

Korean Perception of Arabic Phonemes

Seung-ah Hong and Priyankoo Sarmah

Hankuk University of Foreign Studies

hongseungah@hufs.ac.kr, pri@hufs.ac.kr

Abstract

This study aims at figuring out how native Korean speakers perceive and categorize Arabic phonemes into their phoneme inventory. Two perception experiments are conducted on two groups of Korean speakers (22 in total) with stimuli based on 25 of Arabic consonant phonemes, designed into 176 pairs. Experiment I is an AX discrimination task designed to tease out the acoustic cues relevant to native Korean speakers in identifying Arabic phonemes. Experiment II is designed to see how Arabic phonemes are mapped onto Korean phonemes by means of orthography and to see Korean speakers' perceived similarity of Arabic phonemes to Korean phonemes by a similarity rating task.

1. Introduction

Non-native perception of phonemes has been of considerable interest in the field of speech perception. It has been noticed that speakers usually tend to map the non-native phonemes into the existing phoneme inventory in their first languages (Werker & Tees 2002). On the other hand the same study reports that cross linguistic infant speech perception shows that human beings are aware and able to distinguish a wide variety of speech sounds before they are exposed and adapted to the native language phonology. Adults however, exhibit low levels of sensitivity towards non-native phonemes resulting in difficulty in the perception and in turn production of a foreign language. Part of the problem arises due to the fact that speech sounds contain a variety of acoustic cues and they are only sensitive to the cues that already exist in their native phonology, be it in a phonological or phonetic form (Werker 1984). Hence, the second language (L2) phonemes are interpreted based on the phonemic inventory of the native language (L1). In the categorization of the L2 phonemes, their degree of the similarity to the L1 phonemes serve as a base in the acquisition of L2 (Flege 1995). Native speakers of Korean exhibit similar patterns in the perception of non-native speech sounds; for example Schmidt (1996) reports that native Korean speakers assign higher similarity scores to Korean phonemes for the English /m/, /h/, /t/, /p/, and /j/ phonemes (more than 70% responses as "exactly like"). Similarly, they assign lower similarity scores for /f/, where it received the only 6% response as the "exactly like" rating and also the mean rating was "very different". Apart from that vowel context and token differences also have significant effects on the similarity judgment by the Korean speakers.

Even though there abundant studies on Korean speakers' perception of English phonemes, not much is known about their perception of Arabic phonemes. Hence, in this study we investigate how Korean

speakers perceive phonemes of Arabic which has more variety of phonemes in its inventory than in Korean. Korean has only 19 consonant phonemes (Schmidt 1996), while Arabic has 28 of them (Kong 1993) including four emphatic sounds and six uvulars and pharyngeals. Hence, it is interesting to see how Korean speakers discriminate among the Arabic phonemes and map them into their L1 phonology. It is expected that native Korean speakers would find it difficult to discriminate among some of the Arabic phonemes since there are no phonetic counterparts available in Korean phonology.

We conducted a series of perception tests on Korean learners of Arabic (the KA group) and Korean speakers without any knowledge or experience of the Arabic language (the KN group). A series of discrimination, phoneme identification and similarity rating tests were conducted to see Korean speakers' perception of Arabic phonemes.

2. Methodology

2.1. Stimuli

The stimuli set for this study consists of 25 Arabic consonantal phonemes (1 glottal stop and 2 semi-vowels namely, /ʔ/, /w/ and /j/ are excluded in this study). They are produced by a 23 year old male native speaker of Arabic from Kuwait. Recordings were made digitally at a sampling rate of 44100 Hz in mono channel using Praat 5.1.03 (Boersma and Weenink 2009) in a quiet room. The speaker was specifically asked to pronounce the stimuli in Modern Standard Arabic (MSA). All stimuli are blocked in a vowel, monosyllabic (CV) structure. These consonantal phonemes are originally recorded in six different Arabic monophthong contexts (3 short vowels /a/, /i/, /u/ and 3 long vowels /a:/, /i:/, /u:/) with two iterations each.

