Studies in Second Language Learning and Teaching
Department of English Studies, Faculty of Pedagogy and Fine Arts, Adam Mickiewicz University, Kalisz
SSLLT 3 (2). 261-292
http://www.ssllt.amu.edu.pl

Discrimination of Arabic contrasts by American learners

Mahmoud S. Al Mahmoud
Imam University, Riyadh, Saudi Arabia
mssaam@hotmail.com

Abstract

This article reports on second language perception of non-native contrasts. The study specifically tests the perceptual assimilation model (PAM) by examining American learners' ability to discriminate Arabic contrasts. Twenty two native American speakers enrolled in a university level Arabic language program took part in a forced choice AXB discrimination task. Results of the study provide partial evidence for PAM. Only two-category contrasts followed straightforwardly from PAM; discrimination results of category-goodness difference and both uncategorizable contrasts yielded partial support, while results of uncategorized versus categorized contrast discrimination provided counter-evidence to PAM.

Keywords: discrimination of sound contrasts, Arabic consonant contrasts, perceptual assimilation model

It has long been recognized that second language (L2) speakers perceive and produce non-native sounds under the influence of their first language (L1) sound systems. Polivanov (1931) and Trubetzkoy (1939) were among the earliest to view the native language phonological system as a filter through which L2 sounds are perceived and produced. Lado (1957) formalized this approach in his contrastive analysis hypothesis, which predicts L2 learning difficulty only in sounds and structures different from those in the L1. The poor discrimination of English /l/

and /r/ by Japanese speakers of English (e.g., Best & Strange, 1992; Yamada, 1995) is a classic example of how learners show perceptual difficulty with new contrasts.

However, it soon became clear that a simple contrastive comparison between L1 and L2 sounds is inadequate (Eckman, 1987; Gass & Selinker, 2001) as it falls short of explaining why some L1-L2 differences are easy to learn and why, on the other hand, some L1-L2 similarities still pose a great amount of difficulty for L2 learners (Towell & Hawkins, 1994). Flege (1987) showed, for example, how English speakers of French produce L1-similar French /u/ with second formant values significantly different from native French norms while their production of L1-different French /y/ is native-like (similar findings in Flege, 1995; Flege, Bohn, & Jang, 1997). In perceiving non-native segments, English listeners demonstrate excellent discrimination of novel Zulu click contrasts which do not exist phonemically or phonetically in English (Best, McRoberts, & Sithole, 1988), but poor discrimination of Zulu plosive versus implosive bilabial stops /b/-/ɓ/, even though /b/ exists in English.

Consequently, several models of speech perception and production have been formulated to better predict areas of difficulty in learning a foreign language sound system. Flege's (1987, 1995) psychoacoustic speech learning model claims that the ability to perceive and produce L2 sounds in a native-like manner is contingent upon the establishment of L2-separate abstract phonetic categories,[1] the forming of which is regulated by the similarity relationship between L1 and L2 sounds: "The greater the perceived distance of an L2 sound from its closest L1 sound, the more likely it is that a separate category will be established for the L2 sound" (Flege, 1987, p. 264). The speech learning model predicts, contra a classical contrastive analysis, that *similar* non-native sounds will be problematic for learners, whereas *new* sounds, different from any existing L1 sound, will be easily acquired. The perceptual assimilation model (PAM) advanced by Best et al. (1988) and Best (1995) claims that contrast discrimination varies gradiently as a function of the similarity between native and non-native sounds.

In this study I test the predictions of PAM as they relate to the perception of Arabic consonants by American English (AE) speakers. The paper is organized as follows. First, I lay out some of the basic tenets and claims of the model. After discussing the relationship between Arabic and English phonemes, the predictions of PAM with regard to 9 Arabic consonant contrasts are outlined and their testing in a perceptual experiment involving AE learners

[1] Flege (1995, p. 239) defines *phonetic categories* as "language specific aspects of speech sounds [which] are specified in long term memory representations."

of Arabic is reported. An appraisal of the results in light of the model follows. The paper concludes with a summary of the main findings.

The Perceptual Assimilation Model

The perceptual assimilation model assumes that listeners have direct access to relevant information in the speech signal without the need for intermediate representations. It maintains that auditorily listeners perceive speech sounds in terms of the gestures necessary to articulate them (Best, 1995; Fowler, 1989). The model has undergone a number of revisions since it was first developed to account for English speakers' rather unexpected excellent discrimination of non-native Zulu clicks in Best et al. (1988). Most of the discussion that ensues is mainly based on Best (1995).

The perceptual assimilation model seeks to explain the gradient success listeners demonstrate in perceiving and discriminating non-native segments and contrasts. According to PAM, the ability to discriminate various non-native contrasts follows from implicit or explicit assimilation of each contrasting segment to a native category. A regulating factor in determining assimilability is the degree of phonetic closeness or discrepancy which native and non-native sounds share, as stated in Best (1995):

> The fundamental premise of the perceptual assimilation model of cross-language speech perception is that non-native segments, nonetheless, tend to be perceived according to their similarities to, and discrepancies from, the native segmental constellations that are in closest proximity to them in native phonological space. (p. 139)

The degree of similarity between native and non-native phonemes is defined by "the spatial proximity of constriction locations and active articulators and by similarities in constriction degree and gestural phasing" (p. 194). A non-native phoneme is more likely to be assimilated to a native one when it is perceived as a good exemplar of its native equivalent. A non-native phoneme can be assimilated as (a) an existing native speech sound perceived as being identical, acceptable or deviant exemplar of the native category, (b) a speech sound within the phonological space, but not representative of any particular native category, or can be heard as (c) a non-speech sound and therefore outside of the native phonological space.

Given that contrast discriminability in PAM is predictable from the assimilation of each segment in the contrast, the different combinations of (a)-(c) result in the following pairwise assimilation types, each with its predicted level of discriminatory accuracy (Best, 1995, p. 195):

1. Two-category assimilation (TC type): Each non-native segment is assimilated to a different native category, and discrimination is expected to be excellent. An example of this type is the Tigrinya ejective contrast between the voiceless alveolar /t'/ and bilabial /p'/, assimilated to the English alveolar-bilabial contrast /t/-/p/, respectively (Best, 1993).
2. Category-goodness difference (CG type): Both non-native sounds are assimilated to the same native category, but they differ in discrepancy from the native "ideal" (e.g., one is acceptable, the other is deviant). Discrimination is expected to be moderate to very good, depending on the magnitude of difference in category goodness for each of the non-native sounds. The voiceless ejective and non-ejective velars /k'/-/k/ in Zulu are likely to be treated as voiceless velar /k/ in English, with Zulu /k/ as the good exemplar and ejective /k'/ as the deviant (Best, 1994).
3. Single-category assimilation (SC type): Both non-native sounds are assimilated to the same native category, but are equally discrepant from the native "ideal;" that is, both are equally acceptable or both equally deviant. Discrimination is expected to be poor (although it may be somewhat above chance level). Best (1994) gives the Thompson Salish contrast in ejective velar /k'/ and uvular /q'/ as a SC assimilation type where both sounds are likely to be perceived as deviant exemplars of prototypical English velar /k/.
4. Both uncategorizable (UU type): Both non-native sounds fall within phonetic space, but outside of any particular native category, and can vary in their discriminability as uncategorizable speech sounds. Discrimination is expected to range from poor to very good, depending upon their proximity to each other and to native categories within native phonological space. The well-known difficulty in distinguishing the English /l/-/r/ contrast by Japanese speakers reported in Best and Strange (1992) and Yamada and Tohkura (1992) can be an example of this type; neither liquid is assimilated to a good Japanese equivalent.
5. Uncategorized versus categorized (UC type): One non-native sound is assimilated to a native category, and the other falls in phonetic space, outside native categories. Discrimination is expected to be very good. The English /r/-/w/ distinction for Japanese listeners fits this type where, unlike English /w/, which is assimilated as Japanese /w/, English /r/ is not assimilable to any Japanese category (Guion, Flege, Yamada, & Pruitt, 2000).
6. Nonassimilable (NA type): Both non-native categories fall outside of speech domain being heard as non-speech sounds, and the pair can vary in their discriminability as non-speech sounds; discrimination is

expected to be good to very good. English speakers' discrimination of the Zulu clicks which, for an English listener, do not resemble any speech sound, falls into this category of assimilation (Best et al., 1988).

