ACOUSTIC EFFECTS OF DURATIONAL CUES IN THE PERCEPTION OF *NAJDI* ARABIC VOWEL CONTINUA

Mahmoud S. Al Mahmoud
Imam University, Riyadh, Saudi Arabia
E-mail: msmahmoud@imamu.edu.sa

Abstract

This paper explores how the manipulation of vowel duration as a perceptual cue influences listeners' perceptual ability. Four native speakers of *Najdi* Arabic, a well-known variety of Arabic in the Arabian Peninsula, were tested on the perception of /a/ vs. /ɛ/ vowels. Listeners' identification and discrimination rates along each vowel continuum showed a clear effect of duration on the perception of /a/-/ɛ/ contrast. In each vowel continuum, listeners were more inclined to classify stimuli as belonging to one vowel or the other based on relative proximity to the steady-state vowel duration. Perceptibility naturally improved as duration approximated the normal duration of either vowel. Listeners' perceptual judgments in the identification and discrimination of the vowels were swayed by their aural sensitivity to perceptual shifts (/a/-/ɛ/ at 185-195ms; /ɛ/-/a/ at 195-205ms). Moreover, findings of the identification task followed predictably from the discrimination task; this could be taken as evidence for the existence of categorical perception. Results aggregately indicate that perception of the two *Najdi* Arabic vowels proceeded as a function of duration.

1. Introduction

Much of the phonetic work on Arabic has focused on the standard form of the language overlooking different dialects that have diverged considerably from standard Arabic. It is simply impossible to assume that all speakers of Arabic use one variety of the language, and as such, it becomes essential to explore in more detail the phonetic intricacies of dialectal Arabic. This paper examines one of the most prominent phonetic cues in the perception of vowels, namely duration. The variety of Arabic in question is *Najdi* Arabic (NA), a dialect widely spoken in the central part of the Arabian Peninsula, nowadays Saudi Arabia. There is quite a dearth of phonetic studies on Arabic dialects, particularly Najdi, with most of the work being descriptive in nature (Al-Ani, 1970 and Mitleb, 1984). The paper explores how instrumental duration in and of itself can be in influencing native speakers' ability to perceptually identify, categorize and discriminate contrasts involving two well-recognized, although frequently confounded, vowels, namely, /ɛ/ and /a/. To this end, identifiability as well as discriminability of these two vowels by native speakers of *Najdi* Arabic are assessed along durational continua where duration emerges as an essential cue for perception. The paper attempts to determine if and where a perceptual shift between /ɛ/ and /a/ exists, and how informative this transition is for listeners in the distinction of these two vowels.

2. Literarure Review
2.1 Effects of Duration on Perception

In the identification of consonants and vowels there exist a number of factors that contribute to the robustness of the acoustic signal. For vowels, energy is found essentially below the 1 kHz threshold and it declines considerably around -6 dB/oct with frequency. This energy is concentrated at multiple levels known as the formants (multiples of F0). The first formant (F1) is:

> "readily identified in time plots of many vowels as the inverse of the period of dominant oscillation within a pitch period. Front vowels in particular have a wide separation between F1 and F2, and the lowpass nature of the glottal source causes F1 to have much more energy than higher formants in these cases." (O'Shaughnessy, 2000: 58).

In addition to the locations of the first three fundamental frequencies (formants), duration plays an important role in distinguishing vowels. Vowels are inherently longer than consonants and their intrinsic duration could vary according to the external factors such as the phonetic context in which they occur. Stress as well as speaking rate can influence the length of a vowel substantially, so can the voicing status of the consonant following the vowel with relatively longer vowels before voiced consonants than before voiceless ones. Vowels, especially stressed ones, average about 100-130ms, although variations in length, as well as formant frequencies, are common due to many variables such as the speaker's age, gender or mode of conversation. Duration in vowels is also greater when they occur before continuants than before stops (Borden, Harris, & Raphael, 2003).

A number of studies have reported on the effect of vowel duration in the perceptibility of consonants. Hogan and Rozsypal (1980) systematically reduce the length of the vowel using 24 English monosyllabic words, and show that vowel duration is an important cue in the voicing distinction of word final consonants. Ainsworth (1981) explores the relationship between duration and the identification rate of synthetic vowels, and concludes that the ability to identify and classify synthetically modified vowels is closely related to duration, since recognition of vowels varied as a function of durational differences.

Tsukada (2009) examines vowel length contrasts in three different languages, Arabic, Japanese and Thai. The findings indicate that duration, among other spectrally related cues, can assist in vowel distinction. Listeners systematically relied on duration in perceptually discriminating vowels (also Ueyama, 2003). Luo, Li, and Mok (2019) test native Mandarin speakers' ability to distinguish vowel length contrasts in Cantonese. Even though vowel duration is not contrastive in Mandarin, listeners were able to exploit durational differences in discriminating vowels contrasts.

Hillenbrand, Clark, & Houde (2001) investigate the effect of duration on vowel recognition. In order to evaluate the role of duration in the perception of vowels, fifteen native speakers of American English trained in phonetics participated in a listening experiment that involved identification of CVC syllables with variable synthesized durations generated at multiple intervals. Their findings, although not fully supportive of duration as a primary perceptual identification cue, establish a clear effect for duration in the recognition of the pair /æ/-/ɛ/, among other vowels.

In their examination of Australian vowels, Watson and Harrington (1999) report an effect of duration in the classification of different vowels, specifically when duration complements other acoustically significant measures such as formant trajectories. Dupoux,

Kakehi, Hirose, Pallier, and Mehler (1999) study the effect of different durations on vowel perception by French and Japanese listeners. A six-step vowel continuum was generated (from zero vowel e.g. *ebzo* to full vowel e.g. *ebuzo*) to gauge the discoverability of the vowel. Their results naturally show that the longer the duration of the vowel is, the better identification rates are. In a more recent study, Mok (2011) investigates how vowel duration, in addition to vowel quality, influences vowel-to-vowel coarticulation in Thai, but concludes that by itself duration has little effect, and that other vowel qualities contribute to the degree of overlapping.