It has been attested that Arabic pharyngeal contrasts in consonants are spread into adjacent vowels. In Al-Masri & Jongman (2003), it is seen that the Arabic pharyngeal consonants lower the second formant (F2) of the following front vowels. As a result of this perturbation, the front vowels following pharyngeal consonants are phonetically distinct from front vowels occurring in other environments. Among the three consonants of Arabic, F2 lowering is noticed to be lowest for the vowel /u/ (Jongman et al. 2007). Hence, to prevent such a possible influence, only the stimuli with the nucleus /u/ are used in this study. To further confirm that the vowel used in this study does not vary in itself significantly, we conducted an acoustic analysis of the vowels used in the stimuli with Praat 5.1.03. In Table 1, we report the F1 and F2 values of the Arabic /u/ vowels in the stimuli and compared them with the values of the Korean /u/ vowel as reported in Yang (1996). It is noticed that the F1 and F2 values of Arabic /u/ are only marginally different from the Korean /u/. Hence, using the /u/ vowel in the stimuli also reduces the possibility of Korean speakers identifying two stimuli as different, based on vowel characteristics.

For Experiment I, the 25 stimuli are arranged into 63 minimal pairs in two vowel contexts (C₁u vs. C₂u and C₁u: vs. C₂u:) and 25 control pairs (C₁u vs. C₁u and C₁u: vs. C₁u:) resulting in a total of 176 pairs

of stimuli. Experiment I is an AX discrimination test, where each of the 176 stimuli is presented twice in a random order. For experiment II, the 25 stimuli are used for a cross linguistic consonant mapping and similarity rating task.

Table 1. Average duration, F1 and F2 of Korean /u/ (Yang 1996) compared to that of Arabic /u/

		Duration (ms) [<i>SD</i>]	F1 (Hz) [<i>SD</i>]	F2 (Hz) [<i>SD</i>]
Korean		NA	369 [43]	981 [141]
Arabic	Long	330 [32]	326 [58]	799 [133]
	Short	128 [21]	381 [53]	986 [144]

2.2. Subjects

Twenty two native speakers of Korean from the Seoul area (21 male, 1 female) participated in this study. The participants are in the 19 to 35 years age group, with a mean age of 26.6 years ($SD= 3.9$). None of the subjects reported any medical conditions related to hearing or having a cold at the time of testing. The Participants are divided into two groups based on their experience with the Arabic language. The KA group consists of ten college students majoring in Arabic at the Hankuk University of Foreign Studies (HUFS), Korea. On average, they received 3.3 years ($SD = 0.5$) of Arabic instruction. The KN group consists of twelve participants who do not have any prior knowledge of Arabic. All the 22 speakers have a minimum of college level of education and phonetically untrained. Most of the participants in the KN group are monolinguals with minimal experience in a foreign language. Although all subjects participated in experiment I, one participant from the KA group failed to participate in experiment II.

2.3. Procedure

Two experiments were conducted in this study. In experiment I the participants were required to perform a discrimination task. The 352 stimuli (63 minimal pairs and 25 identical pairs in 2 different vowel contexts, long and short /u/, played twice) were and randomly presented to the subjects with an inter stimulus interval (ISI) of 2500 milliseconds. They listened to the stimuli presented using Praat's MFC (Multiple Forced Choice) tool on a computer, through a pair of headphones. The order of experiments was counterbalanced across the participants. The forced-choice discrimination (AX) task is given to the participants, where they were asked to decide whether the two phonemes are same or different by clicking an appropriate box on the computer screen. The average duration of experiment time was about 40 minutes for each participant. Participants were allowed to take breaks as many times as they wanted during the test. The results of the test were collated to a spreadsheet for further analyses.

Experiment II is an identification and rating task performed on the 25 stimuli originally produced by

a native Arabic speaker. For this task, only the stimuli with long /u/ vowels were used. They were played in the same order to all the participants. Participants were allowed to replay the stimuli as many times as they wanted. They were then asked to write down what they hear in Korean orthography. Participants were also asked to evaluate the similarity of the sounds are to the Korean sounds they relate to, using a scale of 0 to 5 (0 = no match, 5 = exact match).

3. Results

3.1. Experiment I

Correctness between KA and KN group

Both the groups demonstrated high level of discrimination in both the short and long vowel contexts (above 90% correct). The average percentage of correctness of both the KN and KA groups was 93% ($SD = 25$). When compared across all vowel contexts, the two groups demonstrated no significant difference in terms of the percentage of correctness [t-test, $p = 0.947$].