The assimilation types just listed describe how non-native contrasts can vary in their discriminability as a function of the gestural (dis)similarity each member of the contrast bears to a native segment. Interference from the native phonological system can lead to a clear perceptual benefit in distinguishing contrasts when the contrast involves two separate, clearly defined categories in the native language (TC). It can cause, however, a perceptual detriment when the contrast elements correspond to one native phoneme (SC). Or it can be irrelevant as in the NA type. The assimilation types TC, CG, and SC refer to non-native contrasts which can be assimilated to the native phonological space, and therefore follow from degree-of-similarity type (a). UU is concerned with contrasts that fall outside of the native domain but are still speech-like as stated in (b). UC is characterized by both (a) and (b), whereas a contrast of the NA type is solely based on (c). The different predictions made by PAM on the three assimilation types that occur within the native phonological domain suggest the following hierarchy of difficulty in the discrimination of non-native contrasts, where TC is easiest: TC > CG > SC.

Several hypotheses of PAM have been tested. In one of the earliest studies that have led to the development of PAM, Best et al. (1988) examined the discrimination of Zulu (a Bantu language) place and voicing click contrasts by native speakers of American English. Despite the fact that clicks do not exist phonemically or phonetically in English, an AXB task showed that English listeners had little difficulty in discriminating the Zulu click consonant contrasts, even when amplitude differences (a crucial acoustic cue for clicks) were leveled in a subsequent experiment. In both experiments adult English native speakers' success on the discrimination tasks, which amounted to 80% correctness in the natural condition and 78% in the modified one, is predictable as an NA assimilation type. Best et al. (1988) argue that since clicks are gesturally very distant from, and therefore cannot assimilate to, any English phoneme as their articulation involves ingressive suction followed by loud release, they were most likely perceived by English listeners as non-speech sounds that do not belong to any native category, in which case listeners were more reliant on the auditory and phonetic properties of clicks, which make them highly discriminable.

More recently, Best, McRoberts, and Goodell (2001) evaluated the predictions of PAM on the TC, CG and SC assimilation types by examining English listeners' perception of Zulu and Tigrinya consonant contrasts. For native English listeners the contrasts between the voiceless and voiced Zulu lateral fricatives /ɬ/-/ɮ/ as well as the ejective bilabial and alveolar stops in Tigrinya /p'/-/t'/ were expected to fit a TC assimilation type,[2] in which case Zulu fricatives would be equated with English voiceless versus voiced apical fricatives and Tigrinya /p'/-/t'/ with non-ejective bilabial versus alveolar stops. Moreover, English listeners were predicted to show a CG type in their assimilation of Zulu voiceless aspirated and ejective velar stops /kʰ/-/k'/ to American English /k/. Finally, a discrimination pattern consistent with SC was believed to emerge in the Zulu contrast between plosive and implosive voiced bilabial stops /b/-/ɓ/, given that both sounds would most likely be perceived as English /b/. An AXB test showed all English listeners had more difficulty discriminating the Zulu aspirated-ejective contrast in velar stops than the voiceless-voiced contrast in lateral fricatives (89.4% vs. 95% correct discrimination). Even worse, though not generalizeable to all subjects, was their ability to distinguish the plosive-implosive contrast in bilabial stops (65.9%).

The advantage of Best et al.'s study is that it clearly shows how the PAM predictions for these three assimilation types were upheld within the same language and by the same group of listeners, eliminating the possibility that the observed pattern of discrimination was the result of different languages and/or subject populations. A second experiment using the same AXB paradigm revealed that English listeners' discrimination of the Tigrinya bilabial-alveolar place contrast was consistent with the Zulu voicing contrast: both TC types. In general, English listeners' performance on the non-native Zulu and Tigrinya contrasts followed straightforwardly from the patterns outlined in 1-3, along with their predicted discrimination levels, confirming the TC > CG > SC discriminability ranking suggested by PAM.

Further evidence for PAM is found in Guion et al.'s (2000) assessment of the UU and UC types. Guion et al. hypothesize that the distinction between English /r/ and /l/ would fall into the UU type (both uncategorizable) for Japanese listeners, given that the only liquid Japanese has is the alveolar retroflex tap [ɽ], which is phonetically deviant from its English equivalent. Further, two more contrasts, English /r/-/w/ and /s/-/θ/, were evaluated and assumed to follow the UC assimilation type based on the fact that Japanese /w/ and /s/ are good exemplars of their English equivalents. The findings of Guion et al. are generally in line with PAM; Japanese listeners were more successful in distinguishing English /r/ from /w/ in contrast to the poorly discriminated liquids. An exception was listeners'

[2] In determining what sound contrasts belong to which assimilation types, gestural similarities and differences among English, Zulu, and Tigrinya sounds were discussed in Best et al. (2001, pp. 778-779).

discrimination of /s/-/θ/, which contrary to a UC type prediction, was rather difficult, arguably due to the phonetic proximity of the segments in the pair.

An earlier study by Best and Strange (1992) involving Japanese speakers yields results that are different from those of Guion et al. (2000), yet consistent with the PAM predictions. In Best and Strange (1992), Japanese listeners perceived and discriminated English /r/-/l/ and /r/-/w/ as SC and CG contrasts, respectively. However, it should be noted that in both studies the contrast between /r/ and /w/, classified as either UC or CG, was consistently differentiated better than the liquid contrast, classified as UU or SC. Other cross-language speech perception studies also report findings that are supportive of PAM (e.g., Tigrinya ejective contrasts for English speakers in Best, 1990; English /r/-/w for French speakers in Hallé, Best, & Levitt, 1999; English final obstruent voicing contrasts for Malay speakers in Pilus, 2002; and Hindi dental-retroflex for English speakers in Polka, 1991).

Given the need for further evaluation of PAM on novel languages as asserted by Best (1995, p. 198), the present study tests AE listeners' ability to successfully discriminate a number of diverse contrasts in L2 Arabic, a language not yet examined in light of PAM.

Language Background

Consonant Inventories of Arabic and English

Modern Standard Arabic (or Arabic) is the official language of instruction, media and science in the Arab world, and the target language for the majority of L2 Arabic learners in the US and elsewhere. For comparison purposes, the consonant inventories of Arabic and English are both outlined.[3]

The phonemic inventory of Arabic consonants is represented by a wide range of sounds, among which are the universally less frequent uvular, pharyngeal and glottal places of articulation. There are 28 consonant phonemes as shown in Table 1, adopted from Al-Ani (1970) and Ingham (1971) with modification. English, on the other hand, has 24 consonant phonemes as illustrated in Table 2.

[3] The vowels of Modern Standard Arabic include /a/, /i/, /u/ and their long counterparts /aː/, /iː/ and /uː/.

Table 1 Consonant inventory of Arabic

	Stop	Affricate	Nasal	Fricative	Trill	Approximant
Bilabial	b		m			w
Labiodental				f		
Interdental				θ ð ð̲		
Dental	t d t̲ d̲					
Alveolar			n	s z s̲	r	l
Palatoalveolar		dʒ		ʃ		
Palatal						j
Velar	k					
Uvular	q			χ ɣ		
Pharyngeal				ħ ʕ		
Glottal	ʔ			h		

Note. Underlining represents emphatic consonants. Phonemes to the left in pairs are voiceless.

Table 2 Consonant inventory of English (Ladefoged, 2001)

	Stop	Affricate	Nasal	Fricative	Central approximant	Lateral approximant
Bilabial	p b		m		w	
Labiodental				f v		
Dental				θ ð		
Alveolar	t d		n	s z		l
Palatoalveolar		tʃ dʒ		ʃ ʒ		
Retroflex					ɹ	
Palatal					j	
Velar	k g		ŋ			
Glottal				h		

Note. Phonemes to the left in pairs are voiceless.

Comparing the inventories of the two languages, there are a number of phonemes that are found in Arabic only but not in AE. These include the (pharyngealized) emphatics /ð̲/, /s̲/, /d̲/, /t̲/; the uvulars /q/, /χ/, /ɣ/; the pharyngeals /ħ/, /ʕ/ and the glottal /ʔ/. The phoneme /r/ is realized as the retroflex approximant [ɹ] by many American speakers of English (Ladefoged, 2001), but it is an alveolar (or sometimes dental) trill [r] for the majority of Arabic speakers (Al-Ani, 1970; Amayreh, 2003). Much controversy surrounds the pharyngeal sound /ʕ/. Although traditionally described as a voiced fricative, Al-Ani (1970, p. 62) has concluded on acoustical basis that it is best characterized as a voiceless stop visible on the spectrogram in the form of a 40-50 ms burst accompanied by noise. Similarly, Thelwall (1990) argues that the pharyngeal phoneme /ʕ/ is realized in many Arabic dialects as a pharyngealized glottal stop. However, Ladefoged and Maddieson (1996) consider /ʕ/ an epiglottal fricative.

