PERCEPTION AND PRODUCTION OF GEMINATE CONSONANTS BY ENGLISH SPEAKERS AND KUWAITI ARABIC SPEAKERS

by

NINA RINALDI

A THESIS

Presented to the Department of Linguistics
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Arts

April 2014

An Abstract of the Thesis of

Nina Rinaldi for the degree of Bachelor of Arts
in the Department of Linguistics to be taken June 2014

Title: Perception and Production of Geminate Consonants by English Speakers and Kuwaiti Arabic Speakers

Approved: _______________________________

Professor Melissa A. Redford

By comparing the ability of English speakers and Kuwaiti Arabic speakers to perceive and reproduce a novel length contrast encoded in new words, the present study aimed to determine the implications of phonemic status of consonant length on perception and production. In a condition where long consonants were derived from singletons and paired with their natural counterparts (Natural S condition) English speakers were able to perceive the length contrast encoded in the stimuli some of the time. They did not, however, reproduce a length contrast. Kuwaiti Arabic speakers were able to discriminate the length contrast in the Natural S condition most of the time. In a condition where short consonants were derived from natural geminates and paired with their natural counterparts (Natural G condition), Kuwaiti Arabic speakers were able to discriminate the contrast almost all of the time. Kuwaiti Arabic speakers reproduced a length contrast in both conditions. Furthermore, for the Kuwaiti Arabic speakers, perceptual ability predicted the magnitude of G:S duration ratio reproduced by the speaker.

Table of Contents

List of Figures

List of Tables

I. Introduction

1.1 Geminates and the Phonemic Status of Segmental Length

A *geminate* is a consonant held in closure for an audibly longer period of time than a short, or singleton, consonant. In many languages, use of the short or long version of the same consonant in otherwise identical words can create two different words with completely different meanings. For example, the Arabic word *ħamaam* with a short medial *m* means “pigeons,” while the word *ħammaam*, with a long medial *m*, means “bathroom.” Ham (2001) found the duration of Arabic geminate consonants to range from approximately 2 to 2.8 times the length of a singleton consonant in the language, depending on the geminate’s position within the word. Khattab (2007) found Lebanese Arabic intervocalic geminate consonants were on average 1.82 times the length of singletons in spontaneous speech and 2.5 times the length of singletons when speakers read carefully from word lists. Similar geminate-to-singleton duration ratios were found by Miller (1987) and Al-Tamimi (2004) for Levantine Arabic, and by Hassan (2002) for Iraqi Arabic. The meaningful contrast generated by the length of a sound means that Arabic speakers necessarily attend to this length as they process speech. Because of this, Arabic speakers are said to have a *phonemic length contrast,* meaning that they treat long and short versions of the same sound as distinct categories, rather than as variants of a single category. In Linguistics, two sounds that are separately categorized are known as distinct *phonemes*.

English does not have a phonemic length distinction. To illustrate what it means to lack a phonemic length distinction: I can say, “dog” the normal way or I can draw it out to say, “doooog,” and although I may be communicating something by doing so (e.g., I want you to pay attention to what I am saying), I am not changing the meaning of the word. However, if I say “dig” instead of “dog,” now I have changed the meaning of the word with just one sound. In other words, the vowels *o* and *i* are phonemic in English, but segment length is not. The mental organization of sounds into distinct phonemes and allophones constitutes a concrete, but unconscious, patterning.

To reiterate the essential point regarding the phonemic status of segmental length: depending on their language, speakers either do or do not have separate categories that depend on length. In the following section, I show how the categorical status of length as implications for a listener’s ability to discriminate sound based on length. The current study was designed to investigate effects of categorical status of length on perception and production.

1.2 Prior research: perception of geminates as a function of phonemic discrimination

The bulk of prior work addressing the perceptual sensitivity to a length contrast of speakers with and without a phonemic length category has compared the performances of native Japanese speakers (who have a phonemic length category) and English speakers. These studies generally show that speakers without a phonemic length category display poor sensitivity to the Japanese segmental length contrast (Enomoto 1992; Kato et al. 2001; Kato & Tajima, 2002; Tajima et al. 2003; Hayes-Harb, 2005; Hisagi & Strange, 2011). However, perceptual sensitivity to the contrast increases with Japanese proficiency level (Hayes-Harb, 2005, Hayes-Harb & Masuda, 2008; Hardison & Saigo, 2010). Moreoever, even English speakers with no prior exposure to Japanese are able to discriminate the Japanese length contrast very well when their attention is drawn to the nature of the contrast in brief training sessions (Hisagi & Strange, 2011). Likewise, a study investigating perception of the Finnish length contrast by speakers of L1 Russian with differing amounts of Finnish exposure (Ylinen, et al. 2005) found that the speakers of L2 Finnish with a longer period of exposure to the language made more consistent category judgments than did naïve Russian speakers or those with a shorter period of exposure to Finnish. In fact, the L2 Finnish speakers with a longer period of exposure performed similarly to native speakers in terms of consistency in judgments, but had a different category boundary. Additionally, reaction time for this group of L2 Finnish speakers peaked at judgments on segments whose duration fell near the category boundary.

These studies establish the differential ability of speakers with and without a phonemic length category to discriminate between long and short segments. In the following section, I discuss a study examining English speakers' ability to reproduce a phonemic length contrast. That study also forms the basis for the current study.

1.3 Prior research: production of geminates by speakers without a phonemic length contrast

Although English does not have true geminates, long segments do arise when identical short segments occur in sequence due to affixation (Abercrombie, 1967; Trask, 1996; Kreidler, 2004; Kaye, 2005; Oh & Redford 2012). These segments are known as *fake geminates*, because they are phonetically long but their length does not serve as a phonemic contrast, as does length in a true geminate. For example, the word "unnoticed" is formed by prefixing *un-* to a word beginning with the same sound that the prefix ends on. The two adjacent *n*’s are realized in speech as a single long consonant. In English, long and short versions of the same consonant are allophones.

Redford & Oh (2011) found that although English speakers do not have a phonemic category for length, they are still able to immediately repeat a segmental length contrast encoded in nonsense words. However, they are unable to perceive the same contrast, and their short and long versions of the segments are less different from one another than the two distinct lengths encoded in the stimuli. In their study, participants were presented with each stimulus two separate times, and repeated each word once directly after hearing it. That English speakers consistently reproduced a segmental length contrast without consciously detecting the contrast is in keeping with the status of English fake geminates as allophones rather than distinct consonants.

A major limitation of Redford & Oh’s study is that eliciting only one repetition after each target word confines the scope of investigation to speakers’ imitations of the nonsense words, rather than how the speakers have mentally encoded the words. Elicitation of a single repetition is in line with other work in the area (Goldinger, 1998; Shockley, et al. 2004). For example, Shockley et al. captured English speakers’ imitations of lengthened voice onset time (VOT) in just one repetition of an auditory stimulus. However, if speakers had produced the same words a day later without access to the original stimulus, would they still have repeated the lengthened VOT? Several changes to the structure of this type of repetition experiment could be made in order to investigate the speakers’ mental representation of the new words, as well as their representations of the length contrast encoded in the words. First of all, eliciting multiple repetitions of each stimulus would help elucidate the representation of the target word. This is because the original auditory stimulus recedes in time over the course of the repetitions, so the later repetitions can be taken to signify how the speaker has mentally encoded the word, rather than a purely vocal imitation not drawn from a mental representation. This assumption has its basis in the idea that vocal repetition is not necessarily linked to meaning extraction or even the ability to recall the repeated items: people can repeat words or sentences in a foreign language, for example, without understanding their meaning or recalling them later. Moreover, evidence from patients of aphasia suggests that vocal repetition involves a pathway that "transfers phonological codes…towards areas of motor programming, representation, and execution" (Bernal & Ardila, 2009). Patients of conduction aphasia produce fluent spontaneous speech and can comprehend speech, but have difficulties with vocal repetition (Kohn, 1992; Ardila, 2010). Therefore, the pathway transferring phonological codes to areas of executive function must function as a separate neural subsystem from the ones involved, for example, in the control of vocalization or access to the lexicon.

Second, comparing the performance of English speakers to speakers of a language who have a phonemic length category could highlight how the presence or absence of phonemic status for segmental length leads a speaker to encode, or not encode, a phonemic length contrast in representations of new words. Therefore, the present study expands on Redford and Oh's (2011) work by contrasting the ability of English speakers and Arabic speakers to reproduce a novel length contrast across multiple repetitions of the same stimulus.

1.4 Current study

The current study consists of three experiments. Experiment 1 will investigate Kuwaiti Arabic speakers' production of Kuwaiti Arabic words and English speakers' productions of English words, in order to evaluate the properties of the segmental length contrasts in each language. The Arabic words will contain true geminates and singletons, and the English words will contain fake geminates and singletons. We expect the results of Experiment 1 to confirm our assumption that singletons and true geminates in Kuwaiti Arabic form a contrast of greater magnitude than the contrast formed by singletons and fake geminates in English.

Experiment 2 will investigate the ability of English and Arabic speakers to discriminate a segmental length contrast. We will investigate a segmental length contrast of smaller proportions that the contrast that occurs in Arabic, because we expect that Arabic speakers will be able to easily discriminate a length contrast that mirrors the one in their own language. It has been established previously that English speakers cannot discriminate a segmental length contrast to the proportions of the one in their own language (Redford & Oh, 2011) but conducting a perception experiment with the English speakers is still important to the current study, because it will allow us to compare their performance to that of the Arabic speakers. We predict that the Arabic speakers, but not English speakers, will be able to discriminate the novel contrast.