Effect of Vowel Context

The results showed that the vowel context has a significant effect on the correctness judgments across both participant groups. Overall, participants in both groups performed better in the short vowel context (94% correctness) than in the long vowel context (92%). An independent t-test confirmed that the correctness of the participants differed significantly in terms of vowel length [$t(7742) = 2.63, p < .05$].

Contrast Types and Perception

In order to see the effect of consonant types on correctness, the Arabic stimuli sets were categorized into 13 distinct categories based on voicing, place of articulation, manner of articulation and pharyngealization. Table 2 shows the different categories and their corresponding numerical indexes. As noticed in Table 2, type 1 is the category assigned to stimuli sets where both the consonants are identical. In case of type 2 the two consonants differ in terms of the presence and absence of pharyngeal contrasts. Similarly, type six consists of stimuli where the two consonants differ both in terms of place of articulation and manner. The stimuli categories therefore can range from having none (type 1) to four distinct contrasts (type 13).

The perception of each contrast type was calculated by speakers' responses in each subject group (KA and KN) by the two vowel duration contexts namely, long and short. The results of the correct judgments by the two groups of speakers are presented in a percentage scale in Figure 1 and Figure 2. As noticed from the two figures, the two groups demonstrate varying degrees of correctness among the 13 types.

Generally, the KA group performs better when the consonant is followed by a short vowel. The KA

group performs the worst in discriminating between pharyngeal consonants and their non-pharyngeal counterparts. The KA group performs below 95% scores in contrast types 3 to 7 indicating that a single dimension of contrast is not sufficient for them to discriminate one token from another. It also indicates that manner and place contrasts pose a challenge for the KA group in the correct discrimination of tokens. However, voicing contrasts reinforced by another dimension may help in correct perception of the tokens as seen in case of contrast types 8 to 11. In case of contrast type 12, the occurrence of the place and manner dimensions at the same time negatively interfere the correct discrimination of the speech tokens.

Table 2. Contrast types and numerical indexes

Type	Contrasts	Type	Contrasts
1	Identical stimuli	8	Voicing & pharyngeal
2	Pharyngeal	9	Voicing & manner
3	Manner	10	Voicing & place
4	Place	11	Voicing, place & pharyngeal
5	Place & manner	12	Voicing, place & manner
6	Place, manner & pharyngeal	13	Voicing, place, manner & pharyngeal
7	Voicing		

Compared to the experienced KA group, the naive KN group performs in a more consistent manner. The KN group even fails to correctly identify the same phonemes as same, resulting in 92% of correctness in the identical pairs. Surprisingly, they perform better than the KA group in discriminating the pharyngeal phonemes from their non-pharyngeal counterparts. They also perform better than the KA group in discriminating between phonemes that differ in their manner of articulation. Usually, higher correct discrimination rates are noticed in case of short vowel context. In spite of their better performance, single contrasts and ones accompanied by place and manner dimensions pose a challenge to the KN group too. In both the participant groups, the longer vowel duration seems to facilitate better identification of voicing contrasts. Nevertheless, vowel length impedes correct perception in both the groups.

An analysis of variance (ANOVA) test was conducted on the data set to see if the types of contrasts are significantly different from each other in terms of their corresponding correctness scores. The ANOVA test conducted with correctness scores as dependent variable and types of contrasts as independent variable showed significant interaction between the two variables [$F(12, 7744) = 22.35, p < .01$]. Another ANOVA test conducted to see if there is a participant group and contrast types interact together with the correctness scores. The results showed that groups and contrast types do interact together with the correctness scores significantly [$F(12, 7744) = 8.13, p < .01$].

For a more detailed analysis of the discrimination task performed by the native speakers of Korean, the stimuli pairs were grouped into 28 different categories as suggested by Boudraa et. al (2008) based on

the distribution of features (see Table 4). The stimuli sets were further categorized according to existence or absence of these features. For example, type 1 stimuli consist of two identical tokens whereas stimuli of the type 2 category differ from each other only in terms of the FL feature. Similarly, members of the type 18 category differ from each other in terms of four different features. As noticed from the results represented in Figure 3 and Figure 4, when categorized by feature differences, the naive group KN performs better than the experienced KA group.