The rest of the Arabic consonant phonemes exist in English although their phonetic realizations in each language may differ slightly. Because /t/, /d/, /θ/, /ð/, /t̲/, /ð̲/, /h/, /k/, /q/, /ħ/, /χ/, and /ɣ/ are the only subset of Arabic consonants examined in this study, the discussion below will be limited to how each of these consonants is realized phonetically, if any, in each language; their similarity to and discrepancies from English should aid in predicting how native speakers of English would assimilate them, when possible, to their native phonological categories.

Similarity Between Arabic and English Consonants

In determining if and how Arabic consonant contrasts would be perceived and assimilated to the English sound system the first step is to examine the native phonetic realization of each individual phoneme in each language, before even discussing what contrasts are to be tested. While this study acknowledges the similarity measure set forth by Best (1995), in which gestural specifications necessary for the proper articulation of each segment have to be assessed for the language pair in question, a detailed gestural account of Arabic consonants is not possible for a small scale study. The comparison below is based on the articulatory and acoustic measurements of Arabic consonants provided by Al-Ani (1970) as well as on other descriptive work in Arabic linguistics (e.g., Al-Karouri, 1996; Bateson, 1967). For English most of the discussion is based on Ladefoged (2001, 2005).

A first glance at the Arabic consonants examined in this study renders the following unequivocal classification, according to their phonemic status in the English inventory:

- those that exist as separate phonemes in English: /t/, /d/, /θ/, /ð/ /k/, /h/;
- those that do not exist in English as phonemes: /t̲/, /ð̲/, /q/, /χ/, /ɣ/ and /ħ/.

Among those that exist in English, Arabic /t/ and /d/ are often produced with the tip of the tongue touching the posterior part of the front teeth (i.e., the dental place of articulation; Al-Ani, 1970; Bateson, 1967), but according to Ladefoged (2001) their articulation in English involves a fully alveolar gesture. The stop /t/ is aspirated initially in both languages, but only in Arabic is it often released in final position. When aspirated the burst intensity for /t/ is concentrated in the 3 kHz range for Arabic (Al-Ani, 1970, p. 45), but in English the energy is in the higher 3-5 kHz (Ladefoged, 2005, p. 53) or 3.9 kHz (O'Shagaussey, 2000, p. 66) frequencies.

In Arabic the fricatives /θ/ and /ð/ are always described as being interdental sounds, produced with the tip of the tongue placed between the upper and lower front teeth (Al-Ani, 1970; Al-Karouri, 1996; Bateson, 1967). In English, however, many native speakers produce them with the tongue tip placed at the back of the upper front incisors (dental), although for some American speakers they are produced interdentally (Ladefoged, 2005, pp. 119-120). For the velar /k/, the most common realization in both languages is the voiceless aspirated stop made by bringing the tongue into contact with the velum although a palatalized or fronted variant occurs in Arabic when next to high front vowels. The energy associated with the velar aspiration has its intensity in the 2-3 kHz range for both languages (Al-Ani, 1970, p. 32; Ladefoged, 2005, p. 52). The stop /k/ is often released initially in English, but in Arabic, like other plosives, it is released both in initial and final positions. The voiceless fricative /h/ is characterized by random noise that is "caused by the movement of the air across the edges of the open vocal folds and other surfaces of the vocal tract" (Ladefoged, 2005, p. 58). The noise is most intense in the 3 kHz level for English, and is slightly lower for Arabic, around 2.7 kHz.

Among the phonemes that do not exist in English are the Arabic (pharyngealized) emphatics /ṯ/ and /ð̠/, for which the most common allophones are a voiceless unaspirated post-dental stop and a voiced interdental fricative. The emphatic /ṯ/ is the pharyngealized counterpart of /t/, but unlike /t/ it is often articulated with the tongue tip positioned further back, is unaspirated, and has a shorter burst duration (20-30 ms vs. 40-60 ms) concentrated usually at a slightly lower frequency range (1500-2400 Hz) than that of /t/ (1600-2700 Hz). When next to /a/, the second formant values are approximately 1500-1550 Hz for /t/, but drop to the 1150-1250 Hz range for /ṯ/. The emphatic /ð̠/, on the other hand, is more similar to its non-pharyngealized counterpart /ð/: Both are produced interdentally and have the same noise duration (100-160 ms) with similar first and third formant resonances at around 275 and 2350 Hz, respectively. However, the second formant noise resonances for /ð̠/ are between 900-1000 Hz, compared to /ð/'s 1500 Hz. A dampening effect is also observed for the pharyngealized fricative /ð̠/, but not for /ð/, evident in the second formant onset values for an adjacent /a/ (1150-1200 Hz vs. 1500-1600 Hz).

The rest of the phonemes which do not exist in English consist of the gutturals, which are in fact entirely new to English. The uvular /q/ surfaces as a voiceless unaspirated stop acoustically resembling a typical stop accompanied by a robust release burst spectrally visible around 3000 Hz followed by a period of silence with a duration of 30-40 ms. The most common allophone of /χ/ is a voiceless uvular fricative which appears spectrally as random noise averaging 100-160 ms. When next to /a/, the fricative noise condenses around the 1500 Hz baseline and is almost undetectable below 1000 Hz. A fronted velarized variant also exists when next to the high front vowels /i/ or /iː/. The fricative /ɣ/ is the voiced counterpart of /χ/ and is realized either as uvular with /a/, /aː/, /u/ and /uː/, or as a fronted velar sound when next to the vowels /i/ and /iː/. This sound is similar to /χ/ in duration, but has its noise concentrated at a lower frequency range (1300 HZ). The pharyngeal /ħ/ is realized as a voiceless constricted fricative with an average duration of 100-150 ms. It is set off from /h/ by its somewhat constricted articulation which shows more intense noise often at lower third formant frequencies (Klatt & Stevens, 1969). The constriction involved in producing /ħ/ is formed by bringing "the dorsum of the tongue against the posterior wall of the pharynx where the movements of the pharyngeal muscles play an important role" (Al-Ani, 1970, p. 60). When intervocalic, /ħ/, as /h/, tends to be voiced. In general, as Peterson and Shoup (as cited in Al-Ani, 1970) say, pharyngeal (and glottal) sounds are distinguished from other sounds by their "vertical" places of articulation (i.e., along the back of the throat from the palate to the glottis), as opposed to horizontal places of articulation which extend from the lips to the uvula.

Testing PAM

Determining L2 Arabic Contrasts and their Assimilability to English

Table 3 includes 9 Arabic consonant contrasts that were examined.[4] The pairs in the table are distinguished either by voicing (1-3), manner of articulation (4-6), or place of articulation (7-9). The manner contrasts are differentiated either by the feature [emphatic] (pharyngealized) as in 4-5 or [continuant] as in 6; the place contrasts differentiate the pharyngeal /ħ/ either from the glottal /h/ as in 7 or from the uvular /χ/ as in 9, and the velar /k/ from the uvular /q/ as in 8.

[4] Initially some 20 pairs were tested, but since all involved the same contrasts tested here (i.e., the same place, manner or voicing), only a subset is reported on.

Table 3 Tested Arabic consonant contrasts

No.	Contrast	Type of contrast	Description
1	/t/-/d/	voicing	voiceless vs. voiced
2	/θ/-/ð/	voicing	voiceless vs. voiced
3	/χ/-/ɣ/	voicing	voiceless vs. voiced
4	/t̠/-/t/	manner	emphatic vs. non-emphatic
5	/ð̠/-/ð/	manner	emphatic vs. non-emphatic
6	/q/-/χ/	manner	plosive vs. continuant
7	/ħ/-/h/	place	pharyngeal vs. glottal
8	/k/-/q/	place	velar vs. uvular
9	/χ/-/ħ/	place	uvular vs. pharyngeal

Based on the previous discussion of the similarities and discrepancies between Arabic and English in the realization of these sounds, I make the predictions shown in Table 4 with regard to the plausible assimilability of each Arabic consonant to its closest English counterpart. Given that no two sounds are exactly identical in different languages (as is well known, but see e.g., Rochet, 1995), the assimilability shown in Table 4 is both approximate and predictive. It is hypothesized on the basis of the acoustic and articulatory phonetic subtleties between Arabic and English discussed in the previous section, as well as on the author's own observations (teacher's observations) of AE speakers' substitution patterns when learning (perceiving and producing) these sounds of Arabic as a second language. In learning Arabic, the majority of English speakers show a clear and consistent tendency to replace the emphatics /t̠/ and /ð̠/ with their non-emphatic counterparts, namely /t/ and /ð/. In addition, though less consistently, they substitute the velar /k/ for the uvular /q/, and the glottal /h/ for the pharyngeal /ħ/.[5]

Table 4 Assimilability of Arabic consonants to English

Arabic	Assimilable to English	Goodness of fit
/θ/	/θ/	excellent exemplar
/ð/	/ð/	excellent exemplar
/k/	/k/	excellent exemplar
/h/	/h/	excellent exemplar
/t/	/t/	good exemplar
/d/	/d/	good exemplar
/t̠/	/t/	poor exemplar
/ð̠/	/ð/	poor exemplar
/q/	/k/	very poor exemplar
/ħ/	/h/	very poor exemplar
/χ/	/h/	extremely poor exemplar
/ɣ/	/g/	extremely poor exemplar

[5] Sometimes /g/ and /χ/ are substituted for /q/ and /ħ/, respectively.

