THE INTERLANGUAGE SPEECH INTELLIGIBILITY BENEFIT: ARABIC-ACCENTED ENGLISH

by

Zachary Bret Rasmussen

A Senior Honors Thesis Submitted to the Faculty of
The University of Utah
In Partial Fulfillment of the Requirements for the

Honors Degree of Bachelor of Arts

In

Linguistics

Approved:

Rachel Hayes-Harb
Supervisor

Randall S. Gess
Chair, Linguistics

David E. Iannucci
Departmental Honors Advisor

Martha S. Bradley
Director, Honors Program

May 2007

ABSTRACT

This study investigated the intelligibility of native and Arabic-accented speech for native English and native Arabic listeners, in order to better understand factors that are related to differential intelligibility effects in the literature. The phonemic contrast between /b/ and /p/, which exists in English but not in Arabic, was considered. Native Arabic listeners and native English listeners in a forced choice word identification task in which they were presented with individual English words from minimal pairs, e.g. "pat" or "bat, and were asked to identify the words. The words were produced by both native English talkers and native Arabic talkers. With these particular talkers and listeners, there was no evidence for an interlanguage speech intelligibility benefit for listeners (i.e. native English listeners were more accurate at identifying Arabic-accented English words), but there may have been evidence for an interlanguage speech intelligibility benefit for talkers (i.e. Arabic-accented English words were just as intelligible as native English words for native Arabic listeners). Acoustic measurements of the native and Arabic-accented words revealed that the two groups of talkers manipulated voice onset times differently. While the native English talkers showed very distinct categories for /p/ and /b/ along the voice onset time continuum, the native Arabic talkers showed an overlap of the two categories.

.

For Mom and Dad. Thanks for everything

TABLE OF CONTENTS

I. Introduction

Second language learners are generally perceived to have non-native accents, and these accents can decrease the intelligibility of their speech to native listeners. Some aspects of accented speech are generally attributed to differences in the phonemic inventories of a speaker's native language and their second language. For example, French does not have the phoneme /ð/ (the first sound in the English word "this"), whereas English does. In French-accented English, the French phoneme /z/ is often produced as a substitute for the English phoneme /ð/. Aspects of non-native speech such as sound substitution can be described at the level of individual phonemes.

There are also characteristics of accented speech that are sub-phonemic; that is, they are the result of differences in the phonetic implementation of phonemes in the speaker's native language (L1) and their second language (L2), even for phonemes that are shared by both languages. For example, languages vary considerably in terms of voice onset time (VOT), the measure of time between the release of a stop consonant and the onset of voicing, or glottal pulsing, associated with a following voiced sound (e.g., a vowel). It has been shown that native speakers of Spanish, a language with considerably shorter VOT values than English, generally produce English voiceless stop consonants (/p t k/) with significantly shorter VOT values than do native speakers of English (Flege and Eefting, 1987). This means that even when native speakers of Spanish produce phonemes that are shared by both English and Spanish, such as /p t k/, they tend to produce them differently than native speakers of English. The problem here is with the

phonetic implementation of a phoneme, rather than the production of a different phoneme altogether.

Differences in the phonetic implementation of phonemes in a learner's L1 and L2 can result in non-target-like production of these sounds in the second language, which can lead to misperception of these sounds by listeners. It can also make it difficult for L2 learners to identify the phonemes when they are produced by native speakers of the second language. For example, MacKay, Flege, Piske, and Schirru (2001) found that Italian-English bilinguals who had been living in Canada for less than eight years often identified English voiced stop consonants (*e.g.*, /b d g/) in native English speech as their voiceless counterparts (*e.g.*, /p t k/). This misperception was attributed to the fact that Italian voiced stop consonants are generally pre-voiced, meaning that glottal pulsing begins *before* the articulation of the consonant, whereas English voiced stop consonants have short-lag voicing, meaning that glottal pulsing generally begins shortly *after* the articulation of the consonant. Native speakers of Italian and native speakers of English both have the same phonemic categories for voiced and voiceless stop consonants. However, because of differences in the phonetic implementation of these consonants (*i.e.*, VOT), they are perceived differently by speakers of each respective language. This is evidence that non-native accents are not only manifest in the production of L2 speech, but also in the way that L2 speakers perceive native speech in the L2. In other words, non-native accents occur both in speaking and in listening.

Consonant sounds are perceived categorically in a speaker's native language (Schouten and van Hessen, 1992). For example, native English speakers perceive all

bilabial stop consonants as being either a /p/ or a /b/; not something in between. This becomes more complex when we consider that languages vary according to how phonemes are categorized. In Italian, for example, /b/ is a pre-voiced consonant, whereas /p/ is a short lag consonant. In English, on the other hand, /b/ is a short lag consonant and /p/ is a long lag consonant. Italian and English, considered individually, have distinct categories for /b/ and /p/. However, the Italian category for /p/ and the English category for /b/ overlap respective to VOT values. This is what leads many native Italian speakers to identify English /b/ as /p/ (MacKay *et al*., 2001).