Figure 1. Correctness (%) for contrast types of the KA group

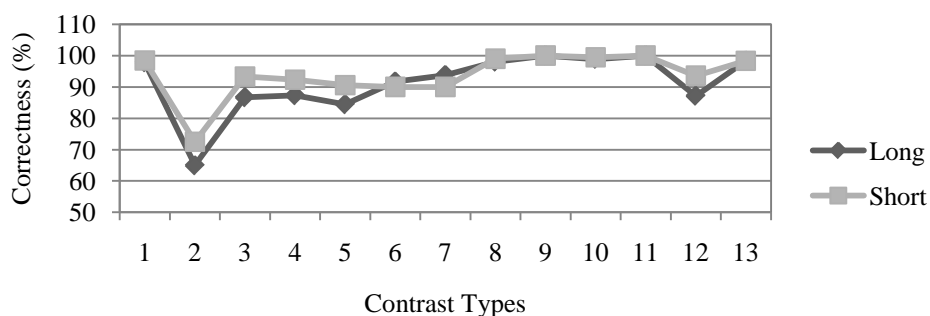


Figure 2. Correctness (%) for contrast types of the KN group

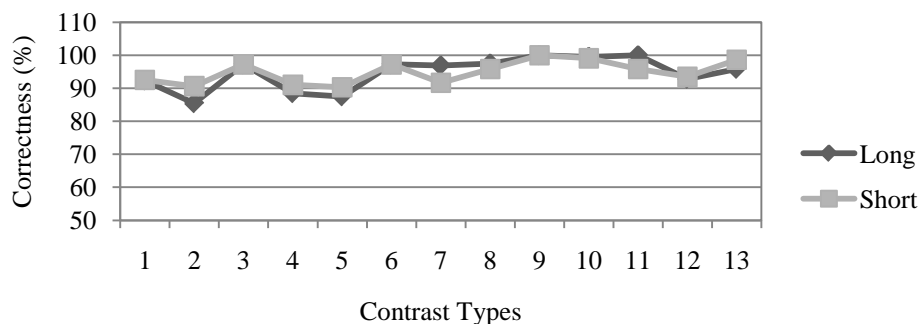


Table 3. Features for Arabic suggested by Boudraa et. al (2008)

Feature	Contrasts	Feature	Contrasts
VC	Vocalic vs. Non-vocalic	AC	Grave vs. Acute
CT	Continuant vs. Abrupt	TN	Tense vs. Lax
FL	Flat vs. Plain	IT	Interrupted vs. Non-interrupted
NZ	Nasal vs. Oral	ST	Strident vs. Mellow
CM	Compact vs. Diffuse		

Both groups score less than 90% correctness scores in both vowel contexts for the stimuli that differ only in tenseness (type 5) and for the stimuli that differ in interruptedness and continuance (type 15). Additionally, the KN group scores less than 90% scores for the stimuli of types 8, 11 and 21.

Noticeably, the KA group also scores less than 90% scores in these three categories but only in the long vowel contexts. Similarly, the KA group scores less than 90% scores in the flatness feature category (type 2) in both vowel contexts. However, the KN group scores less than 90% in the same category on in the context of long vowels. Apart from that the KA group scores less than 90% in long vowel context in type