Experiment 3 will investigate repetitions by English speakers and Arabic speakers using the same stimuli employed in Experiment 2. This experiment investigates how the presence or absence of a phonemic length category leads speakers to differently encode information about length in new words. Multiple repetitions will be elicited following the auditory presentation of each word. We predict that the English speakers will imitate the length contrast in the first several repetitions, but that they will neutralize it by the last several repetitions. This means that if the stimulus word contains a long target consonant, they will at first repeat the consonant to the length encoded in the stimulus, and then shorten it to the length of a singleton. On the other hand, we predict that Arabic speakers will at first repeat the long consonants as they are encoded in the stimuli, and then lengthen them over the course of the repetitions to the length of an Arabic geminate consonant.

II. Experiment 1

2.1 Context: Consonantal length distinctions in Arabic versus English

The hypothesis tested in Experiment 1 is that English and Arabic both have long and short consonants, but that the Arabic length distinction is much larger than in English, due to its status as a phonemic contrast. Experiment 1 took the form of a reading task, where participants were presented with a word list and instructed to read each word out loud three times in a frame sentence. This design was chosen instead of auditory presentation of the words to encourage speakers to access the word as they have mentally encoded it, rather than simply repeating it.

We measured the first vowel (V1), medial consonant (C), and second vowel (V2) for each production of the target word. These measurements allowed us to investigate how the geminates and fake geminates differed from singletons both in terms of *absolute duration* and *relative duration*. Absolute duration refers to the geminate-to-singleton duration ratio, and relative duration refers to the relative timing properties of two segments preceding, following, or including the target consonant, such as the duration ratio of the first vowel to the immediately following target consonant (V1:C) or the duration ratio of the preceding vowel to the following vowel (V1:V2).

Examining both absolute and relative durational properties of the Arabic geminates is important because prior work has shown that both types of properties help to distinguish geminates from singletons, but that the relative importance of the two varies from language to language. Studies examining the segmental length contrast in Arabic suggest that the importance of relative durational criteria to the geminate-singleton distinction is subject to cross-dialectal variation. Khattab (2007) found considerable overlap in the absolute duration of Lebanese Arabic geminate and singleton consonants in spontaneous speech, but that the V1:C duration ratio remained significantly different as to distinguish geminates from singletons. Specifically, the V1:C duration ratio is shorter for geminates than singletons, meaning that a vowel preceding a geminate is shorter in relation to the following consonant than a vowel preceding a singleton. Khattab elicited data in two speech conditions: reading from a word list and spontaneous speech. She found that although the absolute durations of V1, C, and V2 were much larger in the word list condition, the relative timing measure of V1:C were was consistent between speech styles as a measure of consonant length category. Al-Tamimi (2004) examined geminates and singletons in Jordanian Arabic following both short and long vowels, and his results similarly support the notion that relative duration is crucial to the geminate-singleton category boundary in Jordanian. Al-Tamimi found a strong correlation between the presence of a geminate or singleton and the length of the preceding consonant, where geminates were preceded by shorter vowels than singletons.

Studies of Iraqi Arabic, on the other hand, suggest that relative timing criteria may not be important to the phonemic length contrast of that dialect. In a study of Iraqi Arabic disyllabic words with first-syllable stress, Hassan (2002) found phonologically short vowels to be shorter when preceding a geminate consonant and longer when preceding a singleton. He ultimately argued, however, that the shorter vowel duration preceding geminates were inconsequential in serving to differentiate the length contrast. After calculating the average durations of pre-singleton vowels and pre-geminate vowels, and singleton consonants and geminate consonants, Hassan found that the durational differences between pre-singleton and pre-geminate vowels fell around the level of "barely noticeable" whereas the durational differences between geminates and singletons were very salient. It would therefore appear that relative durational criteria play a more important role in the segmental length distinction for Lebanese and Jordanian Arabic than for Iraqi Arabic. The experiments presented in this thesis may help to establish the relative importance of absolute and relative timing properties to the Kuwaiti Arabic segmental length distinction.

2.2 Methods

2.2.1 Participants

Participants were a group of 9 native speakers of American English (4 male; 5 female) between the ages of 20 and 25, and a group of 9 native speakers of Kuwaiti Arabic (6 male; 3 female) between the ages of 18 and 21.

2.2.2 Stimuli

For the Arabic stimuli, eighteen Kuwaiti Arabic words were selected which followed the template CVCV(C) or CVCCV(C) (see Table 1). All words were disyllabic and trochaically stressed1. In all cases, the stressed syllable directly preceded the target medial consonant. Ten words contained the medial consonant *m* and eight words contained the medial consonant *n*. Nine words contained a singleton medial consonant and nine contained a geminate medial consonant. The stimuli were

[1] *Trochaic stress* refers to the stress pattern of a stressed syllable preceding an unstressed syllable. The opposite pattern is *iambic stress*, which refers to an unstressed syllable preceding a stressed syllable.

counterbalanced so that the list contained five words with the geminate *m* and singleton *m*; and four words with geminate *n* and singleton *n*. The words were selected to be minimal pairs or near-minimal pairs, when possible differing only in the length of the medial consonant. Words with first-syllable stress were chosen because segment length was not found to be contrastive in words with second-syllable stress for two of three possible phonemically long vowels. It was established during consultation with a native speaker of Kuwaiti Arabic that two of the three phonemically long vowels cannot immediately follow both a geminate and singleton, so minimal pairs on the basis of stress could only be created for one of those three vowels. Because length subject to phonemic contrast in Arabic, as well as serving as a cue to stress, stress must occur on the second syllable of words with a phonemically long second vowel.

Geminates	**Singletons**
M [ʔammə] عمّة *'maternal aunt'* [ħammɛl] حمّل *'make (s.o.) carry'* [djɛmmɛl] جمّل *'make (s.th.) beautiful'* [sɛmmaʔ] سمّ *'make (s.o.) listen'* [lɛmmə] لمّا *'when'*	**M** [ʔamɛj] عمي *'my two uncles'* [ħamɛl] حمل *'carry'* [zɪmɛn] زمن *'time'* [sɛmaʔ] سمع *'listen'* [gumar] قمر *'moon'*
N [ħɛnnə] حنّة *'henna'* [ɣɛnnə] غنّى *'he sang'* [djɛnnə] جنّة *'paradise'* [sɪnnə] سنّة *'path'*	**N** [honə] هنا *'here'* [ɣɛni] غني *'rich'* [mɪnə] منى *name of the tent city for pilgrims in Mecca, Saudi Arabia* [sɪnə] سنة *'year'*

Table 1: Kuwaiti Arabic stimuli, Experiment 1

The same sixteen English words tested by Oh & Redford (2012) were employed as the English stimuli in this experiment (see Table 2). The words were matched as pairs with the same stress where one word contained a long target consonant and the other contained a short target consonant, as established by Redford & Oh (2012). Eight words contained the medial consonant *m* and eight words contained the medial consonant *n*. Likewise, eight words contained a long medial consonant and eight words contained a short medial consonant. The word list was counterbalanced so that half of the *m* words were long and half were short, and likewise for the words containing *n*. The long and short words were orthographically matched so that in all words a double *n* or *m* represented the target consonant.

Fake Geminates		**Singletons**	
M		**M**	
[ɪmuvəbəl]	immoveable	[əmoʊnyə]	ammonia
[ɪmɔrəl]	immoral	[ɪmɛnsli]	immensely
[ɪməmɔɹiəl]	immemorial	[ɪmyunɪti]	immunity
[ɪmmɛʒɚd]	immeasured	[ɪmɪgɹeɪʃənəl]	immigrational
N		**N**	
[ʌnnoʊtɪst]	unnoticed	[ənɛks]	annex
[ʌnneɪmd]	unnamed	[ɪneɪt]	innate
[ʌnnɝv]	unnerve	[ənɔɪd]	annoyed
[ʌnneɪl]	unnail	[ɪnɝv]	innerve

Table 2: English stimuli, Experiment 1

2.3 Procedure

The word list was randomized for each participant and printed on a piece of paper. Each Arabic-speaking participant was instructed to read each word three times in the Kuwaiti Arabic sentence, *"I said ______ again" ("Gilt ______ marra thania").* Each English-speaking participant was instructed to read each word three times using the frame sentences, *"I said ____ one time. I said _____ two times. I said ____ three times."* Participants' speech was digitally recorded using a microphone clipped to a hat that participants were given to wear.

2.4 Measurements and Analyses

Target words were displayed in the phonetics software Praat (version 5.3.32) as oscillograms (waveforms) and spectrograms (an image of acoustic energy distributed over time) (see Figure 2). Using Praat, durations were measured for all of the target medial consonant as well as the vowels immediately preceding and following it (V1 and V2). All three repetitions were measured for each word. Consonant-V1 boundaries were generally identified by the beginning of strong acoustic energy for all formants. Nasal-vowel boundaries were identified by a slight diminishment of acoustic energy at the first *formant* [2] for the duration of the nasal, as well as a more severe diminishment of acoustic energy at the second formant. Figure 1 shows an example of a sentence displayed in Praat, with the segment boundaries delineated within the target word.