Perceptually, Arabic /t/ and /d/ are expected to represent good exemplars of English /t/ and /d/, since in each language only slight discrepancies exist in the canonical articulation of both. Still, the two sounds are predicted to assimilate to two separate corresponding categories in English. The sounds /θ/, /ð/, /k/ and /h/ are predicted to be perceived by AE listeners as excellent instances of their English equivalents (closest to being perceived as identical), as there are minimal to no differences between their realization in each language. The emphatics /t̠/ and /ð̠/ are expected to sound close to non-emphatic /t/ and /ð/, just as the pharyngeal /ħ/ most likely resembles its unconstricted glottalized counterpart, /h/.[6]

The Arabic voiceless uvular /q/ is predicted to be heard by AE listeners as a very poor exemplar of English /k/. This prediction is based not only on the relative acoustical similarity discussed earlier, but also on the fact that in the speech of many English learners of Arabic /q/ is fronted to /k/, as mentioned above. Even in some non-standard Arabic dialects, the voiceless uvular /q/ is almost always substituted for the voiced velar /g/ as in Gulf Arabic (e.g., spoken in Saudi Arabia, Kuwait, UAE, Qatar, etc.). The uvular fricatives /χ/ and /ɣ/, although in rare cases produced as /k/ and /g/ by beginner AE learners of Arabic, bear no clear resemblance to any particular category of the English phonological system and are deemed, therefore, perceptually non-assimilable to any particular category in English; nonetheless, the guttural gesture (which subsumes uvulars) does exist in English in some allophonic environments and must therefore be included within the native phonological space of English (Best, 1995).

Stating the PAM Predictions

Considering the Arabic contrasts stated in Table 3 and based on the predicted assimilability of each contrast member to its closest English equivalent in Table 4, four assimilation types associated with PAM can be identified in the classification of these contrasts.

The first assimilation type that can be discerned in Table 3 is the TC type in which both non-native phonemes assimilate to two separate categories in the native language. The Arabic voicing contrasts /t/-/d/ and /θ/-/ð/ seem to fall well into this type, as every member in these contrasts is assimilable to a single English phoneme. In other words, the two contrasts exist both in Arabic

[6] Note also that to represent the emphatics and the pharyngeal, IPA uses the symbols /t̠/, /ð̠/ and /ħ/ respectively, which are representationally isomorphic to non-emphatic /t/, /ð/ and glottal /h/, perhaps suggestive of their phonetic closeness.

and English. The perceptual assimilation model predicts contrast discrimination in this case to be highly efficient. The second deducible type is CG, according to which the two non-native sounds are perceived by the listener as instances of one native phoneme, but they vary in their goodness of fit: One is acceptable, the other is not. I consider the contrasts between emphatics and non-emphatics in /t̠/-/t/ and /ð̠/-/ð/ as examples of the CG assimilation type for which discriminability would be "moderate to very good," as stated in PAM.

On the other hand, the pharyngeal-glottal /ħ/-/h/ and velar-uvular /k/-/q/ contrasts appear to fit the UC assimilation type. In either contrast, one member is clearly equatable with English: Arabic /h/ and /k/ are almost identical to English /h/ and /k/. The other members /ħ/ and /q/ fall within the English phonological space, yet lack clear correspondents, although anecdotal evidence suggests that English speaking learners of Arabic poorly realize them as /h/ and /k/. Accordingly, it is possible to treat the /ħ/-/h/ and /k/-/q/ contrasts as another example of CG assimilation, with Arabic /h/ and /k/ being the acceptable and /ħ/ and /q/ the deviant exemplars of English /h/ and /k/.

However, given the less frequent phonemic substitution of /h/ and /k/ for /ħ/ and /q/ in the speech of American learners (cf. /t/-/ð/ for /t̠/-/ð̠/), I believe it is more fitting for these two pairs to be classified as UC rather than CG contrasts. This ambiguity in the categorization of /ħ/-/h/ and /k/-/q/ pairs is quite reminiscent of the English /r/-/w/ contrast for Japanese speakers, categorized in Best and Strange (1992) as a CG type, but as a UC type by Guion et al. (2000). Needless to say, the suggestion made in this study is more along the lines of Guion et al. As far as discriminability is concerned, both assimilation types make similar predictions associated with the contrasts: very good contrast discrimination in UC, and moderate to very good in CG.

The last three types of contrasts in Table 3, namely the voicing contrast in /χ/-/ɣ/, the uvular versus pharyngeal in /χ/-/ħ/ and the plosive versus continuant in /q/-/χ/, show a UU assimilation type. The contrast discriminability which PAM predicts for this type of assimilation depends on how well each segment is assimilable, if any, to the native language as well as on the phonetic closeness between the contrast members, and can range, therefore, from poor to very good. In each of these contrasts neither consonant clearly assimilates to a specific category in the English language. This is especially true for the uvulars /χ/ and /ɣ/, which do not correspond to any English sound. The sounds /ħ/ and /q/, on the other hand, can be equated with /h/ and /k/, although not commonly, as mentioned before. Thus, based on their assimilability to the learner's native language in this study, it is hypothesized that within the UU assimilation type the contrast /χ/-/ɣ/ would be less discriminable than /χ/-/ħ/ or /q/-/χ/.

Table 5 sums up the contrasts tested in this study, their assimilation types, and the predicted discrimination level associated with each type. The perceptual assimilation model predicts that the TC contrasts /t/-/d/ and /θ/-/ð/ will be better discriminated than any of the other contrasts in this study. Another prediction that is testable is the increased indiscriminability of /χ/-/ɣ/ relative to the other two UU contrasts, /χ/-/ħ/ and /q/-/χ/.

Table 5 PAM's predictions for the Arabic contrasts

Contrast	Assimilation type	Predicted discriminability
/t/-/d/	TC	excellent
/θ/-/ð/	TC	excellent
/t̠/-/t/	CG	moderate to very good
/ð̠/-/ð/	CG	moderate to very good
/ħ/-/h/	UC	very good
/k/-/q/	UC	very good
/χ/-/ɣ/	UU	poor
/χ/-/ħ/	UU	very good
/q/-/χ/	UU	very good

Method

The relevant predictions of PAM stated above were tested in an auditory perception experiment carried out to assess AE learners' perception of the various Arabic phonemic contrasts, summarized in Table 3. A perception experiment, as opposed to a production one, is believed to provide a better measure of acquisition or learnability since it involves less conscious knowledge of what is being learned (Archibald, 1998; Larsen-Hall, 2004).

Participants

Twenty two American learners of Arabic participated in the perception experiment. The participants were all native speakers of American English and their average age was 22. They were year 1 students at a US college enrolled in the university's Arabic program at the time of the study. The majority of the participants had knowledge of a third language, mainly Spanish or French. It was hard to find participants who spoke English and Arabic only since most had been exposed to a third language in high school. Werker (1986) shows that trilinguals discriminate non-native contrasts no better than bilinguals do.

In other words, it is believed that knowledge of a third language does not significantly affect contrast discrimination.

Participants who received any specialized training in pronunciation or phonetics and those whose length of residence in an Arabic-speaking country exceeded 6 months were excluded from the study. In addition, participants for whom Arabic was considered a heritage language as well as those reporting any hearing difficulties were precluded.[7] All subjects were compensated $10 each for their participation in the experiment. According to self-report, no participant had any hearing difficulties.