2.2 Selection of the NA Vowel Contrast

The durational contrast selected for the identification and discrimination experiments in this study is the *Najdi* Arabic vowel pair /a/ and /ɛ/. NA is a dialectal variety of Modern Standard Arabic and is one of the main dialects spoken by people in and around the central region of the Arabian Peninsula (Ingham, 1994). Vowel length in Arabic is contrastive word finally and more commonly medially, but not initially since no vowel initial words exist in Arabic (Al-Ani, 1970, p. 75). Specifically, NA maintains a length contrast between /a:/ and /a/, /i:/ and /i/, /u:/ and /u/. Front mid /e/ (sometimes diphthongized /ej/), front high mid /ɛ/, and back mid /o/ have no long counterparts. The vowels in question are the short counterpart of /a:/ (i.e. /a/) and the shorter /ɛ/ vowel, which is often treated, and transcribed by some scholars, as /a/.

Acoustic measurements of the NA vowels are needed primarily to determine if the /a/-/ɛ/ vowel pair is ideal for testing durational effects. It is crucial to look for vowel pairs that rely less on spectral formant trajectories in their identification; presumably if two vowels possess formant frequency values of close proximity, their classification, as well as discrimination, would be more reliant on other temporal acoustic cues such as duration. Vowel separability in such cases would be augmented by durational differences, which predictably can alter listeners' perception of the vowel quality. No previous acoustic analysis of NA vowels and their formant dimensions exists, to my knowledge. Therefore, a preliminary analysis of the acoustic measurements of the NA vowels is essential for this study.

To this end, two native speakers of *Najdi* Arabic (ages 24 and 27) were recorded reading a wordlist that contains the five NA vowels: /a/, /i/, /u/, /ɛ/, and /e/. The vowels formed the nucleus of monosyllabic nonsense words with the voiced bilabial /b/ as the initial consonant. The following shows the seven syllables read by the NA talkers: /ba/, /bi/, /be/, /bɛ/, and /bu/. The materials list was made up of 25 test words (five instances of each syllable*5 vowels=25) which were presented to the talkers in Arabic orthography using Praat (version 5.0.32) speech recording and editing software. Test words were embedded in the carrier phrase /hiyah _____ ʔalkɛlɪmɛh/ "the word is _____". Each phrase was randomly displayed using Praat script function with the subjects controlling stimulus presentation by clicking on a button to initiate recording. The recording took place in a quiet library room using an external clip-on PRO 7 Electret condenser microphone and Audacity audio editor and recorder software (version 2.0.0).

The first instance of each of the five vowel syllables recorded was selected for analysis. An exception, however, is the syllable for /u/ where the second instance was chosen since the vowel of the first instance was considerably longer than other instances. Using Praat (version 5.0.32), spectral and waveform representations of the actual syllables were generated as represented in Figures (1-5):
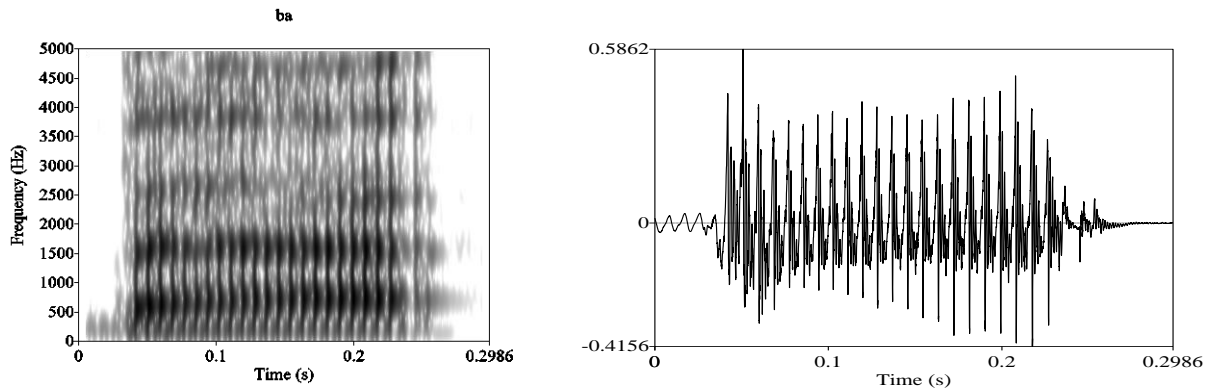
Figure 1: Spectrographic and waveform representation of the syllable /ba/
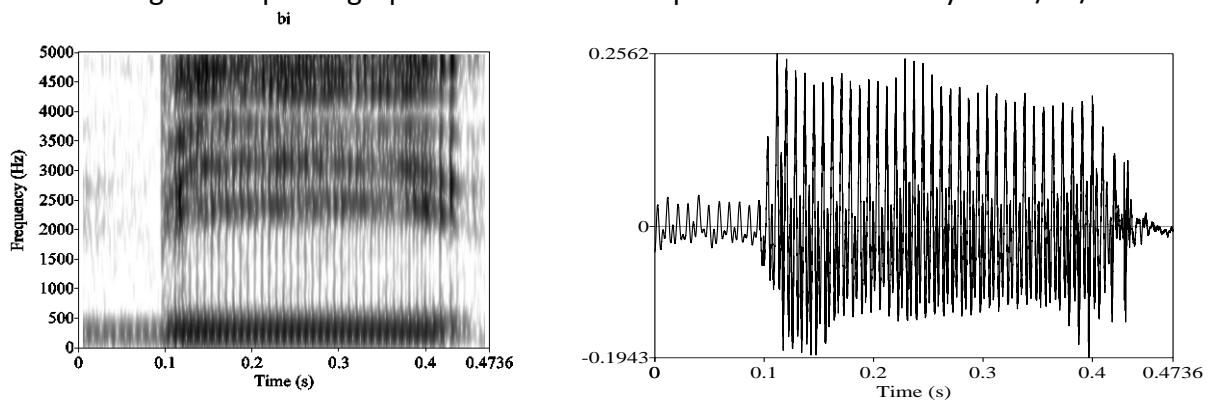