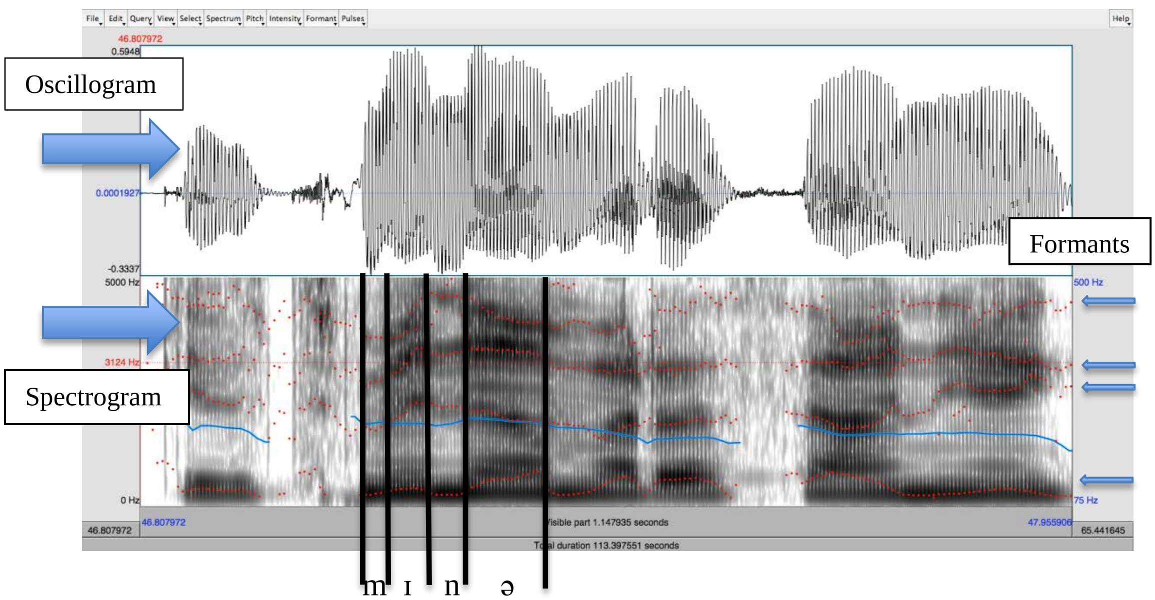

Figure 1: Oscillogram and Spectrogram displayed in Praat.

Sentence displayed: “Gilt [mɪnə] marra thania.” (‘I said [mɪnə] again.’)

2.5 Results

Overall, both English speakers and Arabic speakers consistently produced distinct short and long consonants. The results confirmed our hypothesis that Arabic

[2] Formant= a resonant frequency of the vocal tract. Formant 1 is the formant indicated on the very bottom of the spectrogram, and formant 2 is directly above it.

speakers' productions of long and short consonants would be more different from one another in both absolute and relative timing measures than those of English speakers. Significant effects were found for language, segment, and target length for the absolute duration measures of G:S duration as well as the raw consonant durations; and the relative duration measure of V1:C.

As expected, the geminate-singleton duration ratio is greater for both consonants in Kuwaiti Arabic than in English (Table 3). The geminate-singleton duration ratio is greater for *n* than *m* in both languages. In terms of raw segment durations, Kuwaiti Arabic *m* has a longer mean duration than *n* for both long and short targets. This situation contrasts with English, where *m* and *n* as short targets have almost the same mean duration, but *n* is of longer duration than *m* for long targets.

The results of the relative timing measure V1:C show that in Arabic, long *m* and long *n* are almost the same length, but that short *n* is much shorter than short *m*. We find the same pattern in English, where long *m* and long *n* are almost the same length, but short *n* is much shorter than short *m*. Overall, V1:C proved a strong correlate to length category in Kuwaiti Arabic.

	Arabic	English
M	1.81 *(.57)*	1.17 *(.48)*
N	2.48 *(1.12)*	1.59 *(1.11)*

Table 3: Experiment 1, G:S duration ratios

(Standard deviation in parentheses)

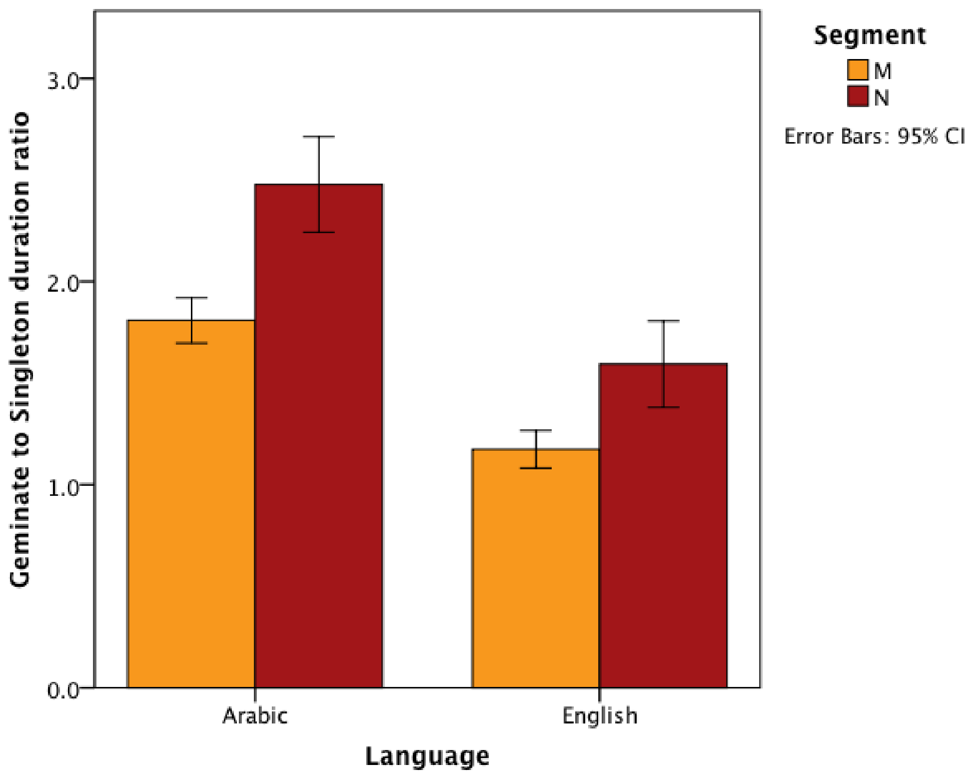

Figure 2: Experiment 1, G:S duration ratios

	Arabic		English	
	Long	Short	Long	Short
M	120.23 *(35.19)*	69.27 *(18.25)*	95.70 *(25.74)*	87.98 *(24.57)*
N	111.24 *(31.13)*	54.82 *(33.05)*	112.63 *(32.54)*	84.28 *(31.30)*

Table 4: Experiment 1, mean raw duration by segment (msec)

(Standard deviation in parentheses)

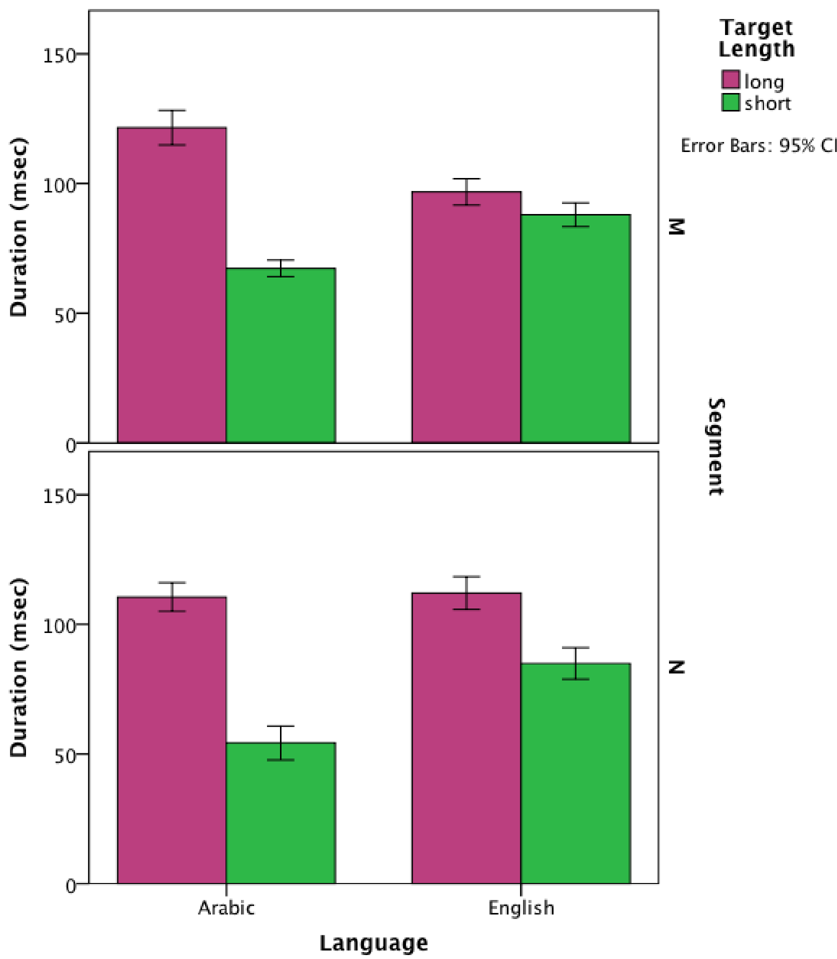

Figure 3: Experiment 1, mean raw duration by segment

	Arabic		English	
	Long	Short	Long	Short
M	.62 *(.31)*	1.01 *(.41)*	.97 *(.87)*	.93 *(.73)*
N	.59 *(.31)*	1.52 *(.84)*	.98 *(.52)*	1.67 *(1.70)*

Table 5: Experiment 1, V1:C duration ratios

(Standard deviation in parentheses)