Materials and Task

Initially 20 phonemic consonant contrasts were tested (see the Appendix). However, only those mentioned in Table 3 are reported on here. In this study, the term *phoneme* refers to the allophone that is used in isolation and is usually more common than other variants, and which Daniel Jones (1967, p. 8) considers as the "principal member" or the "norm" of the phoneme. The pairing of phonemic contrasts in Table 3 depends on potentially confusable consonants that share major place-of-articulation features (coronal, dorsal, guttural). In addition to the nine contrasts in Table 3, four more contrasts were used as distracters: /m/-/n/, /r/-/l/, /w/-/j/, and /dʒ/-/ʃ/.

With an AXB discrimination task, four test items (AAB, ABB, BAA, BBA) were generated for each of the 24 contrasts yielding a total of 96 randomly ordered test trials. Every test trial was a triad consisting of three disyllabic nonsense words following the template ʔaːˈCaː (where C is consonant), with stress being placed systematically on the second syllable, for example, ʔaːˈt̠aː-ʔaːˈt̠aː-ʔaːˈtaː. Thus a total of 288 test words were used in the discrimination task (96 x 3 = 288). Compared to syllable onsets or codas, the intervocalic position was chosen for its ideal environment in the perception of consonants, especially stops (Wright, 1995, p. 35). Also, word-final stop contrasts can be difficult for English listeners to detect due to the fact that stops tend to be unreleased in that position (Selkirk, 1982). The use of nonsense test tokens, as opposed to real words, was intended to minimize any effects word frequency and familiarity may have on L2 learners' discriminative ability.

A native speaker of Arabic trained in linguistics (the present author) produced the test words, which were transcribed in IPA to ensure a more ac-

[7] Initially 24 participants were tested; however, 2 were excluded from the study: one for whom Iraqi Arabic turned out to be a heritage language (parents' native language), and the other had stayed in an Arabic-speaking country (Egypt) for over a year and was married to an Arab.

curate pronunciation. The stimuli were digitally recorded in a sound treated lab using Audacity (Audacity Team, 2008) recording and editing software (version 1.2.4) and a clip-on PRO 7 Electret condenser microphone on a Windows Vista Dell 1420 laptop computer.

Procedure

An AXB forced-choice discrimination paradigm was used to elicit American learners' perception of the different Arabic contrasts in this study. An AXB discrimination task provides a reference point (i.e., X) against which the similarity of stimuli can be gauged by listeners, as opposed to a simple AX discrimination task where listeners may base their same/different responses on non-linguistic factors (Beddor & Gottfried, 1995).

In a quiet library room setting, aural stimuli were presented randomly over headphones (Koss R80) to each participant individually. Each test trial proceeded as follows. The participant listened to all three tokens in each triad (e.g., ʔaːˈt̠aː-ʔaːˈt̠aː-ʔaːˈtaː) and had to indicate on an answer sheet provided whether the first or third word was the same as the second. An inter-stimulus interval (ISI) of 1000 ms followed each token. Longer ISI is believed to encourage phonemic rather than phonetic perception of non-native contrasts (Werker & Logan, 1985). A 3000 ms inter-trial interval separated each trial from the following one. The 96 test trials were administered over two sessions (48 trials each). In order to ensure that each subject understood the procedure, a 3-item practice test was administered to each participant prior to the experiment. For each participant, the experiment lasted an average of 20 min.

Results

By Contrast

Discrimination scores from the AXB task were collapsed across all 22 subjects and pooled for each of the 9 phonemic contrasts stated in Table 3. To determine whether differences in discriminability were significant among the contrasts, discrimination scores were submitted to a repeated measures (within subjects) ANOVA, which tested for the significance of the independent variable, that is, contrast (9 levels). The results indicated a significant effect of contrast, $F(8, 168) = 63.80$, $p < .001$, $\eta_p^2 = .75$. To find out which pairwise comparisons are significant, a series of Bonferroni post-hoc tests compared discrimination scores across all 9 levels of the contrast variable. Table 6 sums up discrimination success rates for each individual contrast as well as the (in)significance of each pairwise comparison. The leftmost column presents percentages of correctly discriminated contrasts. This is depicted graphically in Figure 1. For instance, best discrimination involved the TC contrasts /t/-/d/ and /θ/-/ð/, for which performance was native or near-native. On the contrary, poorest discriminability belonged to the UU contrast /χ/-/ɣ/. Table 6 also shows whether pairwise comparisons (second column vs. top row) among the different contrasts were significant or not. For example, while the difference between /t/-/d/ and /θ/-/ð/, both TC types, was not significant, the difference between these and all other contrasts was. In other words, /t/-/d/ and /θ/-/ð/ patterned similarly in being significantly more discriminable than any other contrast. With some exceptions, it appears that discriminability, to a large extent, varied as a function of the assimilation type.

Table 6 Arabic contrast discrimination results by AE listeners

Discriminability	Contrast	/θ/-/ð/	/t̲/-/t/	/ð̲/-/ð/	/ħ/-/h/	/k/-/q/	/χ/-/ɣ/	/χ/-/ħ/	/q/-/χ/
100	/t/-/d/	ns	*	*	*	*	*	*	*
98.9	/θ/-/ð/		*	*	*	*	*	*	*
64.8	/t̲/-/t/			ns	*	ns	*	ns	*
72.7	/ð̲/-/ð/				*	ns	*	*	ns
37.5	/ħ/-/h/					*	ns	*	*
64.8	/k/-/q/						*	ns	*
36.4	/χ/-/ɣ/							*	*
54.5	/χ/-/ħ/								*
82.9	/q/-/χ/								

ns = insignificant at the level of .05.
* $p < .05$.

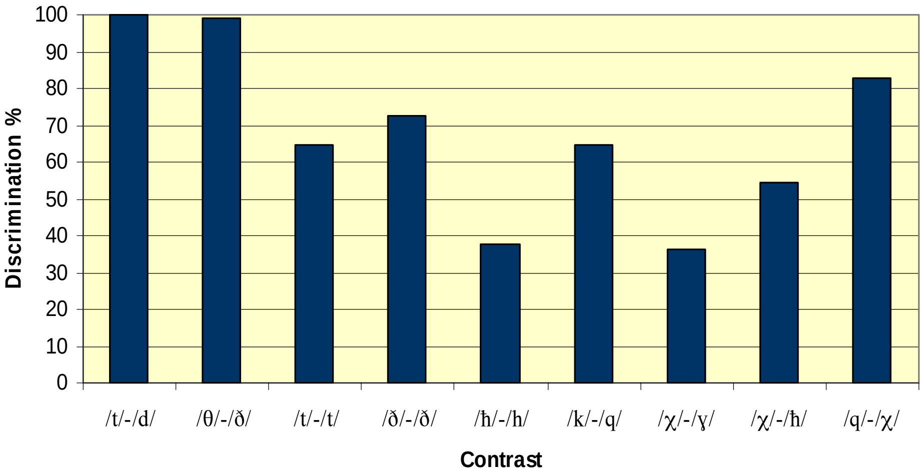

Figure 1 AXB results in the discrimination of Arabic contrasts

By Assimilation Type

Next, contrast discrimination rates were tallied up and averaged for each of the four assimilation types TC, CG, UC and UU, given the a priori classification in Table 5. A repeated measures ANOVA with the independent variable of assimilation (4 levels) revealed a significant main effect for assimilation, with $F(3, 63) = 76.59$, $p < .001$, $\eta_p^2 = .78$. Except for the pairing of UC and UU, which turned out to be insignificant, $p > .05$, post-hoc Bonferroni tests indicated significant differences among all pairwise comparisons with a high confidence level ($p < .001$). This is summarized in Table 7 and illustrated in Figure 2. Averaged discrimination rates show that TC contrasts were significantly more discriminable than other assimilation types. Similarly, CG contrasts had higher overall discriminability than UC or UU contrasts. Although the lowest discrimination rate was found in the UC type of assimilation, it did not differ significantly from that of UU.

Table 7 Discrimination results of the Arabic contrasts by assimilation type

Contrast	Discriminability	Assimilation type	CG	UC	UU
/t/-/d/, /θ/-/ð/	99.4	TC	*	*	*
/t̠/-/t/, /ð̠/-/ð/	68.7	CG		*	*
/ħ/-/h/, /k/-/q/	51	UC			ns
/χ/-/ɣ/, /χ/-/ħ/, /q/-/χ/	57.9	UU			

ns = insignificant at the level of .05.
* $p < .05$.