Figure 2: Spectrographic and waveform representation of the syllable /bi/
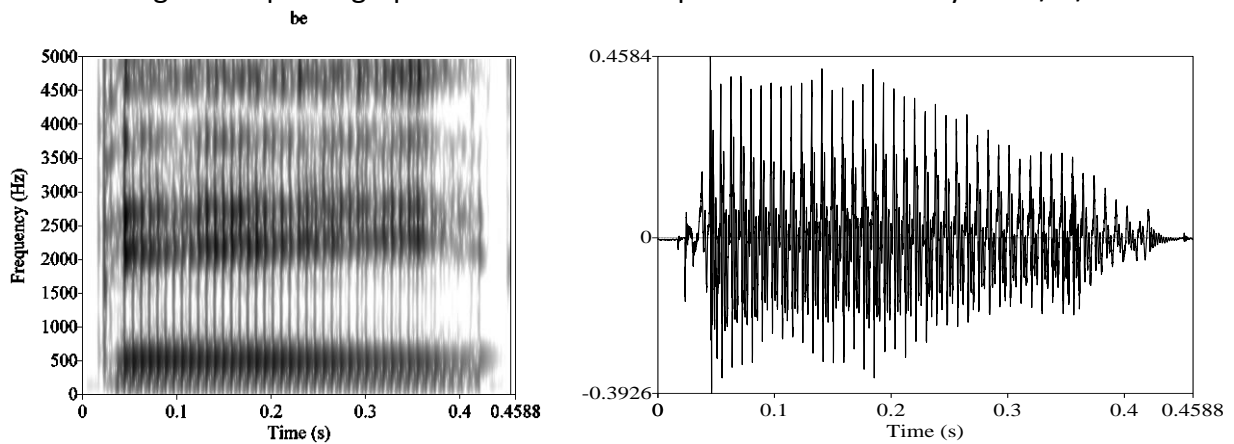

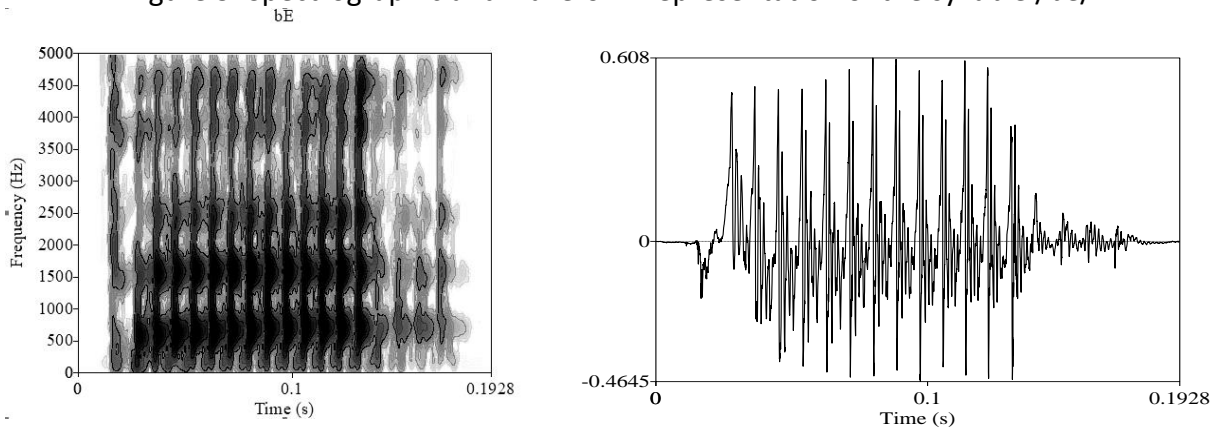Figure 3: Spectrographic and waveform representation of the syllable /be/

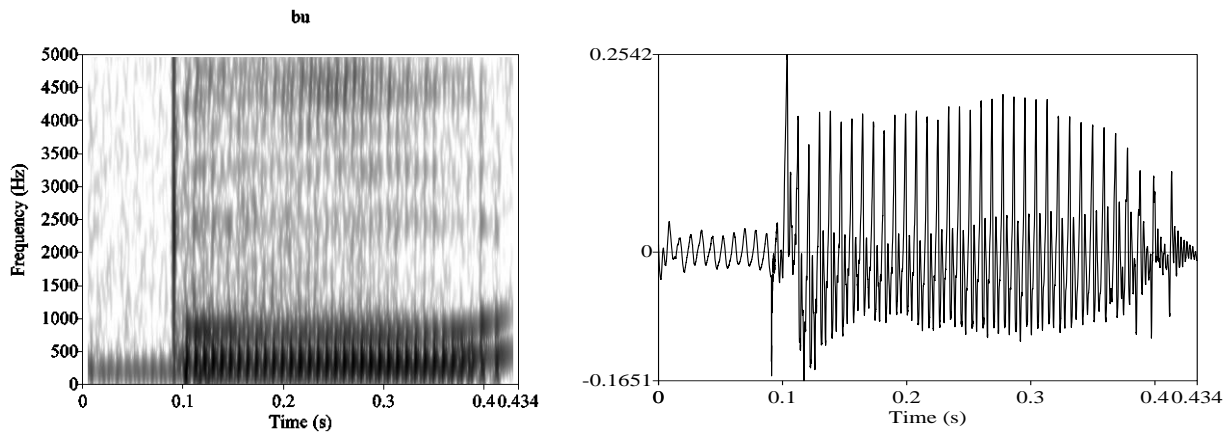Figure 4: Spectrographic and waveform representation of the syllable /bɛ/

Figure 5: Spectrographic and waveform representation of the syllable /bu/

The vowel portion in each syllable was specified by determining the start and end points. Duration and formant frequency measurements were then computed for each of the five vowels. Table 1 summarizes the results of the average duration measurements, F1, F2, and F0 for each of the five vowels tested in this experiment:

Vowel	Duration	F1	F2	F0
a	245ms	727	1550	114
e	380ms	435	2160	112
ɛ	155ms	644	1683	117
i	338ms	239	2402	118
u	327ms	298	730	112