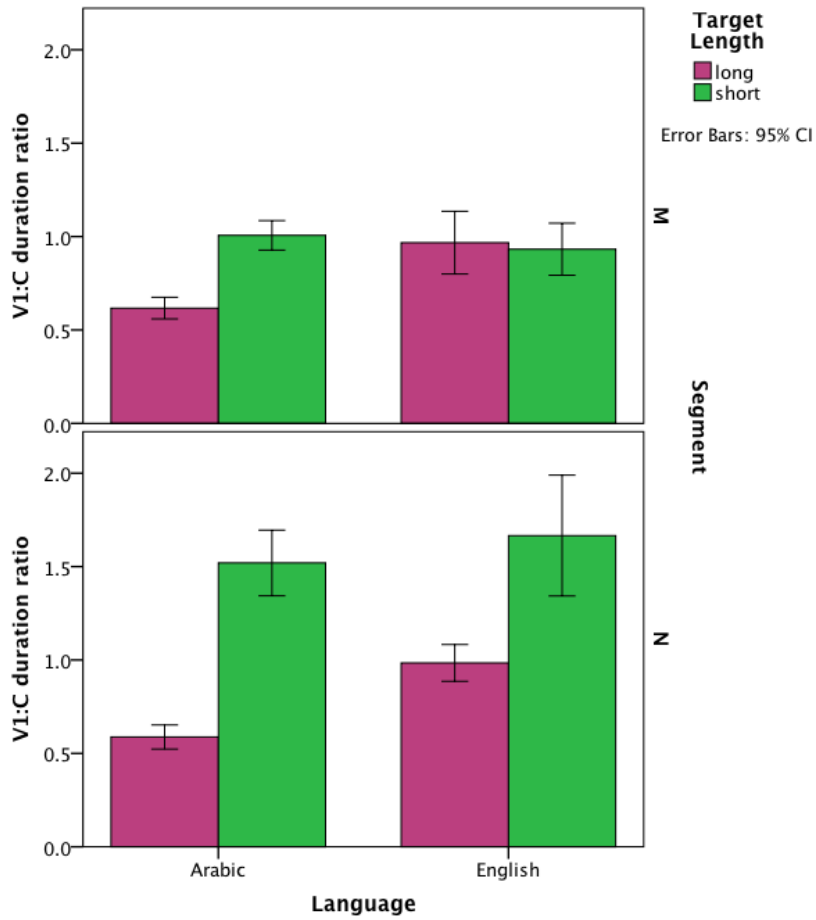

Figure 4: Experiment 1, V1:C duration ratios

2.6 Discussion

If production is taken to reflect representation, then the findings show that both Arabic speakers and English speakers have encoded a length distinction in their representations of the target words. The greater durational differences between long and short targets for Arabic speakers than for English speakers confirm that the encoded length distinction is of greater magnitude in Arabic than in English, presumably because the length contrast in Arabic is phonemic.

The surprising part of these results is how the absolute timing measures align with the relative timing measure V1:C. In Arabic, *m* is of longer average raw duration than *n*. However, *m* also has a smaller G:S ratio than *n*, and long and short versions of *m* are less different from one another than *n* in terms of the measure V1:C. The findings on the absolute timing properties of *m* compared to *n* in Arabic correspond somewhat to those of Khattab (2007) who found the average duration of *m* to be longer than that of *n* only in singleton contexts. The two segments were of similar durations in geminate contexts. Khattab did, however, find that *m* had a greater G:S ratio than *n* (1:2.6 versus 1:2.97 respectively), which contrasts with the results of the current study. In English, short *m* and *n* are of about the same average raw duration, whereas the average raw duration of long *n* is much longer than the average raw duration of *m*. Like in Arabic, the G:S ratio is greater for *n* than *m*. But in terms of V1:C, English long and short *m* are similar durations, while a much greater difference is found between long and short *n*. The absolute duration results align with those of Oh & Redford (2012), who found long *n* to have longer raw duration than long *m*. They attributed the difference to greater decomposability of words prefixed with *un*- compared to *in*- (realized as *im*- in words such as *immeasured*), since segmental duration at a morpheme boundary varies as a function of the word's semantic transparency and productivity (Hay, 2007). They posited that the morpheme boundary in *un*- prefixed words was sometimes, but not always, represented in the speech plan.

With these absolute duration results in mind, it is interesting that for the measure V1:C in both languages, the greater difference between words containing long and short *n* rather than *m* is due to the much shorter relative duration for the short *n.* The long *n* in English is longer than long *m* in absolute terms but nearly the same length in relative terms. That the short *n* in both languages is much shorter than short *m* in terms of the V1:C duration suggests that this relative measure is intrinsic to *n* as a consonant in both languages, perhaps due to physiological differences between the articulation of the two sounds. Oh & Redford could still be correct in proposing that decomposability underpins the difference in absolute duration between segments, but perhaps English *m* and *n* have distinct relative timing properties from one another in both their short and long variations.

In any case, the V1:C duration ratio did prove a reliable signal to length category in Arabic, as it did in English for the segment *n.* We can therefore add Kuwaiti Arabic to the list of languages for which this relative timing measure serves as a durational covariant to segmental length category. The status of V1:C as a correlate to the English segmental length distinction is not so certain, since this measure worked well to differentiate long and short *n* but not *m.*

A caveat to the Arabic results of Experiment 1 is that only one word pattern was tested, where both V1 and V2 were phonemically short vowels. Since Arabic geminates can occur before or after both phonemically short and long vowels, vowel length signaling gemination must interact with phonemic vowel length. Khattab (2007) measured geminates in Lebanese Arabic occurring after both long and short vowels. Her findings indicate that a long vowel immediately preceding a geminate consonant is on average very close in duration to the geminate consonant, whereas a long vowel preceding a singleton consonant is much longer than the singleton. She also found words containing a long vowel-geminate sequence (but not the long vowel-singleton sequence), though attested in Lebanese Arabic, to be extremely rare in spontaneous speech, so it is possible that the word-initial sequence VVCC is perceptually difficult for speakers, who are in the process of reanalyzing it into an alternative timing pattern. Therefore, further testing on different word patterns is required before we are able to conclusively generalize the function of V1:C as serving as an acoustic cue to gemination in Kuwaiti.

Having established that both Arabic and English speakers have encoded a length distinction, we turn to the question of how the different representations of a length distinction, namely phonemic and non-phonemic, correlate with speakers' ability to discriminate a novel length contrast.

III. Experiment 2

3.1 Background: Perception of a novel length contrast

In this experiment, we test the Arabic speakers' and English speakers' ability to perceive the length contrast encoded in nonsense words. Two sets of stimuli were created for Arabic, and one set of stimuli was created for English. In all stimuli sets, the long target consonants were 1.6 times longer than the short targets. This contrast is of lesser magnitude than the phonemic length contrast established in Arabic (2.15), and of greater magnitude than the non-phonemic contrast in English (1.38) as established in Experiment 1. In Arabic, we derived one set of stimuli from singleton productions of the nonwords (Natural S condition) and one set from geminate productions of the nonwords (Natural G condition). We expect the English speakers will be unable to perceive the length contrast. We expect the Arabic speakers to perceive the length contrasts in both conditions.

3.2 Methods

3.2.1 Participants

The same participants from Experiments 1 participated in Experiment 2.

3.2.2 Stimuli

Twenty trochaically-stressed disyllabic Kuwaiti Arabic nonwords were created on the pattern CVCVC. The nonwords were designed to sound like plausible Kuwaiti words. Half contained the medial consonant *n* and half contained the medial consonant *m*. An adult male native speaker of Kuwaiti Arabic recorded each of these words once in the Kuwaiti Arabic sentence, “This is an X” and once in the sentence “Can you say X?” thus producing forty sentences total, including two presentation contexts per word. The speaker then followed the same procedure again but this time produced each nonword with a geminate medial consonant. In order to elicit a geminated target consonant, the speaker was instructed to produce the medial consonant with *tashdeed,* the Arabic grammatical term referring to gemination (lit.: *‘strengthening’*). The singleton word list was then copied, and the duration of the medial consonant of each nonword was measured. The medial consonant of each nonword was then lengthened by 1.6 times its duration. The original words containing natural singleton medial consonants together with their derived geminate counterparts comprised the Natural S condition. The geminate word list was then copied, and the duration of the medial consonant of each nonword was measured. The medial consonant of each nonword was then shortened by 0.6 of its duration. The original words containing the true geminate medial consonants were grouped together with their derived singleton counterparts.

A nearly identical procedure was followed in order to create the English stimuli. Twenty trochaically-stressed disyllabic English nonwords were created on the pattern CVCVC. Half contained the medial consonant *n* and half contained the medial consonant *m*. The other consonants in the words were selected so that the English nonwords matched the Kuwaiti nonwords as closely as possible. However, Kuwaiti Arabic consonants not found in English were replaced with English consonants. An adult male native speaker of American English recorded half of these words once in the sentence, “This is an X” and once in the sentence “Can you say X?” He recorded the other half of the words once in the sentence “That is a X” and once in the sentence “Can you say X?" The word list was then copied and the duration of the medial consonant of each nonword was measured. The medial consonant of each nonword was then lengthened by 1.6 times its duration. As with the Arabic stimuli, the original nonwords containing natural singleton medial consonants together with their derived counterparts make up the Natural S condition, the only condition English speakers were tested in. The English speakers were only tested in the natural S condition because there are no natural geminates in English from which shorter segments could have been derived in order to create a Natural G condition for English. While English speakers do differentiate between short and long consonants in production, there is no way to prompt English speakers to produce a longer target consonant in novel words as there is in Arabic.