In order for non-native speakers to learn to produce and perceive L2 speech sounds like a native speaker, they must develop new phonological categories for L2 phonemes that do not exist in their native language and they must adjust the categories that exist in both the L1 and the L2 to be structured more similarly to those categories in the L2. This is extremely difficult for second language learners to do, but it has been shown that to some extent it is possible. Flege (1980) studied the productions of English /p t k/ by native speakers of Saudi Arabic. He found that the Saudis' productions exhibited many phonetic aspects of native Arabic speech. However, he also found that over time the Saudis gradually acquired the ability to approximate English characteristics of stop production so that their productions were not typical of either native Arabic speech or of native English speech. He determined that the Saudis' productions were the output of what he called an interlanguage phonetic system, a system that exhibits phonetic characteristics that are intermediate to L1 and L2 phonetic norms. This interlanguage phonetic system was not Arabic per se, but it was not quite English either.

The concept of an interlanguage system has been used to explain certain aspects of the intelligibility of non-native speech. Studies have demonstrated that non-native listeners may have an increased ability to understand non-native speech spoken by speakers of their own native language relative to native listeners(Imai, Walley and Flege, 2005; Munro, Derwing and Morton, 2006). For example, English spoken by a native speaker of Spanish might be more intelligible to another native Spanish speaker who is a second language learner of English, than to a native speaker of English. This phenomenon has been attributed to the existence of an interlanguage system that is shared by speakers of the same L1 and has been called the "interlanguage speech intelligibility benefit", subsequently referred to as ISIB (Bent and Bradlow, 2003).

Research into the ISIB has produced a set of complex results. Bent and Bradlow (2003) found that native Mandarin and Korean listeners were just as successful at identifying words in non-native English speech spoken by speakers with whom they shared the same native language as they were at identifying words in native English speech. This type of ISIB refers to the relative intelligibility of speech by non-native and native talkers for non-native listeners. It is referred to as the interlanguage speech intelligibility benefit for talkers (ISIB-T; Hayes-Harb, Smith, Bent and Bradlow, under review). Another type of ISIB is the one mentioned above in which non-native listeners are more successful at identifying non-native speech than are native listeners. This is referred to as the interlanguage speech benefit for listeners (ISIB-L; Hayes-Harb *et al.,* under review). Munro *et al*. (2006) showed that native Japanese listeners outperformed native English listeners at accurately identifying words in Japanese accented English

speech, providing evidence for the ISIB-L. However, they did not produce consistent results with native Cantonese listeners and Cantonese talkers, demonstrating the complex nature of the ISIB.

Hayes-Harb *et al*. (under review) conducted an experiment to better understand the differences between the ISIB-T and the ISIB-L, and the acoustic characteristics of native and non-native speech that might be responsible for the ISIB. They subjected native English listeners and native Mandarin listeners to a test in which they were presented with tokens of native English speech and Mandarin accented English speech and were asked to identify individual words in a forced choice task. They found that the native Mandarin listeners were better at identifying English words produced by native Mandarin speakers than were the native English listeners, providing support for the ISIB-L. However, the native Mandarin listeners were not as successful at identifying words in Mandarin-accented speech as they were at identifying native English speech, evidence inconsistent with the ISIB-T. They also identified acoustic phonetic characteristics of the Mandarin accented speech that may have contributed to the ISIB-L.

This research will follow the methods of Hayes-Harb *et al.* (under review) to investigate the interlanguage speech intelligibility benefit as it relates to English spoken by native speakers of Arabic. It will also explore the aspects of phonetic approximation found in Arabic-accented English that may contribute to the interlanguage speech intelligibility benefit. The focus of the research will be on productions of the English phoneme /p/ by native speakers of Arabic, and how these productions are perceived by both native speakers of Arabic and English.

One reason for choosing the phoneme /p/ is that it exists in English but not in Arabic. Therefore, in order for native speakers of Arabic to produce and perceive native English speech with native like accuracy, they must develop a new phonological category for English /p/. Flege and Port (1981), investigated the productions of English /p/ by native Arabic speakers. They found evidence that the native Arabic speakers understood the phonemic nature of English /p/ and were somewhat successful at developing an L2 category for /p/. The native Arabic speakers produced /p/ with a longer stop closure duration than they did for /b/. This is analogous to the timing contrast that exists in Arabic between voiced and voiceless stop consonants, *i.e.* native Arabic production of /t/ and /k/ have longer closure durations than those of /d/ and /g/. The native Arabic speakers seemed to understand that /p/ is analogous to /b/ in the same way that /t/ is to /d/. This shows that to some degree, the Arabic speakers were able to grasp the phonemic nature of /p/ in English; they were not merely substituting /b/ for /p/. However, they were still unable to produce /p/ with native like VOT and generally produced it with glottal pulsing. Flege cites evidence that L2 patterns of supraglottal timing seem to be easier to learn than those involving glottal timing. He attributes the ability of native Arabic speakers to transfer the voiced-voiceless closure duration contrast to /p/ and /b/, but not the VOT contrast, to this glottal-supraglottal distinction.

Another reason for choosing /p/ was that native speakers of American English have been shown to have more difficulty identifying productions of /p/ by native Arabic speakers than other voiceless stop consonants. Flege and Port (1981) tested the intelligibility of native Arabic productions of English voiceless stop consonants and

found that native English speakers could identify /t/ and /k/ fairly easily, but had trouble with /p/, frequently identifying it as /b/. They attributed this misperception to the presence of glottal pulsing during the stop closure. The native Arabic speakers did produce /p/ differently than /b/. However, this contrast was primarily evident in stop closure duration, rather than in VOT. Native English listeners, while sensitive to VOT contrasts, are not sensitive to timing contrasts in word-initial stop consonants. Therefore, they were usually unable to perceive the difference between the Arabic speakers' productions of /b/ and /p/ and generally identified them both as /b/.