Table 1: Vowel measurements in *Najdi* Arabic

The F1 and F2 values of the vowels in question were then plotted on a vowel formant grid:
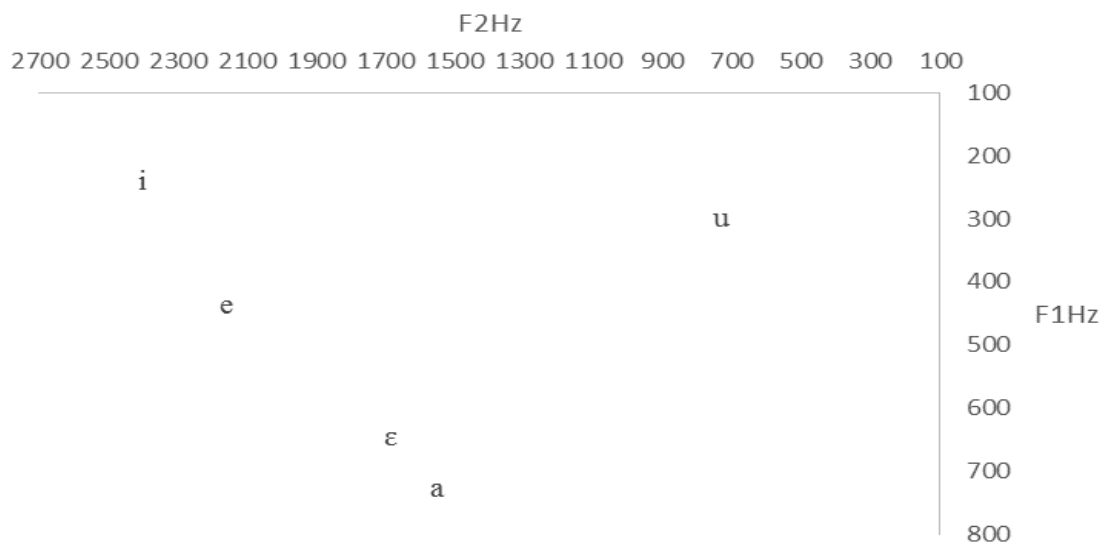

Figure 6: F1 and F2 plot grid of the five *Najdi* vowels

Figure 6 plots all five vowels on a formant grid, with the second formant as the x-axis and the first formant as the y- axis. F0 estimates' differences are irrelevant since they are very small for all five vowels as seen in Table 1, and as such are excluded in the vowel

formant grid. The selection of the pair /a/ and /ɛ/ in this study is thus justified on the basis of their proximity to each other with regard to their formant trajectories. Note how minimal is the difference between /a/ and /ɛ/ in the first and second formant frequencies. Additionally, third formant frequency estimates for this vowel pair are quite close, with F3 valued at 2570Hz for /a/, and F3 at 2441Hz for /ɛ/. It is not unreasonable to assume, owing to the similar formant structure in both /a/ and /ɛ/ vowels, that duration would be an instrumental acoustic cue in the distinction of these two vowels. This hypothesis is appealing especially because the duration measures between the vowel pair appear to be steep. Note that /a/ measures up to 245ms in duration while /ɛ/ only 155ms.

3. Research Method

As discussed in section 2.2, the vowel pair /a/-/ɛ/ represents an interesting case for testing the effects of duration since other spectral information is arguably less definitive due to proximity between the two vowels in their F1 and F2 values. If the formant frequencies in the two vowels are controlled for, perception of the pair will heavily rely on duration. Another reason for the selection of this vowel pair is the fact that although they are used contrastively in *Najdi* Arabic, for example /mɛl/ '(he) got bored' and /mal/ '(he) leaned over', they are commonly confused with each other. It is important to note that Arabic orthography reserves the letter /ٳ/ for /a/ but /ɛ/ is written as a *fatha* diacritic superimposed on the letter /◌/. In fact, many NA speakers are not even aware that /ɛ/ is a vowel in their own language.

In order to appreciate the effects of duration on /a/-/ɛ/ perceptibility, an experiment involving NA listeners' identification and discrimination of the vowel pair is carried out. The study aims to address whether duration solely is influential in the perception of vowel quality differences between /a/ and /ɛ/. Based on the discussion in section 2.2, it is hypothesized that native speakers of NA will be able to identify and discriminate the /a/-/ɛ/ vowel contrast based on durational differences. To this end, identification and discrimination tasks were designed to test *Najdi* listeners' ability to identify and discriminate stimuli along /a/-/ɛ/ vowel continua.

3.1 Participants

Four native speakers of *Najdi* Arabic, different from the ones who took part in the production experiment in section 2.2, served as listeners in all four tasks of the experiment (two identification tasks and two discrimination tasks). The listeners have lived all their lives in the Najd region, Riyadh, and grew up in *Najdi* families. They were graduate students at Imam University with an age range of 27-32, and were recruited via the author's personal contacts. While all listeners reported taking some basic English courses, none has received any specialized training in pronunciation or in Arabic or English phonetics. None has reported participating previously in an auditory experiment, and according to self-report, none suffers from any hearing difficulty.

3.2 Materials

Materials for this experiment include open syllable monosyllabic nonsense words with the voiced bilabial /b/ as the initial constant consonant and the two NA vowels namely /a/, and /ɛ/ as the nucleus. The two vowels are taken from a larger list previously recorded by native talkers of *Najdi* Arabic (see section 2.2). The averaged measurements (from the five tokens) of duration, F1, F2, F3 and F0 values for each vowel are restated in Table 2:

Vowel	Duration	F1	F2	F3	F0
a	245ms	727	1550	2570	114
ɛ	155ms	644	1683	2441	117

Table 2. Vowel measurements for /a/ and /ɛ/ in *Najdi* Arabic

Since the difference between the F1, F2 and F3 values of the endpoint vowels /a/ and /ɛ/ is very minimal, only the duration parameter between the two vowels estimated around 90ms was used to create a ten-step vowel duration continuum for each vowel. Vocalic steady state F1, F2, F3, F0 and duration measures (outlined in Table 2 above) were used as the continuum endpoints. For each vowel, nine intermediate steps with 10ms duration intervals were interpolated by either lengthening or shortening the duration of the vocalic element (i.e. less the stop occlusion portion) using PSOLA in Praat (version 5.0.32). For each continuum, the sound was selected in Praat and then a manipulation object with an empty duration tier was produced. A new duration tier was created and new duration points (i.e. longer/shorter) were added. This new duration tier was applied to the manipulated sound object and the modified sound was then produced separately using the publish synthesis function in the manipulation editor file menu. A Praat script function was used to interpolate the nine steps of the continuum each with 10ms duration interval. The endpoint vowels formed the first steps of each continuum. Spectrographic and waveform representations of every synthesized step were examined and only minor durational discrepancies between the nominal and observed values were tolerated.