Table 6: Experiment 2, Arabic stimuli

Long	**Short**
M	**M**
lɛmmɛtʃ	lɛmmɛtʃ
yɪmmʌx	yɪmʌx
ʔɛmmɛg	ʔɛmɛg
tɛmmɛd	tɛmɛd
kɛmmɛd	kɛmɛd
dɪmmes	dɪmes
zɛmmɛh	zɛmɛh
tʃɪmmʌθ	tʃɪmʌθ
gɛmmɛs	gɛmɛs
zɪmmʌr	zɪmʌr
N	**N**
tʃɪnnʌf	tʃɪnʌf
ɣɛnnɛð	ɣɛnɛð
gɛnnɛr	gɛnɛr
hɛnnɛr	hɛnɛr
lɛnnɛt	lɛnɛt
ʃɛnnɛl	ʃɛnɛl
thɪnnʌg	thɪnʌg
xɪnnʌm	xɪnʌm
tɪnnæʔ	tɪnæʔ
zɪnnʌf	zɪnʌf

Table 7: Experiment 2, English stimuli

Long	**Short**
M **Long** tʃɪmməθ dɪmməs famməz gamməs lammətʃ pɪmməg tamməd yɪmməl zamməf	**M** **Short** tʃɪməθ dɪməs faməz gaməs lamətʃ pɪməg taməd yɪməl zaməf
N tʃɪnnəf gannɚ hannɚ hannəb kannəd lannət lɪnnəm ʃannʌl thɪnnəg zɪnnəf	**N** tʃɪnəf ganɚ hanɚ hanəb kanəd lanət lɪnəm ʃanʌl thɪnəg zɪnəf

Arabic stimuli consisted of all target words from the first presentation (*"This is a X"*) excised from their carrier sentences. Each unaltered word was paired with its derived geminate or derived singleton counterpart, or repeated. This way, each pair consisted of either a minimal pair or two copies of the same word. The minimal pairs (but not the matching pairs) were repeated so the words in the pair could be displayed in two different orders. Thus, the Arabic stimuli consisted of 160 different pairings. Pairs from Natural G and Natural S conditions were mixed in together and both conditions were administered as a single task. The pair ordering was randomized for every participant.

The same procedure was followed to create the English stimuli. However, because the English speakers were tested on only the natural S condition, the English stimuli consisted of 80 different pairings.

3.3 Procedure

The participants completed the perception task at a computer, which played one word pair at a time. As the word pair played, a computer screen displayed two buttons reading “same” and “different.” The participants were instructed to make a judgment on each word pair as to whether the words sounded exactly the same or slightly different.

3.4 Measurements & Analyses

Each correct judgment was coded “1” and each incorrect answer was coded as “0”. The judgments were analyzed in the framework of signal detection theory, a model for representing a person’s ability to correctly identify whether a signal is present or absent in a given system across a number of trials. In the context of Experiment 3, where participants were asked to classify two words as “same” or “different,” a response for a given pairing falls into one of four categories: *hit* (correctly judging “different”), *miss* (judging “same” when the two words are actually the different), *false alarm* (judging “different” when the two words are the same), and *correct rejection* (correctly judging “same”).

A discriminability index, d', is calculated for each participant’s internal response to the signal. The best participant will maximize the hit rate and minimize the false alarm rate, thus minimizing the miss rate and maximizing correct rejection. Therefore, we only need the hit rate and false alarm rate to calculate d'. Misses and correct rejections are redundant because they can be obtained by subtracting the hit rate or false alarm rate from 1. Generally, *d* ' is calculated as the difference between the z-transforms of the hit rate and false alarm rate: d '= z(H) - z(F). The larger the absolute value of *d* ', the more sensitive a person is to the difference between the signal present and signal absent distributions. A value of 1 or above indicates that the participant successfully detected the signal (discriminated the contrast), while a value of zero indicates a chance performance. We used an online *d* ' calculator to find the *d* ' values for each subject (http://memory.psych.mun.ca/models/dprime/). The *d* ' values of several participants were changed to zero after it was established via Fisher test that proportions were not significantly different between the hit rate and false alarm rate.

3.5 Results

Group and individual results are displayed in Tables 11- 13. Overall, the English speakers had the most difficulty correctly identifying the word pairs as same or different, although five of the eight participants were able to discriminate the contrast. The Arabic speakers in the Natural S condition performed more accurately than the English speakers did in the same condition, their performance was still less accurate than in the Natural G condition. While both the English participants and Arabic participants in the Natural S condition were more likely to err on the side of judging two words as “different” than “same,” they were still more conservative about making “different” judgments compared to Arabic participants in the Natural G condition.

Participant	Different_T (Hits)	Same_T (False Alarms)	d'
A06	43.75%	17.5%	0.777
A02	58.75%	3.75%	2.002
A04	81.25%	8.75%	2.243
A07	85.00%	5.00%	2.681
A05	97.50%	15.00%	2.996
A08	90.00%	2.50%	3.242
A03	97.56%	7.69%	3.397
A01	34.57%	0.00%	3.903
A09	71.25%	0.00%	4.861
Group statistics	**Mean: 73.29%** **Std Dev: (23.03%)**	**6.69%** **(6.22%)**	**2.90** **(1.17)**

Table 8: Experiment 2, Arabic Natural G discriminability

Participant	Different_T (Hits)	Same_T (False Alarms)	d'
A03	23.08%	1.22%	0.429
A06	41.25%	23.75%	0.493
A02	27.50%	10.00%	0.684
A05	36.25%	7.50%	1.088
A07	46.25%	11.25%	1.119
A04	38.75%	3.75%	1.495
A08	28.75%	1.25%	1.681
A09	11.25%	0.00%	3.087
A01	20.25%	0.00%	3.467
Group statistics	**Mean: 30.37%** **Std Dev: (11.22%)**	**6.52%** **(7.75%)**	**1.50** **(1.09)**

Table 9: Experiment 2, Arabic Natural S discriminability

Participant	Different_T (Hits)	Same_T (False Alarms)	d'
E04	0.00%	0.00%	0
E09	2.50%	0.00%	0
E08	25.00%	21.25%	0
E02	51.25%	23.75%	0.746
E01	27.50%	6.25%	0.936
E03	16.25%	2.50%	0.976
E05	50.00%	12.50%	1.15
E06	36.25%	2.50%	1.608
Group statistics	**Mean: 25.44%** **Std Dev: (18.29%)**	**7.64%** **(9.34%)**	**0.68** **0.61**

Table 10: Experiment 2, English Natural S discriminability

3.6 Discussion

The results of Experiment 2 contradicted the hypotheses that English speakers would not be able to discriminate the length contrast. The results confirmed the hypothesis that the Arabic speakers would be able to discriminate the length contrast in both conditions. However, the Arabic speakers were not expected to perform significantly better in one condition compared to the other, so their much more accurate performance in the Natural G versus Natural S condition was unexpected. It appears from the Arabic results that participants perceived many of the long segments in the Natural S condition as singletons rather than geminates.

The individual results for each condition shed some light on the group statistics. In the English natural S condition, three participants have the d' value 0. A d' value of 0 where the hit rate and false alarm rate are both zero (such as for participant E04) indicates that all judgments were either *miss* (choosing "same" when two words are different) or *correct rejection* (correctly identifying two words as the "same"). Therefore, a smaller hit rate paired with a smaller false alarm rate reflects conservatism about judging a word pair as *different*. A hit rate and false alarm rate that are both larger and of very similar proportions to one another (such as for E08), reflects a less conservative judge, but one who is still not sensitive to the contrast.

The Arabic natural S individual results show that there were several Arabic speakers who performed substantially better than the English speakers in this condition. In fact, the group d' for the Arabic natural S condition (1.50) is over twice as large as the group d' value for the English Natural S condition (0.68). Participants A01 and

A09 performed much more accurately than the rest of the group, while participants A03, A06, and A02 were the least accurate.

The Arabic natural G individual results show that Arabic speakers were more sensitive to the length difference in this condition compared to the Arabic Natural S condition, with a group d' (2.90) almost twice the group d' (1.50) of the Arabic natural S condition. Again, the most accurate judges in this condition were the participants A01 & A09, and the least accurate judge was A06. In general, however, almost everyone made more accurate judgments than for the Natural S condition. Having established that some English speakers can discriminate the length contrast; that Arabic speakers can mostly discriminate the contrast in the Natural S condition; and that Arabic speakers can fully discriminate the contrast in the Natural G condition, we will now examine participants' reproductions of the same stimuli.

IV. Experiment 3

4.1 Background

In Experiment 3 we investigate how the different groups of speakers encode a novel length contrast over a number of repetitions of the same stimulus. To do this, we presented the target words auditorily and elicited ten repetitions following each presentation. Based on the results of Redford & Oh (2011) we predict that English speakers will reproduce a length contrast immediately. However, we also predict that the contrast will weaken with repetition. As for the Arabic speakers, we predict that they will perform similarly to the English speakers in the Natural S condition, since the results of Experiment 2 show that they had difficulty discriminating the contrast in this condition. Regarding the Natural G condition, we predict that the Arabic speakers will not only reproduce the length distinction immediately, but to also strengthen it over number of repetitions.

4.2 Methods

4.2.1 Participants

The same participants from Experiments 1 and 2 all completed Experiment 3.

4.2.2 Stimuli

The same stimuli used in Experiment 2 were also employed here. The following procedure was carried out for the Arabic Natural S nonwords; the English Natural S nonwords; and the Arabic Natural G nonwords: the singleton or geminate words and their derived counterparts were grouped together and divided into two randomly-ordered lists (List A and List B) of twenty words each. Each list contained 10 short words and 10 derived-long words; or 10 long words and 10 derived-short words. Likewise, each list contained 10 words with the medial consonant *n* and 10 words with the medial consonant *m*. The lists were counterbalanced so that the words which appeared short on List A appeared as short-medium on List B, and vice versa. The nonwords in each list occurred in a randomized order.

4.3 Procedure

The production task was structured as a word-learning task, where participants were instructed to learn the names of Pokemon creatures whose images were presented on index cards. Participants were presented with two cards at once, and the recorded sentences were played for the participants in the following order:

<u>Arabic</u>

“hatha X” (‘This is a X’) [speaker produces 10 repetitions]
“hatha Y” (‘This is a Y’) [10 repetitions]
“tegdar etgol X?” (‘Can you say X?’) [10 repetitions]
“tegdar etgol Y?” (Can you say Y?’)[10 repetitions]

<u>English</u>

“This is a X” [speaker produces 10 repetitions]
“This is a Y” [10 repetitions]
“Can you say X?” [10 repetitions]
“Can you say Y?” [10 repetitions]

Following each presentation, the participant repeated the target word ten times. Participants’ speech was digitally recorded using a small microphone clipped to their clothes or a hat that they were given to wear.