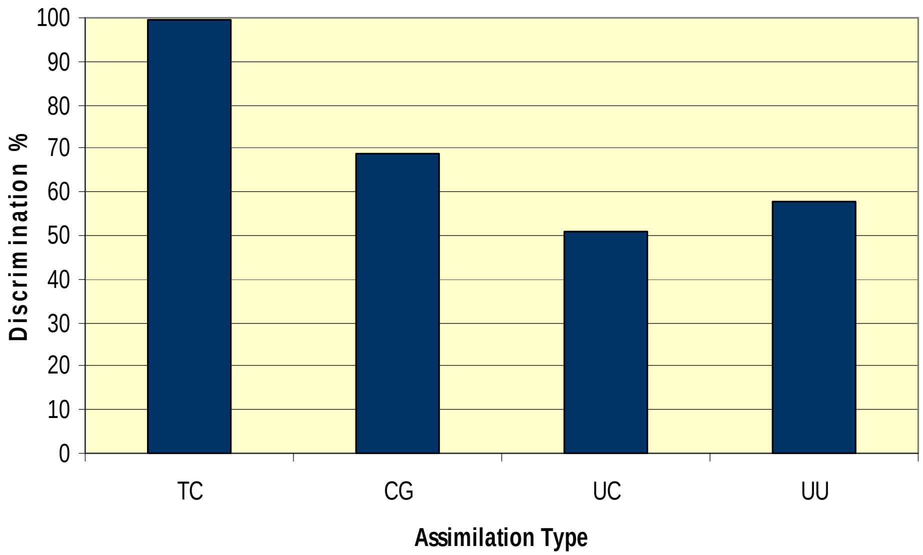

Figure 2 AE listeners' discriminability of Arabic contrasts according to assimilation type

Discussion

A cursory look at Table 6 shows that AE listeners' ability to distinguish non-native contrasts varies significantly as a function of whether or not the two contrasted segments are assimilable to the learner's native language. With few exceptions to be discussed, contrast discrimination seems to generally follow from the PAM predictions stated in Table 5, with regard to assimilation type and discriminability.

/t/-/d/

The contrast between the voiceless /t/ and voiced /d/ dental stops was predicted to follow a TC assimilation type where each segment in the contrast is assimilated to a separate native category, resulting in excellent discriminability. Looking at Table 6, this prediction is borne out; listeners discriminated correctly 100% of the time, and, as post-hoc tests show, their discrimination was significantly better than any other non-TC contrast.

/θ/-/ð/

The voicing contrast between interdental /θ/ and /ð/ was also predicted to show TC assimilation. Results confirm this; subjects were able to discriminate /θ/ from /ð/ accurately most of the time (98.9%). As in /t/-/d/, this result was significantly better than the rest of the contrasts.

TC Assimilation

As predicted by PAM, both Arabic /t/-/d/ and /θ/-/ð/ voicing contrasts behaved similarly in conforming to the TC. They both displayed the highest rate of discrimination and were significantly more distinguishable than other contrasts. The insignificant difference in discrimination between these two contrasts suggests that to American listeners /t/-/d/ was as distinctive as /θ/-/ð/. The success rate averaged for these two contrasts is 99.4%, which is significantly the greatest among other assimilation types. This pattern is in line with the findings of Best et al. (2001), for example, who demonstrated excellent (95% correct) TC contrast discrimination between Zulu voiceless-voiced fricatives /ɬ/-/ɮ/ and Tigrinya ejective bilabial-alveolar stops /p'/-/t'/ by native listeners of English.

/t̠/-/t/

It was predicted that the Arabic emphatic contrast /t̠/-/t/ would be assimilated as a CG type by native speakers of English where, depending on the goodness of fit, discriminability can range from moderate to very good. Results show that listeners had a somehow poor success rate in discriminating /t̠/ from /t/ (64.8%). In addition, discrimination here did not differ significantly from that of UC /k/-/q/ or UU /χ/-/ħ/, perhaps due to the unpredictably poor performance of the latter as well, as will be discussed later.

/ð̠/-/ð/

In contrast to /t̠/-/t/, the distinction between the Arabic emphatic and non-emphatic interdental fricatives was perceived better. Listeners more successfully distinguished /ð̠/ from /ð/ 72.7% of the time, compared to 64.8% for /t̠/-/t/, although the difference between the two emphatic contrasts did not reach significance. However, discriminability for /ð̠/-/ð/ was significantly better than other contrasts except for /k/-/q/, for which discrimination was, as mentioned earlier, counter-predictably low, and /q/-/χ/, a UC type for which PAM makes a similar prediction of very good discrimination.

CG Assimilation

The pattern of perception for the /ð̠/-/ð/emphatic contrast appears to fit the discriminatory level hypothesized for CG assimilation. Listeners' diminished ability to differentiate /t̠/-/t/ compared to /ð̠/-/ð/, however, could have resulted from /t̠/ being perceived as a better fit for /t/ than /ð̠/ for /ð/. In other words, to listeners /t̠/ was perceptually closer to /t/ than /ð̠/ to /ð/, leading to greater confusability in the pair /t̠/-/t/, and thus to lesser discriminability. When combined, discrimination averages 68.7% for the two CG contrasts. This rate is much lower than what is reported in Best et al. (2001) for English listeners who showed a much higher success rate (89.4%) in discriminating the Zulu CG voiceless aspirated and ejective velar stop contrast /k^h/-/k'/. It seems, therefore, that the predictions of PAM concerning CG assimilation are partially supported by the emphatic contrast between the interdental fricatives /ð̠/ and /ð/.

/ħ/-/h/

It was hypothesized that the pharyngeal-glottal contrast would exemplify a UC assimilation for which discrimination should be very good. This prediction did not transpire, however, as AE listeners had great difficulty in distinguishing these two sounds with only 37.5% of the contrasts being correctly discriminated. Except for /χ/-/ɣ/, discrimination of /ħ/-/h/ was significantly worse than all other contrasts, particularly UC /k/-/q/.

/k/-/q/

The velar-uvular contrast was also hypothesized to show very good discrimination, typical of UC assimilation. Although performance significantly improved on this contrast compared to the pharyngeal-glottal, still discrimination was unpredictably poor hovering around the 64.8% range.

UC Assimilation

The manner in which AE listeners discriminated the pharyngeal-glottal /ħ/-/h/ as well as the velar-uvular /k/-/q/ contrasts did not conform to the PAM predictions for UC assimilation; that is, neither contrast showed very good discrimination although for /ħ/-/h/ discriminability was poorer but considerably better for /k/-/q/. Further, discrimination of the two contrasts added up to 51%, which is rather poor and counter-predictive. It is possible that the phonetic proximity of the segments, as evidenced by the phonemic substitution of /ħ/ and /q/ for /h/ and /k/, respectively, has led to the poor discrimination of these UC contrasts. Guion et al. (2000), for example, report similar results for Japanese listeners, who, although successful in discriminating English /r/-/w/, were not able to differentiate the English UC contrast /s/-/θ/ accurately.

/χ/-/ɣ/

The voiceless-voiced uvular contrast /χ/-/ɣ/ was predicted to be a UU assimilation, in which, according to PAM, discriminability can range from poor to very good depending on how well each segment is assimilated, if at all, to the native language, as well as on phonetic similarity between members of the contrast. Recall that within the UU contrasts examined /χ/-/ɣ/ was predicted to be the least discriminable. This was borne out; native AE listeners experienced great difficulty in telling these two sounds apart. They successfully discriminated only 36.4% of the stimuli, the worst of all contrasts.

/χ/-/ħ/

PAM predicted the uvular-pharyngeal /χ/-/ħ/ contrast to be a well-discriminated UU assimilation, since at least one segment (i.e., /ħ/) is remotely assimilable to English /h/. Results show this was not the case, however, with AE listeners discriminating /χ/-/ħ/ poorly, just above the chance level (54. 5%).

/q/-/χ/

The uvular plosive-continuant contrast was predicted to be a UU assimilation as well. Because English /k/ can be representative, although poorly, of Arabic /q/, the contrast was expected to be differentiated rather well. This prediction was borne out. American listeners demonstrated very good, 82.9%, discrimination of the contrast.

UU Assimilation

Of the three UU contrasts investigated in this study, /χ/-/ɣ/ and /q/-/χ/ support the PAM predictions. Averaged discriminability across all three contrasts is 57.9%, significantly lower than TC and CG assimilation types, but not UC, for which discriminability was even worse. As for CG contrasts, it is hard to assess the overall success rate of UU contrasts because of the range of variance in discriminability set by PAM (moderate to very good in CG, poor to very good in UU). Therefore, it is best to examine each contrast individually, as done so far, in order to evaluate the relevant PAM predictions. The results for the UU assimilation type as a whole in this study corroborate those of Guion et al. (2000), whose Japanese subjects, consistent with UU assimilation, failed to perceive the distinction between English /r/ and /l/.