3.3 Stimuli
For the identification task, 100 stimuli for each vowel continuum (10 stepsX10 reps) were created. For the three-step AXB discrimination task, 280 test trials were generated for each vowel continuum (28 trialsX10 reps). The 28 trials are made up of seven stimuli pairs with 30ms duration intervals between them. The 30ms interval was determined in a pilot test after experimenting with 10ms, 20ms and 30ms durations on a subject whose results are not reported here. For each pair, four trials were constructed (AAB, BBA, ABB, BAA). The following table summarizes the stimuli for both the identification and discrimination tasks:

Task	Identification	Discrimination	Discrimination Example
Stimuli	155ms	155ms-185ms	155-155-185, 185-155-155, 185-185-155, 155-185-185
	165ms	165ms-195ms	165-165-195, 195-165-165, 195-195-165, 165-195-195
	175ms	175ms-205ms	175-172-205, 205-175-175, 205-205-175, 175-205-205
	185ms	185ms-215ms	185-185-215, 215-185-185, 215-215-185, 185-215-215
	195ms	195ms-225ms	195-195-225, 225-195-195, 225-225-195, 195-225-225
	205ms	205ms-235ms	205-205-235, 235-205-205, 235-235-205, 205-235-235
	215ms	215ms-245ms	215-215-245, 245-215-215, 245-245-215, 215-245-245
	225ms		
	235ms		
	245ms		

Table 3. Identification and discrimination stimuli for both /a/ and /ɛ/

3.4 Procedure

Two identification tasks were carried out (one for each vowel continuum). ExperimentMFC function in Praat was used to design a two-choice identification task in which each of the 100 stimuli was aurally presented in a randomized fashion. After hearing each stimulus via headphones, subjects were asked to indicate if they heard /a/ or /ɛ/ by clicking on one of two boxes displayed on a computer screen. The boxes showed the syllable with short and long vowels written in Arabic orthography (e.g. [بَ] [بِ], /ba/ and /bɛ/, respectively). The next stimulus was played 500ms after each click was made. An optional subject-controlled break was offered every 50 tokens. The two identification tasks were administered over two sessions with an intervening 5-minute break. A 3-item practice test preceded the experiment to ensure subjects understood the instructions. The two tasks lasted around 15 minutes for each subject.

For the discrimination experiment, two AXB discrimination tasks (one for each vowel continuum) were designed also using the ExperimentMFC function. It is believed that an AXB discrimination task provides a reference stimulus (i.e. X) against which listeners estimate similarity as opposed to a simple AX discrimination task where listeners may base their 'same' or 'different' responses on nonlinguistic factors (Beddor and Gottfried, 1995). The 280 trials were presented via headphones in a random order. Every trial was a triad with three stimuli each separated by 500ms. After listening to each trial, listeners had to decide whether the first or the third word was more similar to the second word by clicking on two boxes shown on a computer screen. The next trial began 500ms after each click was made. Optional breaks every 70 trials were provided. The two discrimination tasks were carried out over two sessions with 5-10 minute breaks intervening. Again, prior to taking the experiment a 3-item practice test was given to ensure subjects understood the AXB task instructions. For each subject the experiment lasted around 25 minutes.

All four identification and discrimination tasks were administered individually for each participant using Koss R80 headphones in a quiet library room setting. None of the participants reported any hearing discomfort or problems after the experiments.

4. Results and Discussion

The goal of this paper is to explore the role of duration as an acoustic cue in the perception of *Najdi* Arabic vowels. It is hypothesized that native speakers of NA will be able to successfully discriminate /a/ and /ɛ/ vowel contrasts and identify them based on durational differences alone. Thus, we would expect NA listeners' performance on the identification as well as the discrimination tasks to be determined to a large degree by variations in the vowel duration. In the experimental design of this study, *Najdi* native speakers' ability to identify and discriminate the two NA vowels was tested along vowel continua with an identification and a discrimination task for each of the two NA vowels. The results here are discussed individually for each vowel continuum.

4.1 The /a/ vowel continuum

NA listeners' identification responses on the 10-step /a/ vowel continuum were tallied for each of the four subjects. Table 4 below shows identification scores on the /a/ duration continuum both individually and averaged among subjects:

	Duration									
Subject	155ms	165ms	175ms	185ms	195ms	205ms	215ms	225ms	235ms	245ms
1	0	0	33	42	69	100	100	99	100	100
2	1	6	40	50	91	100	93	95	100	100
3	4	8	29	43	79	83	92	100	90	98
4	2	5	30	55	77	79	88	95	97	100
Mean	1.75	4.75	33	47.5	79	90.5	93.2	97.2	96.7	99.5
Std. Deviation	1.7	3.4	4.9	6.1	9.1	11	4.9	2.6	4.7	1

Table 4. Identification function scores for the vowel /a/

Identification rates were collapsed across all four subjects and pooled for each of the durational increments in Table 4. To determine whether performance on the identification task significantly follows from durational differences in vowel length, the data were submitted to a one-way repeated measures (within subjects) ANOVA with Duration as the ten-level independent variable. Results indicate a significant effect of Duration on listeners' ability to identify vowels correctly, $F(9, 27) = 194.8$, $p < .001$, $\eta_p^2 = .95$. Figure 7 plots NA listeners' identification rates on each of the ten vowel continuum steps:
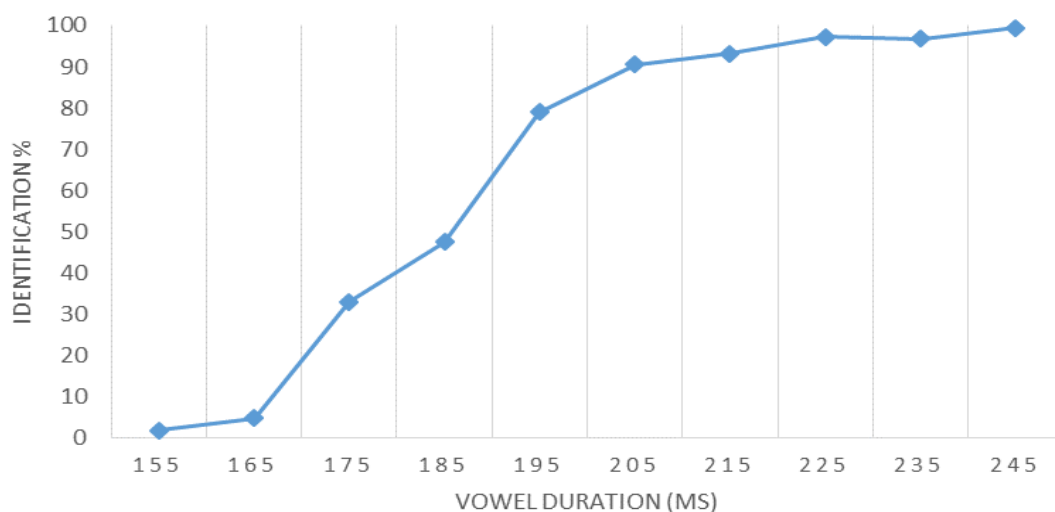