Table 4. Stimuli sets categorized according to feature contrasts

Type	Features	Stimuli Sets					
1	-	identical stimuli pairs					
2	FL	ð vs. ð [̣]	ð vs. t [̣]	t vs. t [̣]	d vs. d [̣]	s vs. s [̣]	
3	ST	θ vs. ð	t vs. d	t [̣] vs. d [̣]	s vs. z	ħ vs. ʕ	k vs. q x vs. ʁ
4	ST+FL	θ vs. ð [̣]	t vs. d [̣]	d vs. t [̣]	s [̣] vs. z		
5	TN	f vs. h	ð vs. z	ħ vs. x	ʁ vs. ʕ		
6	TN+FL	ð [̣] vs. z					
7	TN+ST	ħ vs. ʁ	x vs. ʕ				
8	CM	f vs. x	s vs. ʃ	h vs. h			
9	CM+FL	s [̣] vs. ʃ					
10	CM+ST	f vs. ʁ	ʃ vs. z	h vs. ʕ			
11	CM+TN	f vs. h	h vs. x				
12	CM+TN+ST	f vs. ʕ	h vs. ʁ				
13	CT+VC+NZ	m vs. b					
14	CT+CM	l vs. r					
15	IT+CT	θ vs. t	ð vs. d	ð [̣] vs. d [̣]	q vs. ʕ		
16	IT+CT+FL	θ vs. t [̣]	ð vs. d [̣]	ð [̣] vs. d			
17	IT+CT+ST	θ vs. d	ð vs. t	ð [̣] vs. t [̣]	ħ vs. q	k vs. ʕ	
18	IT+CT+ST+FL	θ vs. d [̣]	ð [̣] vs. t				
19	IT+CT+TN	k vs. x	q vs. ʁ				
20	IT+CT+TN+ST	b vs. f	k vs. ʁ	q vs. x			
21	IT+CT+CM	z vs. dʒ					
22	IT+CT+CM+ST	h vs. q					
23	AC	n vs. m					
24	AC+TN	dʒ vs. q					
25	CT+TN+ST	dʒ vs. k					
26	AC+IT+CT	dʒ vs. ʁ					
27	AC+IT+CT+ST	dʒ vs. x					
28	AC+IT+CT+TN	dʒ vs. ʕ					

19 and 14 where the stimuli differ in three and two features, respectively. In summary, it is evident that the two groups perform significantly lower scores in long vowel contexts. Both groups perform less than 75% of correctness scores in the type 15 category which makes it evident that both groups have problems in discriminating fricatives from their stop counterparts.

Figure 3. Correctness (%) for contrast types based on the features for the KA group

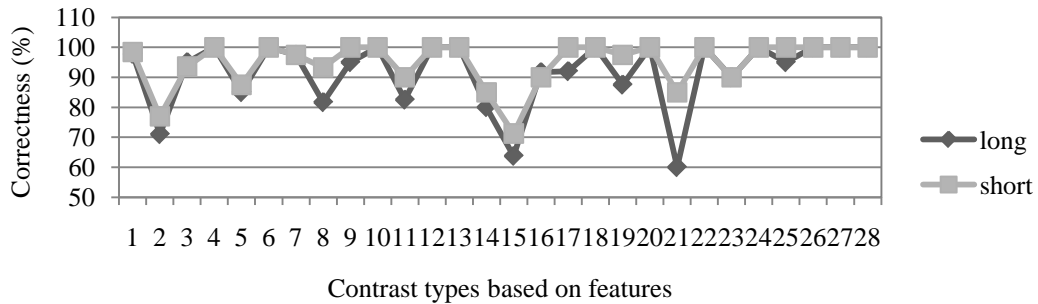
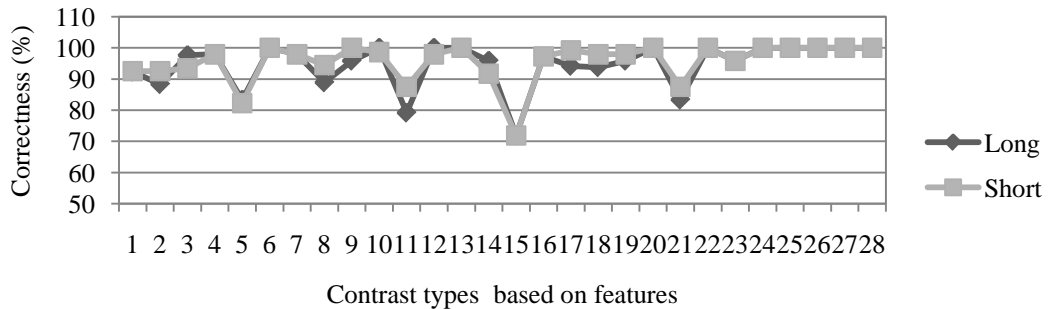


Figure 4. Correctness (%) for contrast types based on the features for the KN group



An ANOVA test was conducted to see if correctness interacts significantly with the two groups (KA and KN) and feature contrast types. The results of the test showed there is a significant difference between the two speaker groups in correctness scores for each feature contrast type [$F(27, 7744) = 4.1, p < .01$].