II. Research Questions and Hypotheses

Flege and Port showed that native English speakers had difficulty identifying English /p/ produced by native Arabic speakers. This leads to the question of how this /p/ is perceived by native listeners of Arabic. According to the ISIB-L, they should be better than native English listeners at identifying /p/ in Arabic-accented English. The evidence that native Arabic speakers are able to grasp the phonemic nature of English /p/ and thereby produce it with some of the correlates of the Arabic voiced-voiceless contrast suggests that they should be able to perceive this contrast as well when produced by other native speakers of Arabic. According to the findings of Hayes-Harb *et al*. (under review) regarding the ISIB-T, Arabic listeners should also be better at perceiving native English speech than Arabic-accented speech.

This research will test perceptions of /p/ and /b/ in Arabic accented English by both native English listeners and native Arabic listeners in order to investigate the two

types of the ISIB previously discussed. The following two hypotheses were posed concerning the ISIB-L and the ISIB-T:

1) Native Arabic listeners will be more successful at identifying words in Arabic accented English than native English listeners. (ISIB-L)

2) Native Arabic listeners will be more successful at identifying words in native English speech than in Arabic accented speech. (ISIB-T)

In other words, we expected to find evidence in favor of the ISIB-L (Hypothesis 1), but not in favor of the ISIB-T (Hypothesis 2). These hypotheses were based on the findings of Hayes-Harb *et al*. (under review) because the methodology and tasks used to collect data in the current study were the same as those used by Hayes-Harb *et al*.

Additionally, we analyzed the productions of /p/ by both native Arabic speakers and native English speakers and compared these productions to productions of /b/ by native Arabic speakers in both an Arabic and English context. It was hypothesized that certain acoustic measurements, particularly stop closure duration and VOT, would show signs of phonetic approximation towards English norms. Specifically, the natve Arabic speakers' productions of /p/ should have acoustic characteristics intermediary to their productions of /b/ and to the native English productions of /p/. Also, as further support for the existence of an interlanguage system, the Arab's productions of /b/ in the context of English should show some approximation towards native English /b/. The third hypothesis is as follows:

3) The productions of /p/ and /b/ by native speakers of Arabic will be the product of an interlanguage phonetic system that is consistent with neither Arabic nor English phonetic norms but is intermediate to them.

III. The Experiment

The experiment followed the same methods and procedures used by Hayes-Harb *et al*. (under review).

1. Stimuli

Productions of word-initial /p/ and /b/ were recorded by three native speakers of Arabic and three native speakers of English producing English minimal pairs. The Arabic talkers were all speakers of the Gulf dialect of Arabic and were from Qatar. They reported being in an English speaking country for six months, one year, and a year and a half, respectively. The minimal pairs that were produced are: *pat* and *bat*; *pan* and *ban*; *pet* and *bet*; *pie* and *buy* (word frequency counts are reported in the Appendix). The words were read in the context of the sentence, "I'm going to say ____ and then I will say ____." The words were read four times, twice in each sentence position, making a total of sixteen productions of /p/ and sixteen productions of /b/ for each speaker. No word existed twice in the same sentence. From these productions, three tokens of each of the target words were randomly extracted from the productions of each speaker, making a total of 144 different stimuli tokens to be used in the listening experiment.

The Arabic talkers also produced Arabic sentences containing /b/ so that acoustic measurements from these sentences could be compared to those from the English productions (Hypothesis 3). They produced the words *bahr* (sea), *baarid* (cold), *balad* (country), and *beit* (house) in the context of "uhibb kalimata _____" (I like the word ____). They produced each word three times, making a total of 36 productions of Arabic /b/ for all three speakers. The individual auditory stimuli were extracted from these recordings using PRAAT software (Boersma, 2001).

2. Subjects

Eleven native English subjects and seven native Arabic subjects participated as listeners in this study. The native English listeners (5 female, 6 male) were recruited from linguistics courses at the University of Utah and were given course credit for their participation. Although some of them had studied foreign languages, none of them had studied Arabic or any related language. The native Arabic listeners (7 male, 0 female) were mainly recruited through student organizations, and an ESL class at the University of Utah and were paid for their participation. They were all from countries in which Gulf Arabic is the main spoken dialect. Five were from Saudi Arabia, one from Qatar, and one from the United Arab Emirates. All of the native Arabic listeners reported being in an English speaking country for less than one year (2-8 months). All of them were functionally fluent speakers of English (e.g., they were able to communicate with the experimenter about the experiment and task directions, and were able to fill out a

questionnaire). Individual Arabic subject language characteristics and daily language use information are reported in the Appendix.

3. Procedure

Subjects participated in a forced choice task in which they heard a single word and were asked to identify it as one member of a minimal pair. For example, they heard the word "pet" and were asked to identify it as either "pet" or "bet". Each of the two possible choices was presented on the computer screen: one on the left side and one on the right side. Subjects chose their answers by pressing the shift key (either right or left) that corresponded to the written word on the screen that they believed to be the correct answer. The stimuli were presented using DMDX software (Forster and Forster, 2003). Each of the 144 stimuli was presented once in random order in each of four different blocks of stimuli. Therefore, each subject heard a total of 576 stimuli, each token being heard four times. There was no time limit for subjects to respond to the stimuli. The experiment lasted about thirty minutes, and there was a break between each of the four blocks.