In general, the data support the hypothesis that L2 Arabic contrasts are perceived by AE listeners within the confines of their native language phonological system. Their ability to discriminate the Arabic contrasts in this study relied crucially on how similar or different (assimilable or not) the contrastive sounds are to the listeners' native language. L2 Arabic contrasts involving sounds that are identical or similar to English tended to be perceptually much more discriminative as in /t/-/d/ and /θ/-/ð/. Other contrasts varied in their discriminability from poor (UC /ħ/-/h/, UU /χ/-/ɣ/ and /χ/-/ħ/) to good (CG

/t̲/-/t/, UC /k/-/q/) to very good (CG /ð̲/-/ð/, UU /q/-/χ/) as an apparent function of their assimilability to the listener's native language as well as the phonetic proximity of the contrasted segments to each other.

Such effects of the native language sound system on contrast discrimination and more generally speech perception have been well established. The loss of sensitivity to foreign contrasts has been attributed to interference from the ambient (native) language and thus, to use Kuhl's (1993) terminology, the once "citizen of the world" becomes a "culture-bound" perceiver. For adults, continued exposure to the native language, while rendering their auditory system more attuned, and thus confined, to the native language sounds, costs them the ability to detect minor phonetic differences that subsequently arise in learning a second language sound system. Strange (1995) summarizes the L2 perception experience as follows:

> Between early infancy and adulthood, then, children's interactions with their linguistic environment while acquiring their first languages produce significant changes in the perception of speech sounds. There is a "loss" in the ability to differentiate phonetic categories perceptually that are not phonologically distinctive in the native language, while native contrasts may become more highly differentiated. (p. 19)

Table 8 sums up the findings of this study for each of the 9 Arabic contrasts tested. It is clear that discrimination results for five of the contrasts in this study are in line with their relevant PAM predictions. The rest, however, contradict PAM, with discrimination in each case being less than expected, especially in the UC contrast /ħ/-/h/.

Table 8 PAM predictions vs. AE listeners' discrimination of Arabic contrasts

Contrast	Assimilation type	Predicted discriminability	Attested discrimination level
/t/-/d/	TC	excellent	100
/θ/-/ð/	TC	excellent	98.9
/t/-/t/	CG	moderate to very good	64.8*
/ð/-/ð/	CG	moderate to very good	72.7
/ħ/-/h/	UC	very good	37.5*
/k/-/q/	UC	very good	64.8*
/χ/-/ɣ/	UU	poor	36.4
/χ/-/ħ/	UU	very good	54.5*
/q/-/χ/	UU	very good	82.9

Note. Asterisks (*) mark discrimination contradicting PAM.

That the UU contrast /q/-/χ/ was significantly better discriminated than /χ/-/ɣ/ and /χ/-/ħ/ is not surprising if within-contrast confusability is taken into account, as recognized by PAM, although it is not entirely clear how this "proximity" or lack of it factors into the whole process of contrast discrimination. Presumably, the less similar two segments in a new contrast are to each other, the less confusable they become making it easy for listeners to tell which is which. It is possible that non-native contrasts involving manner of articulation are more distinctive perceptually than place or voicing contrasts: Of the three UU contrasts, /q/-/χ/ is the only one that distinguishes manner of articulation.

Perhaps another reason why listeners were biased to poorly discriminate /χ/-/ɣ/ and /χ/-/ħ/ compared to /q/-/χ/ is the orthographic similarity which the segments share. In Arabic, /χ/ is typically written as خ (when isolated) or خـ (when connected), /ɣ/ as غ or غـ, and /ħ/ as ح or حـ. On the other hand, Arabic /q/ is written ق or قـ. Table 9 illustrates the orthographic differences for these pairs. The first column shows the Arabic graphemes as letters written in isolation. The next three columns show how each of these graphemes is transcribed in connected writing according to its position in the word. Phonemes in word initial and medial positions are orthographically more similar to each other than in final position, which in turn is very similar to the phoneme in isolation. The last column presents the IPA symbols often used to transcribe the phoneme. As can be seen in Table 9, the phonemic-orthographic relation between these graphemes is clear: The uvular fricatives /χ/, /ɣ/ and the pharyngeal /ħ/ all resemble each other orthographically in initial, medial, and final positions, whereas the distinctive style in which /q/ is written sets it apart from others. It is possible, therefore, that the orthographic disparity between /q/ and /χ/ and lack of it in /χ/-/ɣ/ and /χ/-/ħ/ may have played a role in how learners perceived the contrast in these pairs.

Table 9 Orthographic representation of the Arabic phonemes: /χ/, /ɣ/, /ħ/and /q/

Isolated	Connected			IPA
	Initial	Middle	Final	
خ	خـ	ـخـ	ـخ	χ
غ	غـ	ـغـ	ـغ	ɣ
ح	حـ	ـحـ	ـح	ħ
ق	قـ	ـقـ	ـق	q

Research has shown that orthography is not an unlikely factor in non-native acquisition. Atkey (2001), for example, maintains that the orthographic conventions of the Czech writing system influence the reading ability of early Czech learners. Others have reported effects of orthography on speech perception (e.g., Dijkstra, Roelofs, & Fieuws, 1995; Taft & Hambly, 1985; Ziegler & Ferrand, 1998), and speech production (Lupker, 1982).

As for CG /t̠/-/t/ and UC /ħ/-/h/ and /k/-/q/, each consisting of one consonant present in the listeners' native language phonology paired up with another new (unfamiliar) albeit distantly related consonant, lesser discriminability or confusability between the consonants may have arisen primarily due to the new consonant being well assimilated to an already familiarly known segment. The pattern in which listeners incorrectly responded on the AXB task suggests that they significantly settled the contrast in each of these pairs in favor of /t/, /h/, /k/, respectively. That is, they overwhelmingly judged the English-like sound as the more similar one.

This aural penchant for the native segment varied, however, from one contrast to another. It was strongest for the pharyngeal-glottal pair /ħ/-/h/, with /h/ being chosen almost 89% out of the 55 incorrect responses reported (49/55). Next was the emphatic /t̠/-/t/ with /t/ adding up to 77.4% (24/31), and last was the velar-uvular contrast for which /k/ was selected 64.5% of the time (20/31). As such, this bias may have contributed to AE listeners' unexpectedly poor performance in the discrimination of these contrasts.

To conclude, findings of this study suggest the following discriminability hierarchy: TC > CG > UC, UU. The TC type contrasts significantly emerged as more highly discriminative than the CG, UC, or UU contrasts corroborating the TC > CG discriminability ranking introduced by PAM (Best, 1995) and reported in Best, McRoberts, and Goodell's (2001) examination of English listeners' perception of Zulu and Tigrinya consonant contrasts. The results of this study are also in line with the PAM predictions for the perception of nasal consonants in Malayam Marathi and Oriya by Malayalam, Marathi, Punjabi, Tamil, Oriya, Bengali and American English speakers (Harnsberger, 2001), non-native syllable structure and voicing contrast perception by Japanese learners of English (Nagao, Lim, & de Jong, 2003), and perception of obstruent word-final voicing contrasts by Malay speakers of English (Pilus, 2002).

Although PAM makes no explicit claims regarding the relative difficulty in the discriminability of CG, UC and UU types, which ranges from poor to moderate to very good, it states that discrimination for contrasts of the UC type should always be "very good." However, AE listeners in this study treated UC contrasts no better than CG or UU contrasts. In fact, there was new evidence suggesting better discrimination of CG contrasts compared to UC and UU types. Finally, the study has failed to maintain the distinction between UC and UU assimilation types, reported in Guion et al. (2000) for the discrimination of English /r/-/w/ and /r/-/l/, respectively, by Japanese listeners.

Conclusion

This small study tested the predictions of Best's (1995) PAM in the perception of Arabic consonant contrasts by American learners. With some exceptions, the pattern in which the different Arabic contrasts examined in this study were discriminated followed straightforwardly from PAM: TC > CG > UC, UU. In general, excellent discrimination was significantly associated with TC contrasts. Interestingly, CG contrasts were also significantly more distinctive than UC or UU contrasts. No distinction between UC and UU types emerged, however; both were poorly discriminated by AE listeners.