3.2. Experiment II

Cross-language Consonant Mapping

The two subjects groups (KA vs. KN) show difference in the number of mapping relationships they used to map Arabic consonants to Korean consonants. The participants of KN group show more variety in labeling Arabic consonants to Korean consonants. As seen in Figure 5, participants in KA group exhibit more consistency in labeling the Arabic consonant to Korean consonant phonemes. For all of 25 stimuli used in the experiment II, KA subjects labeled each stimulus no more than a one-to-three mapping.

Moreover, 52% (13 stimuli out of 25) showed one-to-one relationships, while the Subjects of KN Group has mapped the Arabic consonants in a more variety of ways (up to one-to-five mapping). Figure 6 to 8 present detailed information on the each Arabic phonemes and their correspondence to Korean phonemes.

Some Arabic consonants have one-to-one relationship in both groups of subjects. Arabic liquids and nasals are generally consistently mapped onto one or two Korean consonant phonemes, which are also liquid and nasal Korean phonemes. Similarly, both the labial consonants (/f/ and /b/) are mapped onto the Korean bilabial stop consonants. The Arabic /b/ is mapped into Korean voiceless unaspirated bilabial stop, while /f/ shows more variety in labeling choice but most of the times identified as the aspirated bilabial stop /p^h/. In Figure 6, we can also see that the participants' best mapping for the Arabic alveolar fricatives is onto the Korean alveolar fricatives in both KA and KN groups. While the majority of the KA group maps both the plain and the emphatic alveolar fricatives onto the tense Korean alveolar fricative, most of the participants in the KN group mapped the Arabic emphatic alveolar fricative as a lax Korean alveolar fricative. In addition to this, the inter-dental fricative /θ/ in Arabic is primarily mapped onto Korean alveolar fricatives. These mapping patterns are consistent with that of Korean speakers' in perceiving English phonemes as reported in Schmidt (1996).

However, as seen in Figure 7, the two groups of Korean speakers exhibit a distinct pattern in the mapping of the dental emphatic phonemes (/t^ɖ/, /d^ɖ/ and /ð^ɖ/). While the participants in the KA group judged the phonemes more close to the dental or alveolar sounds, the majority of KN subjects mapped them to the bilabial phonemes in Korean.

Similarity Rating

The Korean speakers in both groups were asked to rate the similarity between the Arabic stimuli they hear and the Korean consonant to which they map the particular stimuli on a scale of 0 to 5. The participants were asked to rate the mapping as 0, if they considered the stimuli and the mapped Korean phoneme to have no similarity. On the other hand they were asked to rate the mapping as 5 if they considered the two phonemes to be an exact match. In Figure 8 average similarity rating for each of the Arabic stimuli for its closest consonant in Korean assigned by all participants in the study is presented. As seen from the figure, there is no significant rating difference between the two speaker groups. Schmidt (1996) reports that when the native Korean subjects in her study had more experience with English, they rate English sounds to be less similar to Korean sounds. Hence, it was expected that subjects in the KA group would give the lower rating to the similarity of Arabic phonemes to Korean phonemes. However, it was not the case in this study.

Vowel duration and Syllable Restructuring

Although the original stimuli in Arabic are monosyllabic CVV stimuli, native Korean subjects while transcribing in Korean alphabets often transcribe them as the polysyllabic stimuli. This tendency of