IV. Results

1. Identification Accuracy

Accuracy of word identification was calculated by coding listener responses as correct if they matched the intended word produced by the talker. The results were analyzed using a repeated measures analysis of variance (ANOVA) with listener group (two levels:

native English and native Arabic) as a between-subjects factor and talker group (two levels: native English and native Arabic) as a within-subjects factor. This analysis revealed a significant main effect of talker language ($F(1,34)=140.94$, $p< .001$, partial eta squared=.806), a significant main effect of listener language ($F(1,34)=41.706$, $p<.001$, partial eta squared=.551), and a significant interaction of listener language and talker language ($F(1,34)=20.29$, $p<.001$, partial eta squared=.374). The native English listeners were more accurate at identifying words spoken by native English talkers than by native Arabic talkers ($F(1,42)=203.11$, $p<.001$, partial eta squared=.829). There was only a marginal effect of talker language for Arabic listeners, English talkers being more intelligible than Arabic talkers ($F(1,26)=3.651$, $p=.067$, partial eta squared=.123). This is contrary to Hypothesis 2. Native English listeners were more accurate at identifying words spoken by native Arabic speakers than were the native Arabic listeners ($F(1,34)=18.958$, $p<.001$, partial eta squared=.358), contrary to Hypothesis 1, and they were also more accurate at identifying words spoken by native English speakers than were the Arabic listeners ($F(1,34)=51.542$, $p<.001$, partial eta squared=.603). This data is presented in Figure 1.

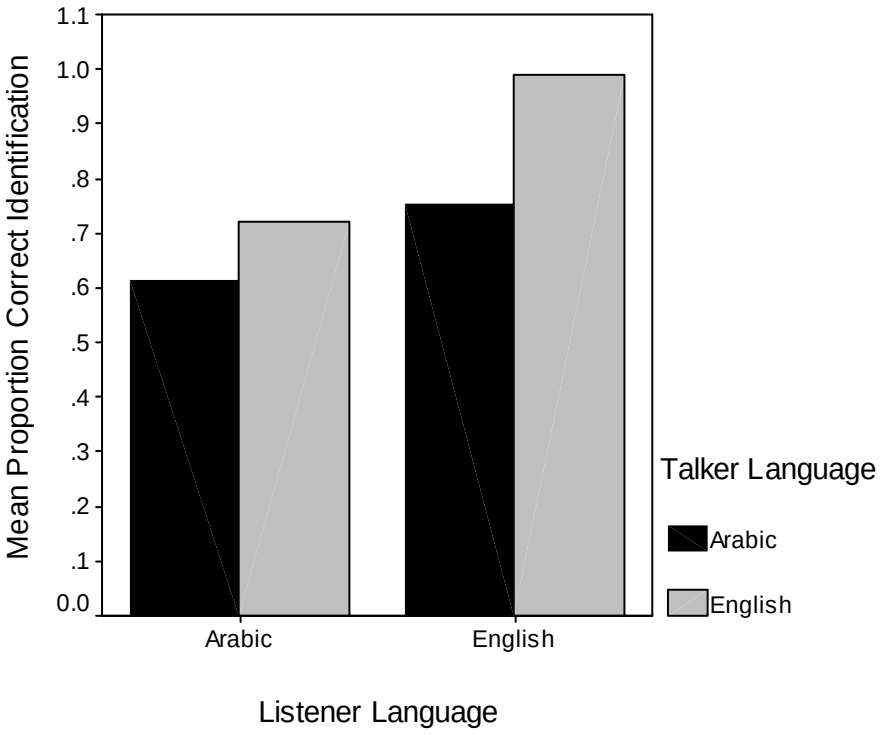

Figure 1. Mean proportion correct identification (bars represent +/- 1 standard deviation), by talker language and listener language.

2. Acoustic Characteristics of the Stimuli

Consonant closure duration, following vowel duration, and VOT were measured from the production data collected from the three native Arabic talkers and the three native English talkers and means were calculated from these measurements. Closure duration means (in milliseconds) are shown in Figure 2. By comparing the consonant closure duration means of both groups of talkers for both /b/ and /p/, we see that neither talker group manipulated closure duration to differentiate between the two phonemes.