4.4 Measurements and Analyses

Target words were displayed in Praat as oscillograms and spectrograms. Using Praat, durations were measured of the target medial consonant as well as the vowels immediately preceding and following it (V1 and V2). Measurements were taken of the first, second, third, seventh, eighth, ninth, and tenth repetitions of each word. The same consonant-vowel boundary criteria were used as in Experiment 1. However, in cases where the initial consonant was a voiceless stop such as *t*, the C-V1 boundary was identified as immediately following the release of the stop. The purpose of this exception was to delineate a V1 for some Arabic words which were reproduced with no formant activity preceding the target consonant, indicating a vowel shortened to the extreme.

4.5 Results

Contrary to the hypothesis, no effect of repetition was found for any of the measures. Overall, the Arabic speakers reproduced a length contrast and the English speakers did not.

Results on the G:S duration ratios are presented in Table 8. Significant effects were found for language by stimulus type, and segment. A two- way interaction was also found between language by stimulus type and segment.

Clear differences are also present in the raw segment durations between the English and Arabic productions from the natural S condition (Table 9). Significant effects were found for language by stimulus type, segment, and target length. Two-way interactions were found for language by stimulus type and target length, as well as language by stimulus type and segment.

The results on V1:C duration ratios are presented in Table 10. Significant effects were found for language by stimulus type, segment, and target length. A two-way interaction was found between language by stimulus type and target length, as well as language by stimulus type and segment.

	Arabic, Natural G	Arabic, Natural S	English, Natural S
M	1.35 *(0.28)*	1.50 *(0.29)*	1.04 *(0.18)*
N	1.30 *(0.28)*	1.14 *(0.19)*	1.03 *(0.09)*

Table 11: Experiment 3, G:S duration ratios by stimulus type

(Standard deviation in parentheses)

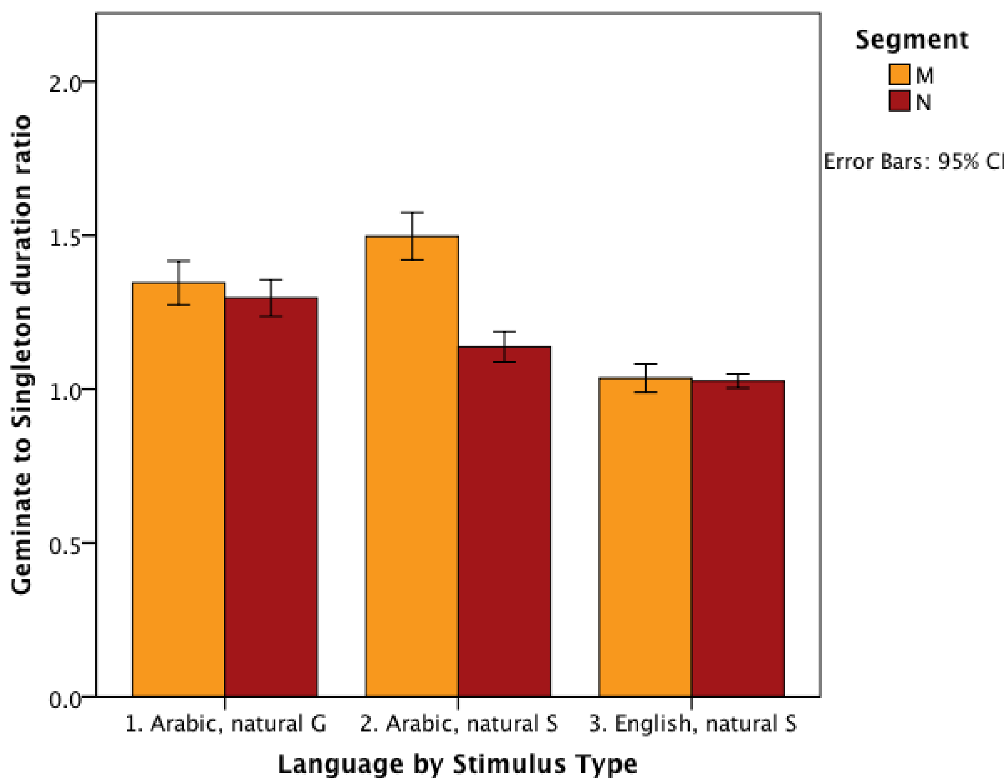

Figure 5: Experiment 3, G:S duration ratio by stimulus type

	Arabic, Natural G		Arabic, Natural S		English, Natural S	
	Long Target	Short Target	Long Target	Short Target	Long Target	Short Target
M	156.56 *(29.10)*	118.35 *(28.54)*	121.89 *(30.28)*	79.01 *(7.67)*	72.61 *(7.82)*	71.38 *(10.43)*
N	147.88 *(32.22)*	115.85 *(31.60)*	60.85 *(11.50)*	52.86 *(6.25)*	53.28 *(7.91)*	52.14 *(7.42)*

Table 12: Experiment 3, mean raw segment duration by stimulus type (msec)

(Standard deviation in parentheses)

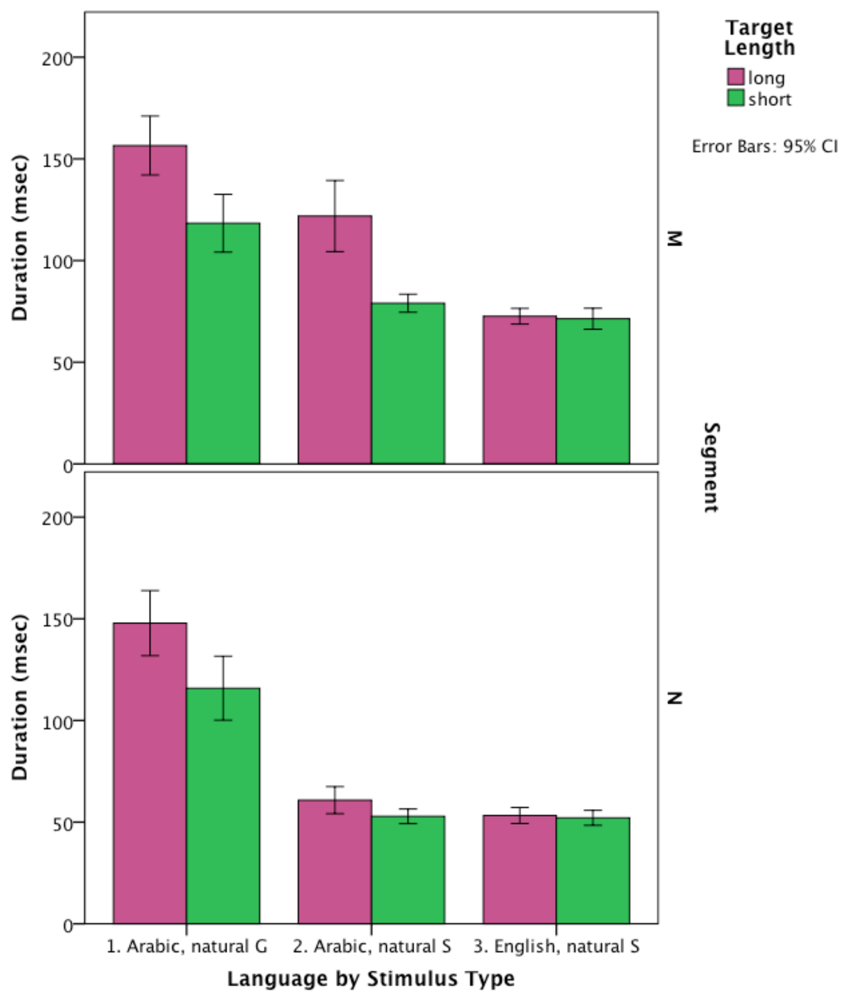

Figure 6: Experiment 3, mean raw segment duration by stimulus type

	Arabic, Natural G		Arabic, Natural S		English, Natural S	
	Long Target	Short Target	Long Target	Short Target	Long Target	Short Target
M	0.38 *(0.15)*	0.68 *(0.40)*	0.73 *(0.59)*	1.05 *(0.46)*	1.53 *(0.75)*	1.49 *(0.78)*
N	0.43 *(0.23)*	0.68 *(0.50)*	1.51 *(0.93)*	1.69 *(0.83)*	2.35 *(1.59)*	2.37 *(1.51)*

Table 13: Experiment 3, V1:C duration ratios by stimulus type

(Standard deviation in parentheses)