Figure 5. Number of distinct perceptions for Arabic stimuli

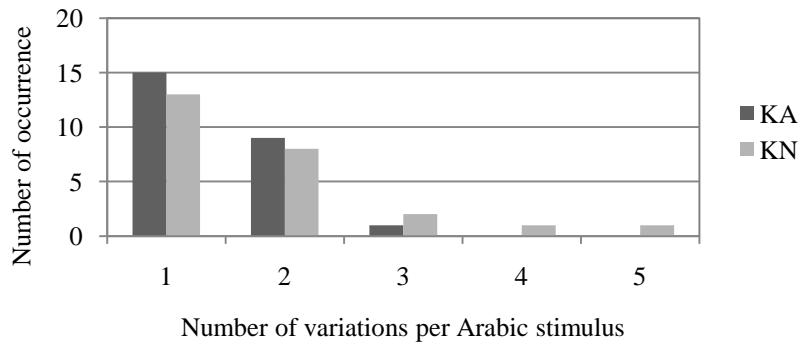


Figure 6. Arabic fricatives mapped onto Korean

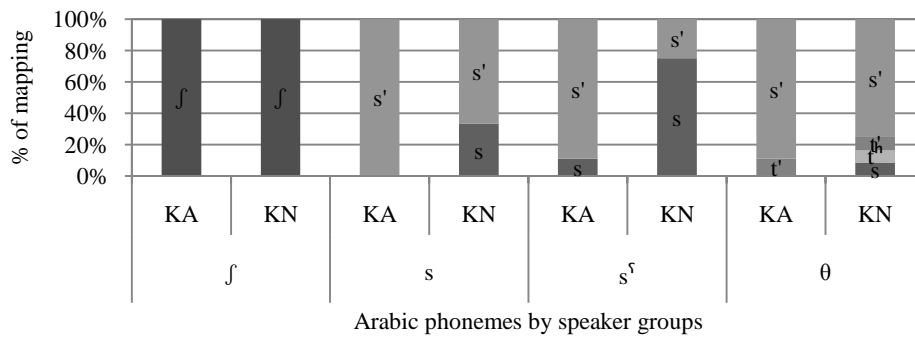


Figure 7. Arabic coronal consonants mapped onto Korean

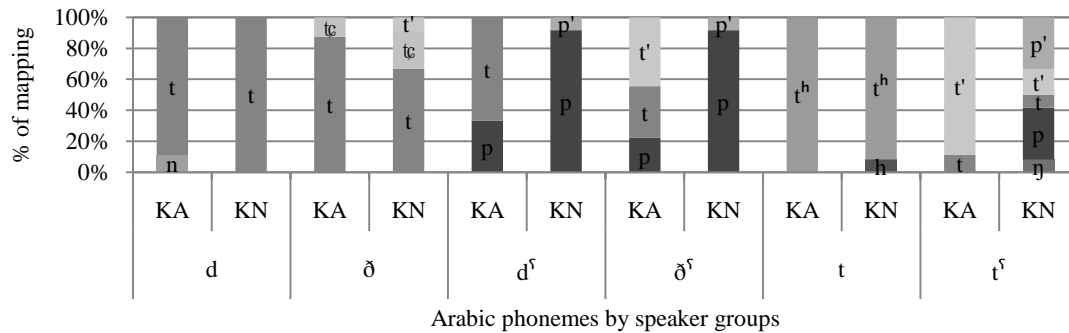


Figure 8. Average similarity rating to the closest Korean phoneme for each Arabic consonant

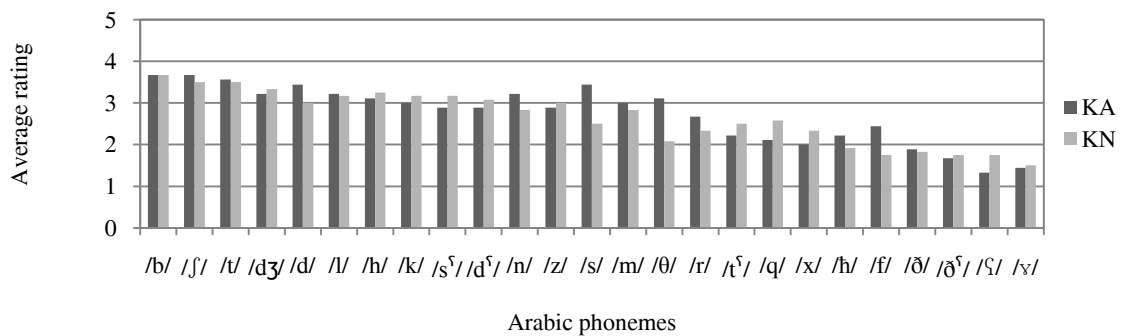


Table 5. Perception of Arabic CVV syllables as various syllable structures

	Monosyllable	Disyllable	Trisyllable
KA	72.2%	23.3%	4.5%
KN	49.8%	44.7%	5.5%

syllabification appeared higher among the subjects in the KN group. The stimuli used in the second experiment are with long vowel contexts. So, the vowel duration would be interpreted as an extra syllable.