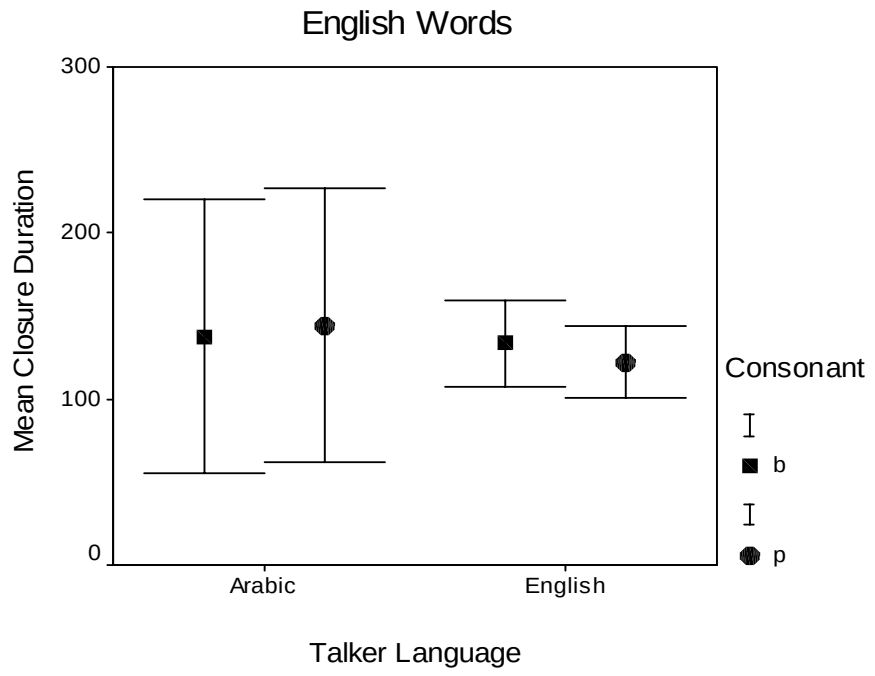

Figure 2. Mean closure duration data (bars represent +/- 1 standard deviation), by talker language and consonant.

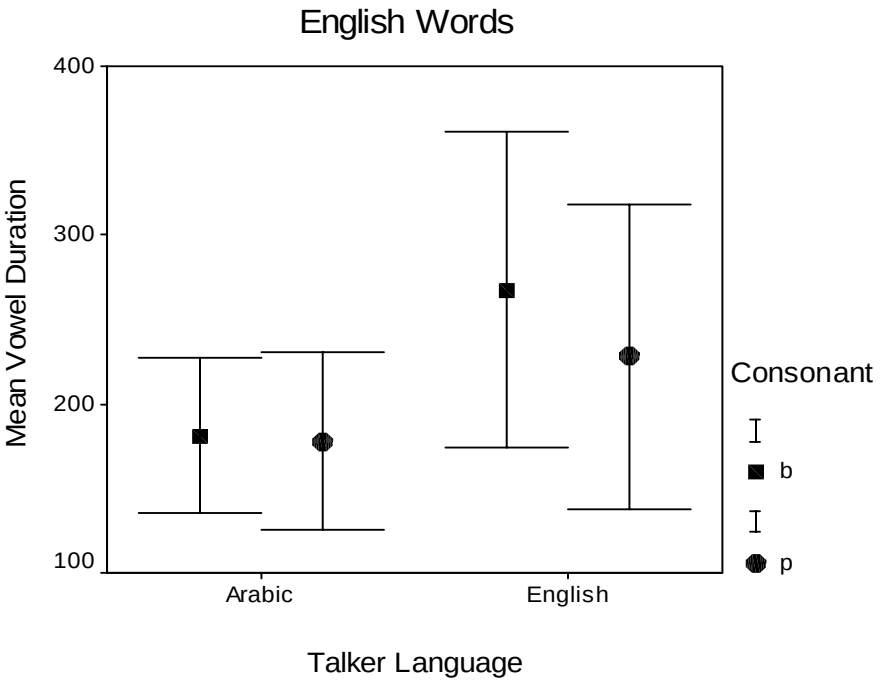

Figure 3. . Mean following vowel duration data (bars represent +/- 1 standard deviation), by talker language and consonant

Figure 3 shows the mean vowel duration measurements (in milliseconds) for the vowels immediately following /b/ and /p/. Neither group of talkers manipulated the duration of these vowels in such a way to distinguish between the two consonants. The data presented in Figure 2 and Figure 3 show us that neither consonant closure duration, nor following vowel duration were acoustic features manipulated by the talkers to differentiate between voiced and voiceless bilabial stops in English.

Means for voice onset time (in milliseconds) for English words produced by both groups of talkers are shown in Figure 4. There is no overlap in the VOT mean +/-1 standard deviation for /b/ and /p/ produced by native English talkers, which means that VOT was consistently manipulated by the native English talkers to differentiate between the two consonants. On the other hand, the data from the native Arabic talkers shows an overlap of VOT mean +/- 1 standard deviation for /b/ and /p/. This means that the native Arabic talkers did not consistently manipulate VOT in such a way to differentiate between /b/ and /p/. However, it is important to note that the native Arabic talker's VOT mean +/- 1 standard deviation for /p/ does not extend below zero, while their VOT mean +/- 1 standard deviation for /b/ does. This means that the native Arabic talkers consistently produced /p/ without pre-voicing (glottal pulsing beginning during the consonant closure), but frequently produced /b/ with pre-voicing. It is also important to note that the VOT +/- 1 standard deviations for /p/ produced by native Arabic talkers and /b/ produced by native English talkers overlap.

