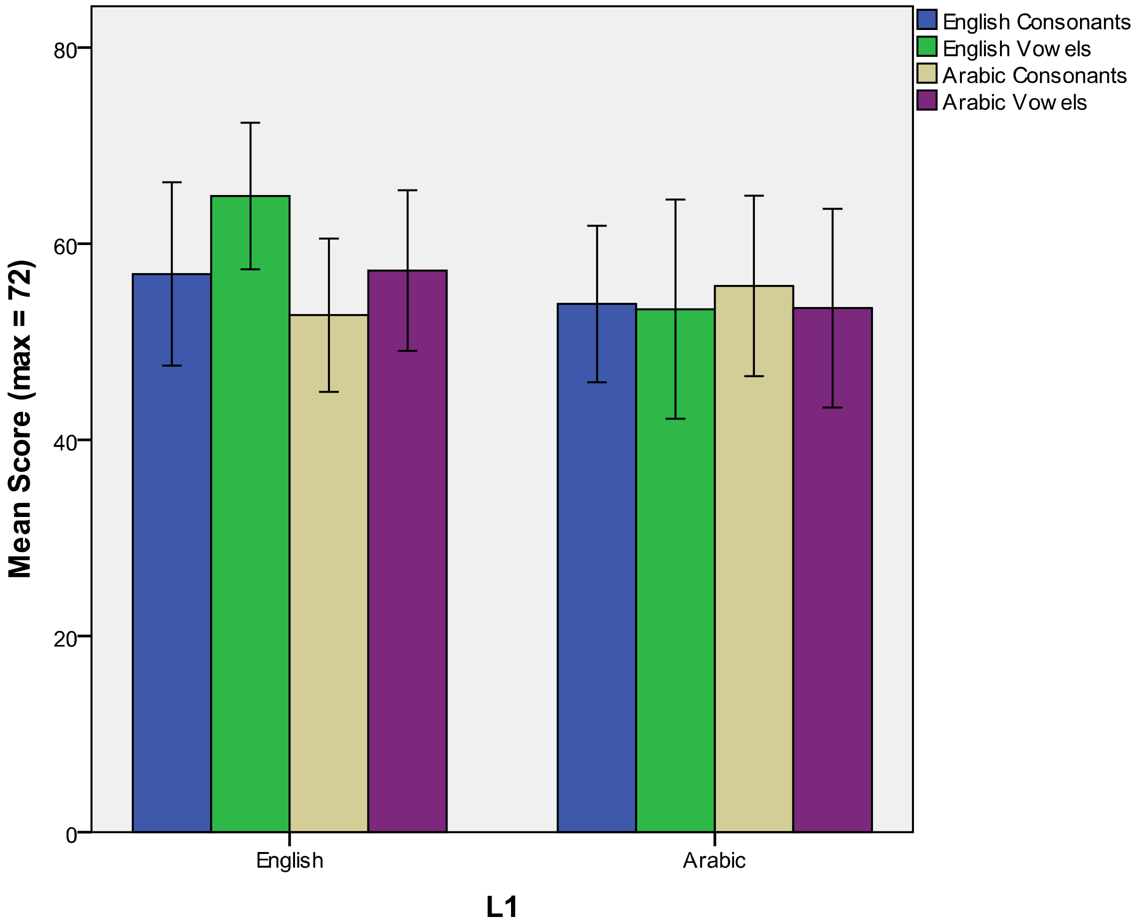

Table 3. Descriptive Statistics of Error Types by Block and L1 Group

	L1 English (n=20)			L1 Arabic (n=19)		
	Mean	SD	% of Block errors	Mean	SD	% of Block errors
English Consonants						
Transposition	11.15	7.32	73.84%	14.58	6.59	80.29%
Incorrect Syllable	2.95	1.99	19.54%	3.26	2.16	17.97%
Missing Syllable	1.00	2.64	6.62%	0.32	0.82	1.74%
English Vowels						
Transposition	4.45	4.65	62.24%	12.95	7.85	69.30%
Incorrect Syllable	1.45	1.54	20.28%	5.11	3.84	27.32%
Missing Syllable	1.25	3.81	17.48%	0.63	2.31	3.38%
Arabic Consonants						
Transposition	15.50	6.81	80.31%	13.00	6.98	79.68%
Incorrect Syllable	2.95	2.21	15.28%	2.63	2.39	16.13%
Missing Syllable	0.85	2.11	4.40%	0.68	1.92	4.19%
Arabic Vowels						
Transposition	10.85	6.04	73.56%	13.84	8.38	74.50%
Incorrect Syllable	2.35	2.13	15.93%	4.37	3.83	23.51%
Missing Syllable	1.55	3.85	10.51%	0.37	1.61	1.98%