The selection of CVV structures among the disyllable structures gives more weight to the hypothesis.

As in Table 5, KA subjects identified the stimuli monosyllables in general. However, a large number of KN subjects identified the stimuli as disyllables, and even as trisyllables in some cases. Hence, the listeners who are less exposed to the L2 phonology tend to interpret the L2 speech sound based on their L1 phonology and thus can be more sensitive to the duration of the vowel.

4. Discussion

This study investigated Korean native speakers' (trained and untrained) perception of Arabic phonemes. It is noticed from experiment I that training does not have significant effect on the Korean speakers' perception of Arabic phonemes. In the discrimination test, it was seen that the KA groups performs worse in discriminating between the emphatic consonants from the non-emphatic ones. On the other hand, the KN group performs significantly better in the same contrasts. In other words, we propose that KA groups' training in Arabic makes them lose the idiosyncrasy of emphatic consonants when compared to their native language, Korean. However, the KN group still finds the emphatic consonants in Arabic as idiosyncratic when compared to their native phonology and hence are able to perform better in the discrimination task. In other words, the KN group is able to successfully identify the acoustic cues for the emphatic sounds that are unavailable in their native language, and use them in discriminating them from the non-emphatic ones.

The results of the perception test also demonstrated that across all subjects Arabic consonants are better perceived in the short vowel context than in a long vowel context. Even though some varieties of Korean use vowel duration minimally, in general, Korean speakers' from the Seoul area cannot perceive the long and short vowel contrast (Park 1994). Moreover, the average duration of Korean vowels [86 ms ($SD= 32$), as in Yang 1996], correspond more closely to the duration of Arabic short vowels. Hence, the long vowel context presumably adds to the difficulty in perception of the Arabic phonemes resulting in lower correctness scores.

In experiment II, it is noticed that the KA group tends to have less number of mappings associated with each Arabic stimuli, whereas that KN group has more variety of mappings. It can be argued here that the KA group has already phonologized the Arabic consonant categories and hence tend to map them

more rigidly than the KN group. The KN group is aware of the consonantal differences however, as they lack training in they are not able to assign the auditory stimuli to rigid categories.

References

- Al-Masri, M. & Jongman, A. 2003. Acoustic correlates of emphasis in Jordanian Arabic: preliminary results. *Proceedings of the 2003 Texas Linguistics Society Conference*, 96-106.
- Boersma, P. & Weenink, D. 2009. Praat.5.01.09. Software. Retrieved May, 1st, 2009.
- Boudraa, B. Boudraa, M. & Guerin, B. 2008. Arabic Diagnostic Rhyme Test using minimal pairs. *Journal of the Acoustical Society of America*, 123(5): 3324.
- Flege, J. E. 1995. Second Language Speech Learning: Theory, Findings and Problems. In Strange, W, ed. *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross Language Speech Research*. Timonium, MD: York.
- Kong, I-j. 1993. *Arabic Phonetics*. Seoul: Songsan.
- Park, J-W. 1994. Variation of Vowel Length in Korean. In Kim-Renaud, Y-K, ed. *Theoretical Issues in Korean Linguistic*. CSLI publications.
- Schmidt, A. M. 1996. Cross-language identification of consonants: part 1. Korean perception of English. *Journal of the Acoustical Society of America*, 99(5): 3201-3211.
- Jongman, A., Herd, W. & Al-Masri, M. 2007. Acoustic correlates of emphasis in Arabic. *Proceedings of the 16th Annual Congress of Phonetic Sciences*, 913-316.
- Werker, J. F. (1984). Phonemic and phonetic factors in adult corss-language speech perception. *Journal of the Acousical Society of America*, 75(6):1866-1878.
- Werker, J. F. and Tees, R. C. 2002. Cross-language speech perception: Evidence for Perceptual reorganization during the first year of life. *Infant Behavior & Development* 25: 121-133.
- Yang, B. 1996. A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, 24: 245-161.