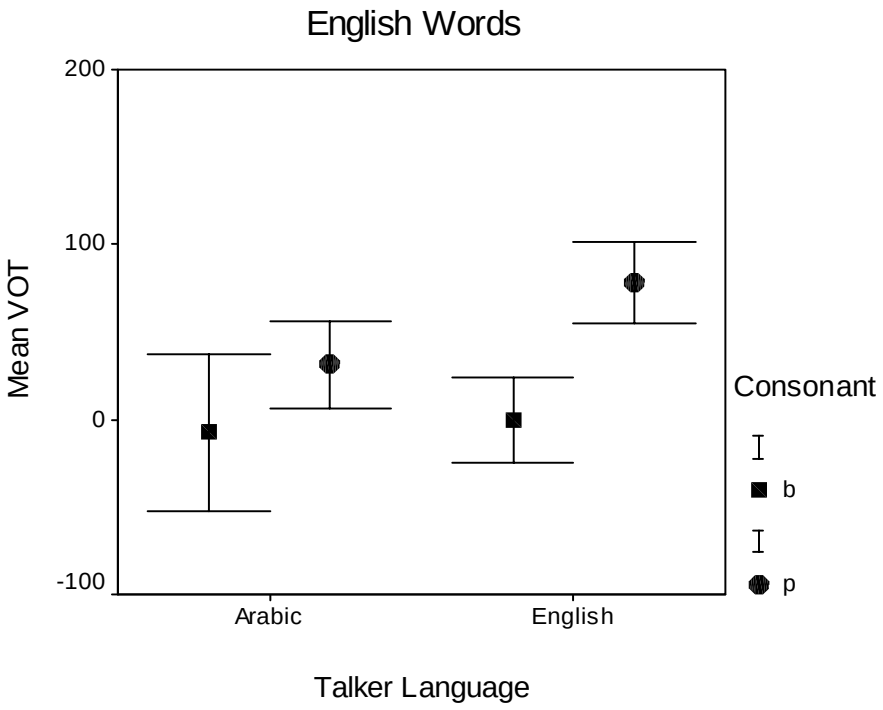

Figure 4. Mean VOT data for English words (bars represent +/- 1 standard deviation), by talker language and consonant.

Figure 4 compares the mean VOT measurements produced by the Arabic talkers for both English and Arabic words. All three VOT mean +/-1 standard deviations overlap to some extent, but there is much less overlap between the productions of /b/ in Arabic words and /p/ in English words than there is between the two consonants in English words only. When we compare Figure 3 with Figure 4 we can see that the mean VOT for the native Arabic productions of /b/ in English words is higher than the mean VOT for native Arabic productions of /b/ in Arabic words and approaches the mean VOT for native English productions of /b/.

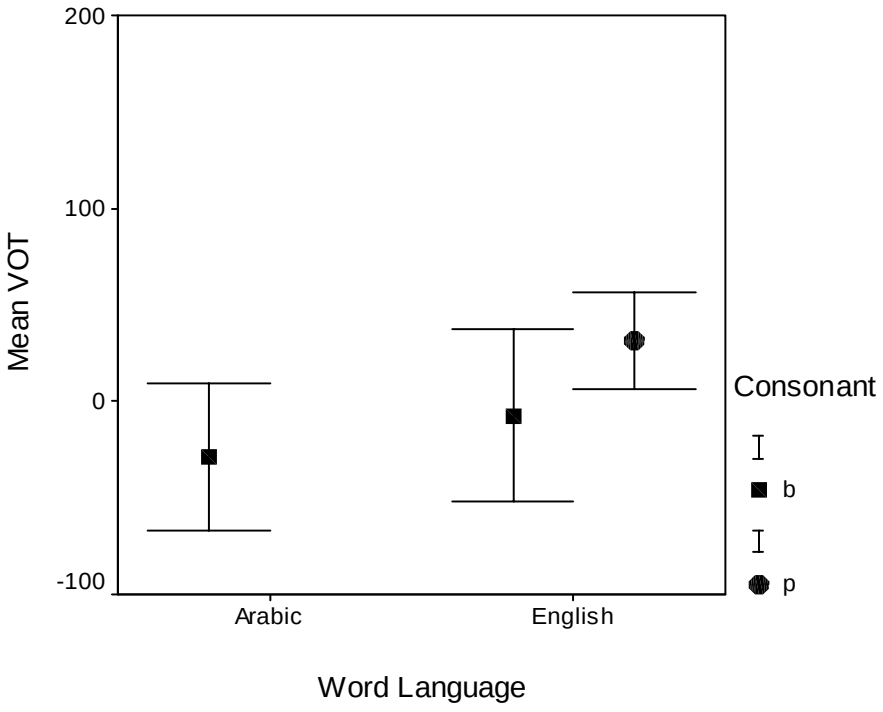

Figure 5. Mean VOT data for Arabic talkers (bars represent +/- 1 standard deviation), by word language and consonant.

V. Discussion

1. Hypothesis 1

Hypothesis 1 was that the native Arabic listeners would be more accurate at identifying words produced by native Arabic talkers than the native English listeners (ISIB-L). The statistical analysis of the data revealed that, contrary to Hypothesis 1, the native English listeners were more accurate than the native Arabic listeners at identifying the target words produced by native Arabic talkers. This is also contrary to the ISIB-L and to the findings of Hayes-Harb *et. al* (under review) that native Mandarin listeners were

significantly more accurate than native English listeners at identifying words produced by native Mandarin talkers.

The differences in findings between the current study and Hayes-Harb *et al*. (under review), may be attributed to some differences in the nature of the experiments and the subjects who participated in the experiments in each study. Hayes-Harb *et al*. tested the intelligibility of words in Mandarin-accented English that contained word-final voiced and voiceless obstruents. Whereas English maintains a voicing distinction for word-final voiced and voiceless obstruents, Mandarin does not allow obstruents word finally (but does allow them in other positions) and Mandarin speakers generally devoice word-final voiced obstruents in English. Hayes-Harb *et al*. tested an aspect of Mandarin-accented English that is the result of differences between the phonotactics of Mandarin and English, whereas the current study investigated an aspect of Arabic-accented English that is the result of a novel phonemic contrast that does not exist in Arabic.