The perceptual assimilation model provides a useful tool in classifying non-native contrasts, and predicting the differential success non-native listeners achieve in discriminating them. The issue of assessing similarity between native and non-native phonemes remains a perennial one, not only to PAM but to any model of non-native speech acquisition. One measure of similarity that appears to be overlooked by PAM is the orthographic conventions of the target language. Evidence from this study suggests a role of orthography in the perception of L2 contrasts. Compared to /q/-/χ/, the voiceless-voiced contrast in /χ/-/ɣ/ and the uvular-pharyngeal contrast in /χ/-/ħ/ were highly indistinguishable to AE listeners presumably due to the orthographic similarity between these phonemes in the Arabic writing system. However, this remains a hypothesis that needs to be examined further. Given the implementation of orthography in the second language teaching curriculum and its importance as a factor in the perception of non-native phonemes, revisions to have it incorporated into models of second language perception ought to be made.

References

Al-Ani, S. H. (1970). *Arabic phonology*. The Hague: Mouton.

Al-Karouri, A. M. (1996). *Phonetics of classical Arabic: A selectional study of the problematic sounds*. Khartoum University Press: Khartoum.

Amayreh, M. M. (2003). Completion of the consonant inventory of Arabic. *Journal of Speech, Language, and Hearing Research, 46*, 517-529.

Archibald, J. (1998). *Second language phonology*. Amsterdam: John Benjamins.

Atkey, S. (2001). *The acquisition of L2 segmental contrasts: English speakers' perception and production of Czech palatal stops* (Unpublished master's thesis). University of Calgary, Canada.

Audacity Team. (2008). Audacity (1.2.4) [Computer program]. Retrieved from http://audacity.sourceforge.net/

Bateson, M. C. (1967). *Arabic language handbook*. Washington: Center for Applied Linguistics.

Beddor, P. S., & Gottfried, T. L. (1995). Methodological issues in crosslanguage speech perception research with adults. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 207-32). Timonium, MD: York Press.

Best, C. T. (1990). Adult perception of non-native contrasts differing in assimilation to native phonological categories. *Journal of the Acoustical Society of America, 88*, 177-178.

Best, C. T. (1993). Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year* (pp. 289-304). Dordrecht: Kluwer Academic.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist view on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 171-204). Timonium, MD: York Press.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109*, 775-794.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345-360.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross language perception on approximants. *Journal of Phonetics, 20*, 305-330.

Dijkstra, T., Roelofs, A., & Fieuws, S. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology, 49*, 264-271.

Eckman, F. (1987). Markedness and the contrastive analysis hypothesis. In I. Georgette & S. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 55-69). Cambridge, MA: Newbury House.

Flege, J. E. (1987). The production of new and similar phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47-65.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 233-277). Timonium, MD: York Press.

Flege, J. E., Bohn, O. S., & Jang, S. (1997). The production and perception of English vowels. *Journal of Phonetics, 25*, 437-470.

Fowler, C. A. (1989). Real objects of speech perception: A commentary on Diehl and Kluender. *Ecological Psychology, 1*, 145-160.

Gass, S. M., & Selinker, L. (2001). *Second language acquisition: An introductory course.* London: Lawrence Earlbaum.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107*, 2711-2724.

Halle, P. A., Best, C. T., & Levitt, A. (1999). Phonetic versus phonological influences of French listeners' perception of American English approximants. *Journal of Phonetics, 27*, 281-306.

Harnsberger, J. D. (2001). On the relationship between identification and discrimination of non-native nasal consonants. *Journal of the Acoustical Society of America, 110*, 489-503.

Ingham, B. (1971). Some characteristics of Meccan speech. *Bulletin of the School of Oriental and African Studies*, *University of London*, *34*(2), 273-297.

Klatt, H., & Stevens, K. N. (1969). Pharyngeal consonants. *Quarterly Progress Report*, *Research Laboratory of Electronics, MIT, 93*, 207-216.

Kuhl, P. K. (1993a). Early linguistic experience and phonetic perception: Implications for theories of developmental speech perception. *Journal of Phonetics, 21*, 125-139.

Ladefoged, P. (2001). *A course in phonetics*. Chicago, IL: Harcourt Brace Jovanovich.

Ladefoged, P. (2005). *Vowels and consonants: An introduction to the sounds of language*. Malden, MA: Blackwell.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Malden, MA: Blackwell.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan Press.

Larson-Hall, J. (2004). Predicting perceptual success with segments: A test of Japanese speakers of Russian. *Second Language Research, 20*(1), 35-80.

Lupker, S. J. (1982). The role of phonetic and orthographic similarity in picture-word interference. *Canadian Journal of Psychology, 36*, 349-376.

Nagao, K., Lim, B., & de Jong, K. (2003, August). *Perceptual acquisitions of non-native syllable structures by native listeners of Japanese*. Poster presented at the 15th International Congress of Phonetic Sciences, Barcelona, Spain.

O'Shaughnessy, D. (2000). *Speech communications: Human and machine*. Piscataway, NJ: IEEE Press.

Pilus, Z. (2002). *Second-language speech: Production and perception of voicing contrasts in word-final obstruents by Malay speakers of English* (Unpublished doctoral dissertation). University of Wisconsin, Madison.

Polivanov, E. (1931). La perception des sons d'une langue´etrang`ere [The perception of non-native language sounds.]. *Travaux du Cercle Linguistique de Prague, 4*, 79-96.

Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America, 89*, 2961-2977.

Rochet, B. L. (1995). Perception and production of L2 speech sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 379-410). Timonium, MD: York Press.

Selkirk, E. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations* (pp. 337-383). Dordrecht: Foris.

Strange, W. (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 2-45). Timonium, MD: York Press.

Taft, M., & Hambly, G. (1985). The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language, 24*, 320-335.

Thelwall, R. (1990). Illustrations of the IPA: Arabic. *Journal of the International Phonetic Association, 20*(2), 37-41.

Towell, R., & Hawkins, R. (1994). *Approaches to second language acquisition*. Clevedon, UK: Multilingual Matters.

Trubetzkoy, N. (1939). *Principles of phonology* (C. A. Baltaxe, Trans.). Berkeley, CA: University of California Press.

Werker, J. (1986). The effect of multilingualism on phonetic perceptual flexibility. *Applied Psycholinguistics, 7*, 141-156.

Werker, J., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics, 37*, 35-44.

Wright, R. (2004). A Review of perceptual cues and cue robustness. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology* (pp. 34-57). Cambridge: Cambridge University Press.

Yamada, R. A. (1995). Age and acquisition of second language speech sounds perception of /r/ and /l/ by native speakers of Japanese. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 306-320). Timonium, MD: York Press.

Yamada, R., & Tohkura, Y. (1992). The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Perception and Psychophysics, 52*, 376-392.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review, 5*, 683-689.

APPENDIX

List of the Arabic contrasts used in the experiment

Contrast	Stimuli
/t/-/d/	ʔaːˈtaː-ʔaːˈdaː
/θ/-/ð/	ʔaːˈθaː-ʔaːˈðaː
/s/-/z/	ʔaːˈsaː-ʔaːˈzaː
/s̱/-/s/	ʔaːˈs̱aː-ʔaːˈsaː
/ṯ/-/t/	ʔaːˈṯaː-ʔaːˈtaː
/ð̱/-/ð/	ʔaːˈð̱aː-ʔaːˈðaː
/ḏ/-/d/	ʔaːˈḏaː-ʔaːˈdaː
/ħ/-/h/	ʔaːˈħaː-ʔaːˈhaː
/q/-/k/	ʔaːˈqaː-ʔaːˈkaː
/q/-/ʔ/	ʔaːˈqaː-ʔaːˈʔaː
/ṯ/-/ḏ/	ʔaːˈṯaː-ʔaːˈḏaː
/χ/-/ɣ/	ʔaːˈχaː-ʔaːˈɣaː
/χ/-/ħ/	ʔaːˈχaː-ʔaːˈħaː
/ɣ/-/ʕ/	ʔaːˈɣaː-ʔaːˈʕaː
/ħ/-/ʕ/	ʔaːˈħaː-ʔaːˈʕaː
/ð̱/-/s̱/	ʔaːˈð̱aː-ʔaːˈs̱aː
/χ/-/ʕ/	ʔaːˈχaː-ʔaːˈʕaː
/ɣ/-/ħ/	ʔaːˈɣaː-ʔaːˈħaː
/q/-/χ/	ʔaːˈqaː-ʔaːˈχaː
/q/-/ɣ/	ʔaːˈqaː-ʔaːˈɣaː
/m/-/n/	ʔaːˈmaː-ʔaːˈnaː
/r/-/l/	ʔaːˈraː-ʔaːˈlaː
/w/-/j/	ʔaːˈwaː-ʔaːˈjaː
/dʒ/-/ʃ/	ʔaːˈdʒaː-ʔaːˈʃaː