Figure 7. Duration effects on /a/ vowel identification

Results from the identification task indicate that duration did play a role in the perception of /a/ vowel. Listeners consistently and significantly relied on durational length differences to draw a distinction between /a/ and /ɛ/. Figure 7 demonstrates that on the lower end of the vowel continuum, and more specifically in the area of 155ms, 165ms, 175ms, listeners were biased towards hearing /ɛ/ rather than /a/ as clear from their low mean identification rates on these durations, 1.75%, 4.75%, 33%, respectively. Even when stimuli length was 185ms, still identification was below chance, 47.5%. It seems that for the NA listeners, a perceptual shift between /a/ and /ɛ/ exists in the 185-195ms threshold. This is where identification rates noticeably improved from 47.5% to 79%. That is, listeners were sensitive to durational differences in length, and the 195ms continuum step marked the beginning of distinction between /a/ and /ɛ/. Note that in longer durations of the vowel, 205ms to 245ms, subjects were successfully able to classify stimuli as /a/. This improvement

in vowel classification is quite expected as stimuli approach higher ends of the continuum close to the normal duration measurement of /a/, which is 245ms, as noted in Table 2.

Next, responses of the discrimination task were computed for each subject and tallied. The following table presents discrimination rates, means and standard deviation values for the vowel /a/:

Subject	Duration						
	155-185ms	165-195ms	175-205ms	185-215ms	195-225ms	205-235ms	215-245ms
1	75	75	80	60	95	75	55
2	60	50	50	60	75	50	40
3	50	80	70	55	80	55	60
4	54	71	75	49	73	69	53
Mean	59.7	69	68.7	56	80.7	62	52
Std. Deviation	10.9	13	13	5.2	9.9	11.7	8..5

Table 5. Discrimination function scores for the vowel /a/

A one-way repeated measures (within-subjects) ANOVA, which tested the significance of different length durations on NA listeners' vowel discriminability, reveals a significant effect of the independent variable Duration (7 levels), $F(6, 18) = 5.6$, $p < .05$, $\eta_p^2 = .97$. Results of the discrimination task are shown in Figure 8:
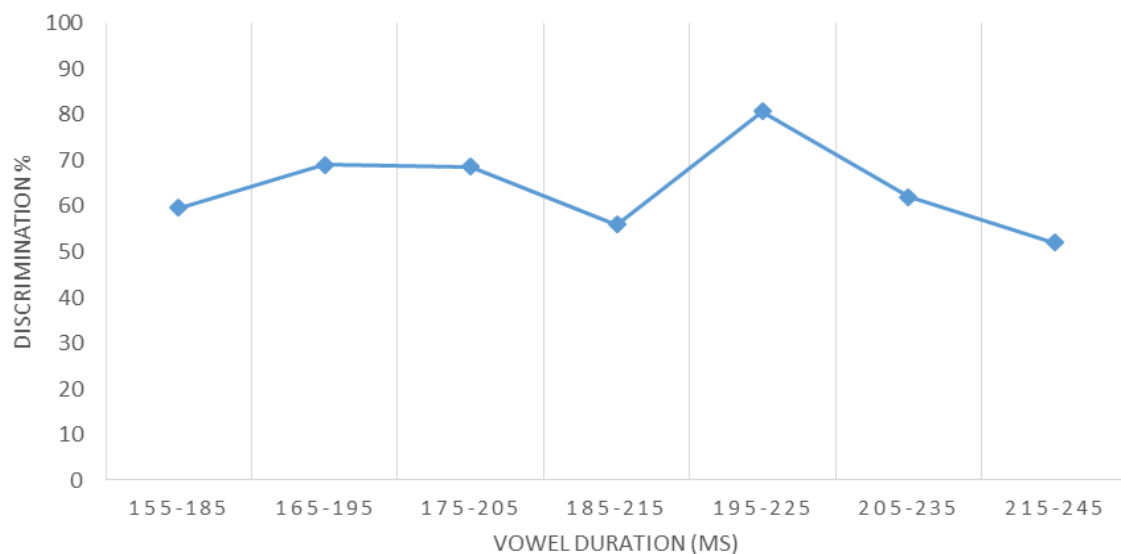

Figure 8. Duration effects on /a/ vowel discrimination

Although discrimination rates in general support the hypothesis in that they follow from durational variations, discriminability appears to be modest across different vowel durations. The 195-225ms length contrast marks a perceptual boundary between /a/ and /ɛ/ as clear from the overall enhanced ability of listeners to discriminate these two vowels, 80.7%. It seems that for NA listeners discrimination is at its best when duration spans the 195-225ms range. This is interesting since in the identification task, 195ms marked a transitional stage as well, and can be taken as evidence for the tendency of NA listeners to perceive the vowel /a/ categorically. Categorical perception is a phenomenon where within-category stimuli are

harder to discriminate than stimuli belonging to two separate phonetic categories (Liberman, 1996). The perception of the NA listeners in the discrimination task, as well as in the identification task, was aided by an abrupt perceptual shift that took place after the 195ms mark in discrimination and the 185ms mark in identification. This perceived dichotomy in the auditory distinction clearly biased listeners' responses on both tasks towards /a/. The existence of such perceptual boundary, as well as the correlative relationship between both tasks are quite characteristic of categorical perception (Repp, 1984).