Other differences between the two studies have to do with the amount of data collected and the subjects who provided both the stimuli and the production data. Hayes-Harb *et al*. used stimuli produced by 12 talkers, 6 native English and 6 native Mandarin. Within each talker group were 3 male talkers and 3 female talkers. The current study used stimuli from 6 talkers, 3 native English and 3 native Arabic. All 6 talkers were male. Hayes-Harb *et al*. collected intelligibility data from 15 native English listeners and 15 native Mandarin listeners, who were each presented with 1,152 stimuli, whereas the current study collected data from 11 native English listeners and 7 native Arabic listeners, who were each presented with 576 stimuli. There are also differences in the L2

backgrounds of the subjects. The native Arabic listeners had all been living in an English speaking country for less than one year. Most of them were students enrolled at the English Language Institute or in an ESL class at the University of Utah. Most of the native Mandarin speakers who participated in Hayes-Harb *et al*. were students enrolled in graduate and undergraduate programs at the University of Utah. The difference in the actual amount of data that was collected in each study and the demographic differences between the different groups of subjects may offer some explanation as to the differences in results. Future research will need to focus on native Arabic talkers and listeners who speak English at a higher proficiency level and who have been living in an English speaking country for a longer period of time.

2. Hypothesis 2

Hypothesis 2 was that native Arabic speakers would be more accurate at identifying words spoken by native English talkers than by native Arabic talkers. This hypothesis is not consistent with the ISIB-T, but is consistent with the findings of Hayes-Harb *et al*. (under review), which found that native Mandarin listeners were more accurate at identifying words produced by native English speakers than by native Mandarin speakers. The statistical analysis of the data collected from the native Arabic listeners revealed that there was no significant effect of talker language on their number of correct responses. However, there was a marginal effect ($p=.067$) in favor of the English talkers. In other words, the native Arabic listeners were marginally better at identifying words produced by native English talkers than by native Arabic talkers. The lack of a significant effect

seems to be consistent with the findings of Bent and Bradlow (2003) that non-native listeners may be just as accurate at identifying non-native speech as native speech (ISIB-T). However, the marginal effect may be evidence contrary to the ISIB-T, and consistent with Hypothesis 2. More data will need to be collected to clarify these findings.

3. Hypothesis 3

Hypothesis 3 was that the productions of /p/ and /b/ by native speakers of Arabic would be the product of an interlanguage phonetic system that is consistent with neither Arabic nor English phonetic norms but is intermediate to them. Flege and Port (1981) found that native Arabic speakers were producing English /p/ with a longer closure duration than English /b/, but that they were not able to manipulate voice onset time in a way that could differentiate between /b/ and /p/. On the other hand, the production data collected for this study showed that the native Arabic talkers were not manipulating consonant closure duration, but were to some extent manipulating VOT. The native Arabic talker's VOT mean +/-1 standard deviations for English /b/ and /p/ overlapped, but were not the same. The mean VOT for /p/ was considerably higher than the mean VOT for /b/. This shows that, while the native Arabic talkers were not able to consistently produce /b/ and /p/ as completely distinct categories along the VOT continuum, they did seem to understand to some degree that /b/ and /p/ are distinguished in English by VOT and their productions demonstrate an attempt to approximate English VOT norms. The Arabic talkers' mean VOT for /b/ in English words is also slightly higher than their mean VOT for /b/ in Arabic, which may also be the result of an attempt to approximate English VOT norms.

The production data show evidence that the native Arabic talker's productions are the product of an interlanguage phonetic system, though that system may be in the beginning stages of development. The subjects who provided the production data had all been living in an English speaking country for less than a year and a half. To better understand the development of interlanguage phonetic systems we will need to collect more data from talkers who have been living in an English speaking country for longer periods of time and who speak English at different levels of proficiency. By comparing the data for higher proficiency speakers with the data that has been collected we should be able to learn more about how interlanguage systems develop, especially those involving a novel phonemic contrast.

VI. Conclusion

The results presented here highlight important issues pertaining to the intelligibility of non-native speech. Two types of interlanguage speech intelligibility benefit, the ISIB for listeners and the ISIB for talkers, were defined and investigated in order to better understand how non-native speech is perceived by both native and non-native listeners. The results showed that the native Arabic listeners who participated in the study had no advantage over the native English listeners at identifying words in Arabic-accented speech. In fact, the native English listeners were significantly more accurate than the native Arabic listeners. This may be the result of low L2 proficiency on the part of the native Arabic talkers and listeners. Also, the native Arabic listeners showed no significant advantage at identifying words from native English speech over words from

Arabic-accented speech. The acoustic analysis of the Arabic and Arabic-accented speech showed that, while the native Arabic talkers were not consistently producing /p/ and /b/ with acoustic features characteristic of native English speech, they were also not producing /p/ and /b/ consistently in the same way, and they produced /b/ in English with a slightly higher mean VOT than /b/ in Arabic. While it appears that an interlanguage phonetic system had been developed to a certain degree, it seems that this system had not yet been developed to the extent that it provided a benefit to the native Arabic listeners.

References

Bennett, J. (2005). "Voicing Analysis of English and Arabic Stop Consonants, Produced by a Native Speaker of Arabic," M.S. Thesis, University of Utah.

Bent, T., and Bradlow, A.R. (2003). "The Interlanguage Speech Intelligibility Benefit," Journal of the Acoustical Society of America 114, 1600-1610.

Boersma, P. (2001). "Praat, a System for doing Phonetics by Computer," Glot International 5, 341-345.