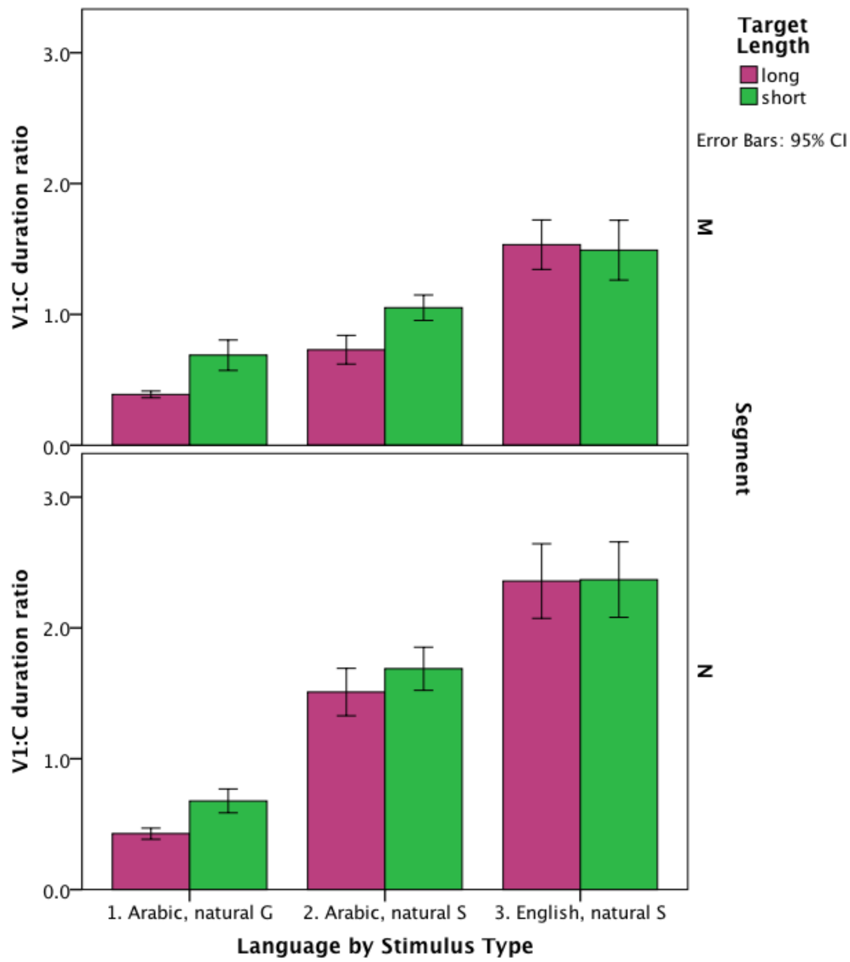

Figure 7: Experiment 3, V1:C duration ratios by stimulus type

4.6 Discussion

4.6.1 Arabic natural S condition versus English natural S condition

The lack of an effect of repetition for either language indicates that neither English speakers nor Arabic speakers affirmed the hypothesis that they would initially repeat a length contrast and then neutralize or emphasize it over the rest of the repetitions. Instead, it suggests that neither group of speakers changed the way they had encoded the new words over the course of the repetitions. Alternatively, the lack of a repetition effect could signal that the impression of the stimulus did not recede in the speakers' memory as it receded in time over the ten repetitions. In this case, the experiment would not promote word learning enough to prompt speakers to mentally encode the new words, and instead would encourage only superficial imitation.

The English speakers' average G:S duration ratios of 1.03 for *m* and 1.04 for *n*, were smaller than the G:S duration ratio obtained by Redford & Oh (1.25, collapsed across segments). It is possible that Redford & Oh's iambically-stressed stimuli allowed English speakers to repeat the contrast since the consonantal lengthening occurred at the onset of the stressed syllable. It has been established that English word-medial consonants at the onset of a stressed syllable are longer than word-medial consonants at the onset of an unstressed syllable (Umeda, 1977). In the stimuli of the current experiment, consonantal lengthening occurred at the onset of the unstressed syllable. Therefore, Redford & Oh's participants may have lengthened the target consonants simply because they perceived a stressed syllable, and imitated the syllable with consonantal lengthening because they would normally produce stressed syllables in this way.

In fact, in the context of second-language acquisition, segmental length variation in English may actually help English speakers learn to perceive and produce an L2 phonemic length contrast, compared to speakers of other language where length does not very systematically. McAllister, et al. (2005) tested a group of L2 Swedish speakers of L1 Estonian, L1 English, or L1 Spanish, on discrimination and production of the Swedish segmental length distinction. The speakers of L1 Estonian, who have a phonemic length category, performed very similarly to the native speaker controls, while the speakers of L1 English and Spanish differed from the control group. The speakers of L1 English, however, performed more similarly to the control group than did the speakers of L1 Spanish. Therefore, it would appear that variation in English segment length as a function of stress lends a prominence to length that can ultimately help English speakers to both produce and perceive a phonemic length category. While it does not appear that an advantage for English speakers as a result of English segmental length variation was borne out in Experiment 3 of the current study, the findings of McAllister, et al. raise the question of how speakers of Spanish or a similar language would have performed. In the Experiment 3 results, English speakers did produce very small durational differences between long and short segments as expressed in the G:S duration ratios and also in the mean raw segment durations. Given that over half of the English speakers were able to discriminate the contrast in Experiment 2, it is likely that they detected the contrast in Experiment 3 as well, but neutralized it in repetition because it did not fit with their usual framework for producing an unstressed syllable. If a group of Spanish speakers had performed the same tasks, they might not have reproduced even small durational differences as the English speakers did here.

As expected, we find a much larger contrast in the Arabic natural S condition than in the English natural S condition, albeit without a repetition effect. The lack of a repetition effect means that Arabic speakers did not gradually lengthen their long consonant productions to the length of their native language length contrast. Instead of adapting the novel length contrast to the standards of their native language contrast, the Arabic speakers were able to repeat a contrast without returning to the Kuwaiti Arabic G:S ratio. To explore the possible reasons why the segmental length contrast embedded in the Natural S stimuli did not drive Arabic speakers to imitate it according to the specifications of their native language G:S ratio, we turn to the Arabic relative timing results.

The V1:C duration ratio in the Arabic natural S condition raises some interesting questions when compared to the real word results from Experiment 1. We see a large shift in the V1:C duration ratio for the long *n* target between Experiment 1, where the ratio is .62, and Experiment 3, where the ratio is 1.51. The short *n* target in Experiment 1 had a V1:C ratio of 1.52, and a V1:C ratio of 1.69 in Experiment 2. The *n* targets show a greater shift overall than the *m* targets according to this relative timing ratio. The shift to greater V1:C ratios for *n* in Experiment 3 means that the V1 is longer in relation to C in the condition where natural singletons were matched with geminates derived from those same singletons. This result is also reflected in the G:S raw duration ratios for the Arabic real word results compared to the Arabic natural S results. The Arabic real word G:S ratio is 1.81 for M, and 2.48 for *n*. The Arabic natural S G:S ratio is 1.5

for *m*, and 1.14 for *n*. A possible explanation for the scale of change for *n* versus *m* between Experiment 1 and Experiment 2 lies in the stimuli construction process, where we lengthened each singleton consonant by 1.6 times its natural length in order to derive the geminates, irrespective of consonant type. When we created the stimuli, we did not realize that Kuwaiti Arabic *m* had a longer mean duration than *n* for both singleton and geminate consonants, nor that *m* had a smaller G:S duration ratio than *n*. Because short *n* has a much larger V1:C ratio than short *m* in the real Kuwaiti Arabic words, in addition to a shorter average duration, lengthening only the consonants yielded a V1:C ratio for derived-geminate *n* that more closely resembled that of singleton *n* than the ratio for derived-geminate *m* resembled singleton *m*. Therefore, it's possible that the listeners perceived the Arabic long *n* targets in Experiment 1 as resembling natural singletons more than natural geminates. Indeed, the V1:C ratio for the long N targets in Experiment 2 closely matches the same ratio for the short *n* targets in Experiment 1 (1.51 and 1.52 respectively).

To summarize, manipulating the absolute rather than relative durational values when deriving the geminates in the Natural S stimuli triggered an asymmetry in both the absolute and relative timing properties of *n* between the Arabic natural S condition of Experiment 2 and the real Arabic words of Experiment 1. This proposal also accounts for why the Arabic speakers discriminated the length contrast less well in the Natural S condition of Experiment 2: Arabic speakers perceived the derived long n consonants as singletons in the Natural S condition, but perceived the long and short segments in the Natural G condition as geminates and singletons respectively. However, without examining the perception data by segment, we cannot confirm the participants' accuracy on *m* versus *n* items.

We do not, however, find the same inverse results by consonant between Experiments 1 and 3 in English as in Arabic. The real word results from Experiment 1 show that English long *n* does has a longer average mean duration than long *m*, although the same is not true for singleton *n* compared to singleton *m*. The same difference in mean duration between *n* and *m* by target length appeared in Oh & Redford's (2012) results, which they attributed to greater decomposability of the *un*-prefixed words. Oh & Redford's proposal appears to be correct, given that the raw segment durations in Experiment 3 show that long and short versions of *n* are both shorter than long and short versions of *m,* in both absolute and relative terms. That is, morphological differences between the *un*- versus *in*- prefixed words drives variation in segment length, causing *n* to be longer than *m*. But when morphological conditions are identical between words containing medial *m* and *n*, as in the Experiment 3 nonwords, *n* is shorter overall compared to *m.*

4.6.2 Arabic natural S condition versus Arabic natural G condition

A comparison of the results between the Arabic natural S condition and the Arabic natural G condition shows that deriving singletons by shortening geminates in the Natural G condition did not yield the same degree of categorical ambiguity for the segment *n* as in the natural S condition. This is likely due to the fact that Kuwaiti Arabic long *m* and long *n* have very similar V1:C duration ratios, unlike short *m* versus short *n*. In the Natural G reproductions, both target lengths of *m* and *n* are longer than their matched counterparts in the Natural S condition in both absolute and relative terms.

Additionally, the timing properties of *m* and *n* are much closer to one another than they are in the Natural S condition or in the real word results from Experiment 1. If the derived geminates in the Arabic Natural S condition (at least for segment *n*) were perceived as singletons, it makes sense that listeners would perceive the Natural G condition as containing two separate length categories: the natural geminates, perceived as geminates, and the derived singletons, perceived as singletons.