4.2 The /ɛ/ vowel continuum
　　Responses on the identification task were tallied for each vowel duration. Standard deviation and means were computed across all subjects as shown in Table 6:

	Duration									
Subject	155ms	165ms	175ms	185ms	195ms	205ms	215ms	225ms	235ms	245ms
1	97	100	88	90	89	48	32	8	0	0
2	93	90	97	88	79	52	22	13	5	2
3	100	98	85	84	82	68	29	17	6	2
4	94	90	92	82	86	59	19	11	4	1
Mean	96	94.5	90.5	86	84	56.7	25.5	12.2	3.7	1.2
Std. Deviation	3.1	5.2	5.2	3.6	4.4	8.8	6	3.8	2.6	.95

Table 6. Identification function scores for the vowel /ɛ/

　　Data submitted to a repeated measures (within subjects) ANOVA show significance of Duration (10 levels) on vowel identification, $F(9, 27) = 276$, $p < .001$, $\eta_p^2 = .99$. Listeners' identification of /ɛ/ along the vowel continuum are depicted in Figure 9:
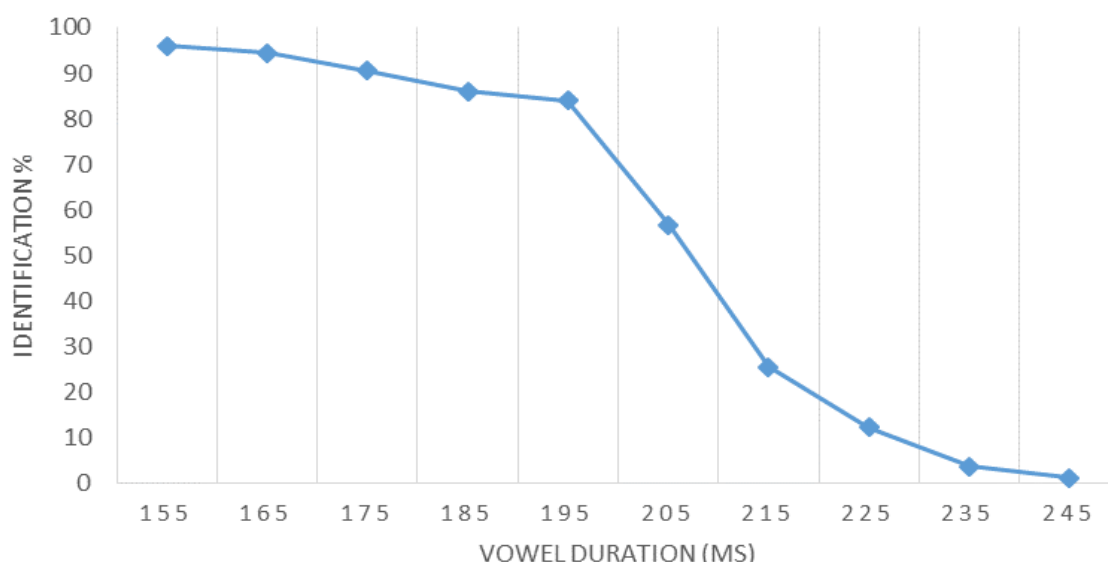

Figure 9. Duration effects on /ɛ/ vowel identification

　　In their identification of /ɛ/ NA listeners predictably were more successful in classifying /ɛ/ in shorter durations. As can be seen from Table 6 and Figure 9, identification rates are highest in the 155ms, 165ms, and 175ms continuum steps. Poor identification seems to take

place somewhere between the 195ms and 205ms range, and declines drastically thereafter. In other words, as the duration of the /ɛ/ vowel gets longer, listeners' penchant to hearing /a/ rather than /ɛ/ becomes stronger. On the 245ms duration step /ɛ/ was identified as /a/ more than 99% of the time. Understandably so since the 245ms length is characteristic of /a/ not /ɛ/. What is important to note here is that a high level of confusion, due to increased similarity, exists after the 195ms boundary.

To find out if listeners were sensitive to durational difference in their discrimination of the /ɛ/ vowel, discrimination rates were analyzed using a repeated measures (within subjects) ANOVA which tested for the significance for the independent variable Duration (7 levels). Results show, however, that duration had no statistical significance on listeners' ability to accurately discriminate /ɛ/ vowel stimuli, $F(6, 18) = 1.66$, *ns*. Table 7 and Figure 10 illustrate subjects' discrimination rates of the /ɛ/ vowel:

Subject	Duration						
	155-185ms	165-195ms	175-205ms	185-215ms	195-225ms	205-235ms	215-245ms
1	40	65	55	52	50	62	65
2	50	41	65	50	45	40	60
3	55	62	40	45	50	54	55
4	32	49	44	47	53	58	69
Mean	44.2	54.2	51	48.5	49.5	53.5	62.2
Std. Deviation	10.3	11.2	11.3	3.11	3.3	9.6	6

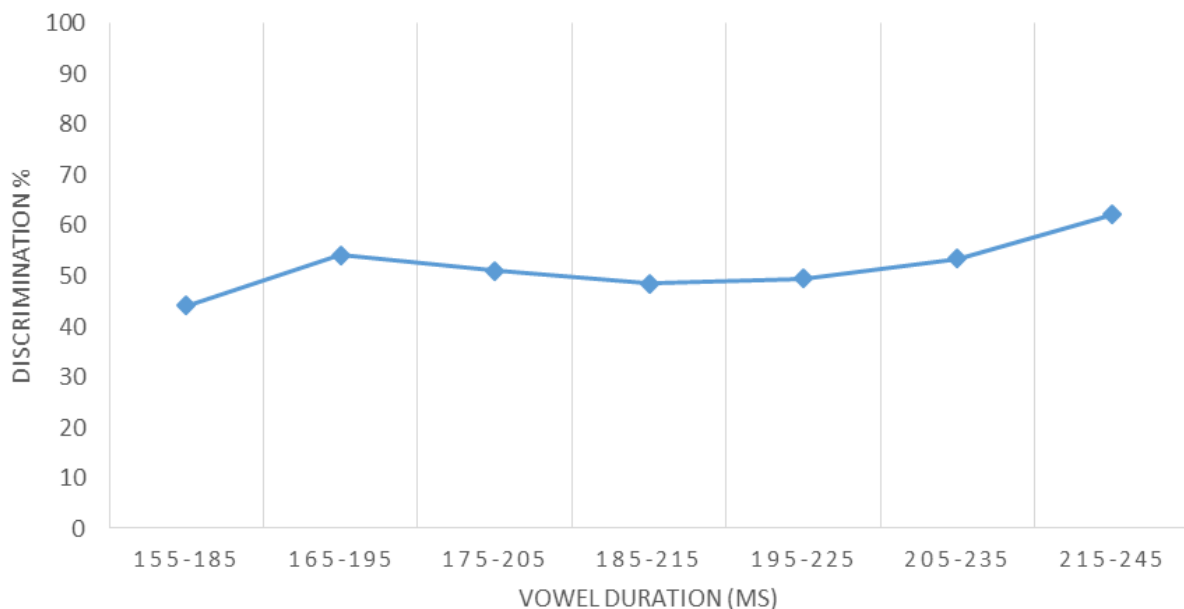Table 7. Discrimination function scores for the vowel /ɛ/