Caramazza, A., Yeni-Komshian, G.H., Zurif, E.B., Carbone, E. (1973). "The Acquisition of a New Phonological Contrast: The case of Stop Consonants in French-English Bilinguals," Journal of the Acoustical Society of America 54, 421-428.

Cho, T., and Ladefoged, P. (1999). "Variation and Universals in VOT: Evidence From 18 Languages," Journal of Phonetics 27, 207-229.

Flege, J.E., (1980). "Phonetic Approximation in Second Language Acquisition," Language Learning 30, 117-133.

Flege, J.E., and Eefting, W. (1987). "The Production and Perception of English Stops by Spanish Speakers of English," Journal of Phonetics 15, 67-83.

Flege, J.E., and Port, R., (1981). "Cross-Language Phonetic Interference: Arabic to English," Language and Speech 24, 125-146.

Forster, K.I., and J.C. (2003). "DMDX: A Windows Display Program with Millisecond Accuracy," Behavior Research Methods, Instruments, and Computers 35, 116-124.

Hayes-Harb, R., Smith, B.L., Bent, T., Bradlow, A.R. (under review). "Production and Perception of Final Voiced and Voiceless Consonants by Native English and Native Mandarin Speakers: Implications Regarding the Interlanguage Speech Intelligibility Benefit."

Keating, P., Linker, W., Huffman, M. (1983). "Patterns in Allophone Distribution for Voiced and Voiceless Stops," Journal of Phonetics 11, 277-290.

Khattab, G. (2000). "VOT Production in English and Arabic Bilingual and Monolingual Children," Leeds Working Papers in Linguistics 8, 95-122.

MacKay, I.R.A., Flege, J.E., Piske, T., Schirru, C. (2001). "Category Restructuring During Second-language Speech Acquisition," Journal of the Acoustical Society of America 110, 516-528.

Munro, M.J., Derwing, T.M., Morton, S.L. (2006). "The Mutual Intelligibility of L2 Speech," Studies in Second Language Acquisition 28, 111-131.

Port, R.F., and Mitleb, F.M. (1983). "Segmental Features and Implementation in Acquisition of English by Arabic Speakers," Journal of Phonetics 11, 219-229.

Schouten, M.E.H. and Hessen, A.J. van (1992). "Modeling Phoneme Perception." I. Categorical Perception. Journal of the Acoustical Society of America 92: 1841-1855.

Smith, B.L., Bradlow, A.R., Bent, T. (2003). "Production and Perception of Temporal Contrasts in Foreign Accented English," *Proceedings of the XVth International Congress of Phonetic Sciences, Barcelona, Spain*. Edited by M.J. Sole, D. Recasens, and J. Romero.

Weinrich, U. (1953). *Languages in Contact: Findings and Problems*. (Mouton, the Hague).

Yeni-Komshian, G.H., Caramazza, A., Preston, M.S. (1977). "A Study of Voicing in Lebanese Arabic," Journal of Phonetics 5, 35-48.

Appendix

Production Lists

1. I'm going to say "cat" and then I will say "ban."
2. I'm going to say "to" and then I will say "buy."
3. I'm going to say "bat" and then I will say "pet."
4. I'm going to say "tie" and then I will say "can."
5. I'm going to say "ban" and then I will say "pat."
6. I'm going to say "die" and then I will say "pet."
7. I'm going to say "bat" and then I will say "to."
8. I'm going to say "pat" and then I will say "tie."
9. I'm going to say "bet" and then I will say "bat."
10. I'm going to say "pan" and then I will say "do."
11. I'm going to say "pie" and then I will say "die."
12. I'm going to say "do" and then I will say "pan."
13. I'm going to say "tie" and then I will say "bet."
14. I'm going to say "pet" and then I will say "cat."
15. I'm going to say "can" and then I will say "pat."
16. I'm going to say "buy" and then I will say "ban."
17. I'm going to say "pat" and then I will say "buy."
18. I'm going to say "cat" and then I will say "pie."
19. I'm going to say "pan" and then I will say "can."
20. I'm going to say "die" and then I will say "cat."
21. I'm going to say "pie" and then I will say "do."
22. I'm going to say "bet" and then I will say "pan."
23. I'm going to say "to" and then I will say "pie."
24. I'm going to say "do" and then I will say "bat."
25. I'm going to say "buy" and then I will say "die."
26. I'm going to say "ban" and then I will say "to."
27. I'm going to say "tie" and then I will say "bet."
28. I'm going to say "can" and then I will say "pet."

١ أحب كلمة بحر

٢ أحب كلمة بار

٣ أحب كلمة بلد

٤ أحب كلمة بيت

٥ أحب كلمة بحر

٦ أحب كلمة بار

٧ أحب كلمة بلد

٨ أحب كلمة بيت

٩ أحب كلمة بحر

١٠ أحب كلمة بار

١١ أحب كلمة بلد

١٢ أحب كلمة بيت

Word Frequency Counts (occurrence per million words)

pan	16
ban	7
pet	8
bet	20
pie	14
buy	70
pat	35
bat	18

Individual Arabic Listener Characteristics

Listener	**More Arabic or English Daily**	**Length of time in English speaking country**
A1	Arabic	7 months
A2	Arabic	7 months
A3	Equal	7 months
A4	Arabic	8 months
A5	English	2 months
A6	English	2 months
A7	Arabic	4 months