V. General Discussion

In summary, the experiments presented here show that both English and Arabic have long and short consonants, but that the length contrast is greater in Arabic than English. Neither the perception nor production results for English speakers aligned with those of Redford & Oh (2011), suggesting that the lack of a phonemic length category in English confines English speakers to reproducing length variation as a function of word stress, as well as in a handful of words where a long consonant is specified at a morpheme boundary that falls at the onset of a stressed syllable. While it appears that English speakers cannot manipulate segment length independent of other factors such as stress or boundaries, the lack of a phonemic length category in English does not necessarily prevent English speakers from discriminating the contrast. With regards to Arabic, the results of the current experiments suggest that the phonemic status of length in Arabic gives rise to perceptual sensitivity to segment length, and also allows speakers to more easily lengthen consonants.

Figure 8 illustrates the relationship of perception to production, for all speaker groups and conditions. Kuwaiti Arabic speakers' discrimination ability predicts the magnitude of G:S ratio produced in Experiment 3, where a better discrimination ability predicts a larger G:S ratio. On the other hand, there does not exist a meaningful relationship between discrimination ability and G:S ratio for the English Natural S condition.

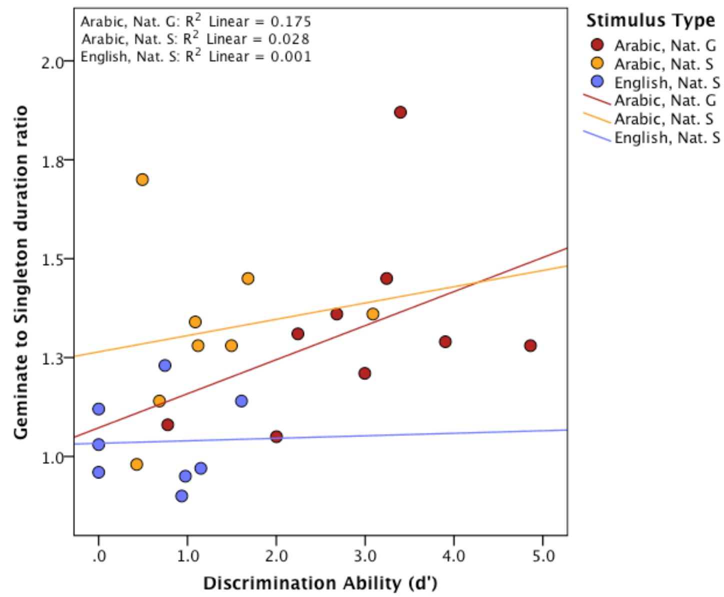

Figure 8: G:S duration ratio and discrimination ability, by participant

“G:S duration ratio” refers to the ratio produced in Experiment 3

Although no repetition effects were found for any condition, it still appears that the experimental design succeeded in examining the effects of a length contrast on imitation and perception abilities. That is, the experiments succeeded in capturing language-based differences in perception and reproduction of a length contrast. Previous work on word and nonword imitation has indeed shown that just one repetition can capture speakers’ ability to imitate fine phonetic detail and sensitivity to word frequency. According to Goldinger’s framework of an episodic lexicon (1998), the process of imitating and auditory stimulus (or “shadowing”) involves accessing a multitude of unique episodic traces for that word. An episodic memory for a given word is created and stored with every exposure to that word. Goldinger found that high frequency words prompt a faster response time, and that low frequency words prompt more accurate imitation in terms of word duration and fundamental frequency. Goldinger found the same results when participants shadowed nonwords that they had been exposed to at low, medium, and high frequencies. Given these results, it does not appear that we would have been able to elicit more accurate imitations by exposing participants to the stimuli more times.

Shockley et al. (2004) built on Goldinger's study in order to examine English speakers' ability to imitate subphonemic variation in voice onset time (VOT). They found that speakers systematically extended VOT in shadowed words depending on whether or not the stimulus contained an extended VOT. Shockley et al. also found that listeners proved perceptually sensitive to the greater similarity between shadowed words and the auditory stimuli, compared to the smaller amount of similarity between baseline word productions and the auditory stimuli.

The Arabic Natural G results make sense in light of Shockley et al.'s (2004) results, since segmental length in Arabic and VOT in English are both phonemic. The results of both the current study and Shockley et al. suggest that the phonemic status of a sound feature allows speakers to accurately imitate it. Although the speakers in Shockley's study were able to imitate subphonemic variations in VOT, this could have been due to the fact that English speakers are accustomed to manipulating VOT as a function of phonemic category structure for voiced and voiceless consonants. In other words, VOT in English is a phonetically relevant dimension of sound that English speakers can potentially use to create a phonemic distinction. The English Natural S results of the current study suggest that the presence of segmental length variation may have sensitized English speakers to length variation on the level of perception, but that allophonic variation alone was not sufficient to allow speakers to reproduce the length variation that they heard.

5.2 Implications of current study

Prior research into second language (L2) Japanese learners' development of a phonemic length category has shown that learners without a phonemic length category in their native language are able to develop a perceptual category for segmental length, but face more difficulty in learning to produce the contrast in a nativelike way. Hayes-Harb & Masuda (2008) found that after just one year of language training, L1 English learners of L2 Japanese performed similarly to native Japanese speakers in terms of discriminating the Japanese segmental length contrast, while novice learners performed at chance. However, the productions of the length contrast by the L2 learners with one year of Japanese training did not resemble the productions of native speakers. Similarly, Mah & Archibald (2003) found in a case study of an intermediate L2 Japanese learner (L1 English) that the learner consistently produced geminate consonants that were significantly longer than singleton consonants, but the durational ratios of her short and long consonants did not match those of native speakers, and were highly variable. Han (1992) found similar production results in an examination of the Japanese length distinction as produced by speakers of L1 English.

These findings, taken together with the results of the current study, suggest that L2 Arabic learners' ability to accurately perceive the Arabic length contrast would precede the ability to accurately produce the contrast. Furthermore, the linkage between English consonant length variation and lexical stress as established by the results of the current study and those of Redford & Oh (2011) suggest that teaching L1 English learners of L2 Arabic to manipulate consonant length might become easier or more successful if initial lessons were grounded in the idea of producing a geminate as the beginning of a stressed syllable, and later expanded to other phonetic contexts.

References

Abercrombie, David. (1967). *Elements of General Phonetics.* Edinburgh: Edinburgh University Press.

Al-Tamimi, F. (2004). An experimental phonetic study of intervocalic singleton and geminate sonorants in Jordanian Arabic. *Al-Arabiyya*, *29*, 37-52.

Ardila, A. (2010). A review of conduction aphasia. *Current Neurology and Neuroscience Reports*, *10*(6), 499-503.

Bernal, B., & Ardila, A. (2009). The role of the arcuate fasciculus in conduction aphasia. *Brain*, *132*(9), 2309-2316.

Enomoto, K. (1992) Interlanguage phonology: the perceptual development of durational contrasts by English-speaking learners of JP. *Edinburgh Working Papers in Applied Linguistics*, *3*, 25–35

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251.

Ham, W. (2013). *Phonetic and phonological aspects of geminate timing*. New York: Routledge.

Han, M. S. (1992). The timing control of geminate and single stop consonants in Japanese: A challenge for nonnative speakers. *Phonetica*, *49*(2), 102-127.

Hassan, Z. M. (2002). Gemination in Swedish and Arabic with a particular reference to the preceding vowel duration: an instrumental and comparative approach. In *Proceedings of Fonetik*, 81-85.

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

Hardison, D. M., & Saigo, M. M. (2010). Development of perception of second language Japanese geminates: Role of duration, sonority, and segmentation strategy. *Applied Psycholinguistics*, *31*(01), 81-99.

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

Hay, J. (2007). The phonetics of 'un'. *Lexical Creativity, Texts and Contexts.* Amsterdam: John Benjamins, 39-57.

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

Kato, H., Tajima, K., & Akahane-Yamada, R. (2001). Native and non‑native perception of phonemic length contrasts in Japanese. *The Journal of the Acoustical Society of America*, *110*(5), 2686-2686.

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

Kaye, A.S. (2005). Gemination in English. *English Today*, *21*(2), 43-55.

Khattab, G. (2007). A phonetic study of germination in Lebanese Arabic. *Proc. of 16th International Congress of Phonetic Sciences*, Saarbrücken, 153-158.

Kohn, S. E. (Ed.). (1992). *Conduction aphasia*. Hillsdale, NJ: Psychology Press.

Kreidler, Charles W. (2004). *The Pronunciation of English.* Oxford: Blackwell.

Ladefoged, Peter. (2001). *A Course in Phonetics* (4^{th} ed.). Fort Worth, TX: Harcourt.

Mah, Jennifer & Archibald, John. (2003). Acquisition of L2 Length Contrasts. *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002),* ed. Juana M. Liceras et al., 208-212.

McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, *30*(2), 229 – 258.

Oh, G. E., & Redford, M. A. (2012). The production and phonetic representation of fake geminates in English. *Journal of Phonetics*, *40*(1), 82-91.

Redford, M. A., & Oh, G. E. (2011). Reproducing singletons and fake geminates. *Proc.of 17^{th} International Congress of Phonetic Sciences*, Hong Kong, 1674-1677.

Shockley, K., Sabadini, L., & Fowler, C. A. (2004). Imitation in shadowing words. *Perception & Psychophysics*, *66*(3), 422-429.

Tajima, K., Kato, H., Rothwell, A., & Munhall, K. G. (2003, August). Perception of phonemic length contrasts in Japanese by native and non-native listeners. In *Proc. of 15th International Congress of Phonetic Sciences*, Barcelona,1585-1588.

Trask, R. L. (1996). *A Dictionary of Phonetics and Phonology*. London: Routledge.

Umeda, N. (1977). Consonant duration in American English. *The Journal of the Acoustical Society of America*, *61*(3), 846-858.

Ylinen, S., Shestakova, A., Alku, P., & Huotilainen, M. (2005). The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. *Language and Speech*, *48*(3), 313-338.