Figure 10. Duration effects on /ɛ/ vowel discrimination

Unlike the results for the identification task, discrimination of the /ɛ/-/a/ vowel contrast here did not significantly react to durational differences. Minor nonsignificant differences only exist as shown in Table 7 and Figure 10. The overall performance of the native NA listeners seems to hover around chance level, indicating a high level of uncertainty among

respondents as they make their choices on the discrimination task. It is not readily clear, however, why listeners were able to correctly classify /ɛ/ in the identification task but failed to do so on the discrimination task. Discrimination in general is a more demanding task than identification as listeners have to process multiple similar stimuli in a short window of time. It is possible that this may have led listeners to confuse /ɛ/ with /a/ and provide answers based on guessing. Recall too that this discrimination task is the last of four subsequent tasks and subjects' performance may have been inadvertently and inevitably undermined due to fatigue. Nonetheless, the findings here cannot be taken to express a significant tendency toward counting duration as a prominent determiner in listeners' overall discriminability of the /ɛ/-/a/ contrast.

5. Conclusion

This paper examined the effect of duration as an acoustic cue on the perceptual ability of native *Najdi* Arabic speakers to classify and distinguish the vowel pair /a/-/ɛ/. Identification results from a 10-step duration continuum for each vowel revealed listeners significantly employed discrepancies in length in their perception of /a/ vs. /ɛ/. Similarly, the seven different duration steps along the vowel continua provided listeners with robust acoustic cues in their discriminatory responses when presented with the auditory /a/-/ɛ/ contrast, although findings were insignificant for /ɛ/. In general, the performance of the NA subjects on both tasks for each vowel suggests that listeners categorically classified each vowel according to the perceived proximity to its steady-state normal duration measurement. For the NA listeners, the perceptual shift from /a/ to /ɛ/ seems to occur somewhere in the 185-195ms range. Whereas for /ɛ/-/a/, the transition occurs in the 195-205ms intervals. It is argued that, in addition to spectral information in the acoustic signal such as formant frequencies, perceptual distinctiveness between the NA front vowels /a/-/ɛ/ varies a function of vowel duration.

References

Al-Ani, S. (1970). *Arabic Phonology: An Acoustical and Physiological Investigation*. The Hague, Netherlands: Mouton.

Ainsworth, W.A. (1981). *Duration as a Factor in the Recognition of Synthetic Vowels*. Journal of Phonetics, 9,333–342. [Online] Available: https://doi.org/10.1016/S0095-4470(19)30976-3

Beddor, P.S. and Gottfried, T.L. (1995). *Methodological Issues in Crosslanguage Speech Perception Research with Adults.* In Strange, W. (ed), Speech perception and linguistic experience: issues in crosslanguage research. Timonium, MD: York Press, 207–32.

Borden, G., Harris, K., and Raphael, L. (2003). *Speech Science Primer: Physiology, Acoustics and Perception of Speech*. Amazon: Lippincott Williams and Wilkins.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., and Mehler, J. (1999). *Epenthetic Vowels in Japanese: A Perceptual Illusion?* Journal of Experimental Psychology: Human Perception and Performance, 25, 1568–1578.

Hillenbrand, J. & Clark, M. & Houde, R. (2001). *Some Effects of Duration on Vowel Recognition.* The Journal of the Acoustical Society of America, 108, 3013-22. [Online] Available: https://doi.org/10.1121/1.1323463

Hogan, J. T., & Rozsypal, A. J. (1980). *Evaluation of Vowel Duration as a Cue for the Voicing Distinction in the Following Word-Final Consonant*. The Journal of the Acoustical

Society of America, 67, 1764–1771. [Online] Available: https://doi.org/10.1121/1.384304

Ingham, B. (1994). *Najdi Arabic: Central Arabian*. Georgetown: John Benjamins.

Liberman, A.M. (1996). *Speech A Special Code*. Cambridge: MA MIT Press.

Luo, J., Li, V., & Mok, P. (2019). *The Perception of Cantonese Vowel Length Contrast by Mandarin Speakers.* Language and Speech, 63, 635-659.

Mitleb, F. (1984). *Vowel Length Contrast in Arabic and English: a Spectrographic Test.* Journal of Phonetics, 12, 229-235.

Mok, P. (2011). Effects of Vowel Duration and Vowel Quality on Vowel-to-vowel Coarticulation. Language and Speech, 54. 527-45. [Online] Available: https://doi.org/10.1177/0023830911404961

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

Repp, B.H. (1984). *Categorical perception: Issues, methods, findings.* In N.J. Lass (ed.), Speech and Language: Advances in Basic Research and Practice, Vol. 10 (pp. 243–335). New York: Academic Press.

Tsukada, K. (2009). *An Acoustic Comparison of Vowel Length Contrasts in Arabic, Japanese and Thai: Durational and Spectral Data.* International Journal of Asian Language Processing, 19, 127-138.

Ueyama, M. (2003). *Duration and Quality in the Production of the Vowel Length Contrast in L2 English and L2 Japanese.* [Online] Available: https://doi.org/10.13140/2.1.2787.9680

Watson, C., & Harrington, J. (1999). *Acoustic Evidence for Dynamic Formant Trajectories in Australian English vowels.* The Journal of the Acoustical Society of America, 106 1, 458-68. [Online] Available: https://doi.org/10.1121/1.427069