Voice Onset Time of Initial Stops in Najdi Arabic

Mohammad, F., Aljutaily[1], & Bader, Y., Alharbi[2]

[1] Department of English Language & Translation, College of Arabic Language & Social Studies, Qassim University, Buraidah, Saudi Arabia

[2] Department of English Language & Translation, College of Arabic Language & Social Studies, Qassim University, Buraidah, Saudi Arabia

Correspondence: Mohammad, F., Aljutaily, Qassim, Buraidah City, Saudi Arabia. E-mail: mjtiely@qu.edu.sa

Received: May 26, 2022 Accepted: July 8, 2022 Online Published: July 11, 2022

doi:10.5430/wjel.v12n6p83 URL: https://doi.org/10.5430/wjel.v12n6p83

Abstract

The voice onset time (VOT), which indicates the durational interval from the stop release to the start of the voicing of the following vowel, has long been a prominent research topic in the field of acoustic phonetics. It has shown to be a reliable universal acoustic parameter in discriminating the stop consonants in numerous languages, like English, German, Korean, and various Arabic varieties. It also has been claimed that other factors, such as place of articulation, gender, and vocalic contexts, may affect the productions of the VOTs of the stop categories. Therefore, the present study aims to examine the VOT values of the Najdi Arabic (NA) stops, particularly the plain stop consonants /b, d, g, t, k/, and the potential effects of the place of articulation, gender, and vocalic contexts on their production. Sixty male and female native NA speakers were recruited in this study. The findings revealed that NA has two types of VOTs: positive VOTs with the voiceless stops and negative VOTs with the voiced stops. They also showed that the VOTs of the NA stops become longer when the articulators move back to the oral cavity and shorter in the front cavity. Lastly, the findings reported that the VOTs with short vowels are significantly shorter than the VOTs with long vowels and that gender does not affect the VOTs of the voiced stops, but does affect the VOTs of their voiceless counterparts. Hopefully, the conclusions reached in this study contribute to a better understanding of this interesting acoustic parameter and to the less acoustic literature on Arabic and its various varieties.

Keywords: VOT, Najdi Arabic, stops, place of articulation, gender, vowel contexts

1. Introduction

The contrast in voicing between stop consonants has been a main research topic in the fields of phonetics and phonology over the past few decades (Khattab, 2002). It also attracts researchers from other related linguistic fields, including language typology (Lisker and Abramson, 1964), psychology (Papanicolaou et al., 2003), and language acquisition (Kehoe et al., 2004). Linguists normally implement the dimension of voicing to categorize consonants as either voiced or unvoiced. Based on articulatory phonetics, the existence of glottal pulses during the closure represents voiced consonants, whereas the absence of pulses represents unvoiced consonants. From acoustic perspective, the voicing can be detected visually in both spectrograms with low frequency components and a blank in the closure interval. However, it has been shown that these cues are not sufficient to detect and classify the stop consonants. Therefore, using another acoustic cue, namely the voice onset time (VOT)[1], together with voicing can adequately classify the set of stops.

In the literature, several studies have been conducted to examine the VOTs of the stop consonants in various languages, such as Korean and Chinese (Lisker & Abramson, 1964), Hebrew (Raphael & Tobin, 1983), German (Jessen, 1999), Swahili (Alsamaani, 2021), English (Öğüt et al., 2006), and various Arabic varieties, such as Modern Standard Arabic (MSA), Iraqi Arabic, Lebanese Arabic, Jordanian Arabic, Qatari Arabic, and Saudi Arabic (AlAni, 1970; AlDahri, 2012; AlDahri & Alotaibi, 2010; Alghamdi, 2006; Flege & Port, 1981; Khattab, 2002; Kulikov, 2021; Mitleb, 2001). However, this area of research has been largely overlooked in the dialect of Najdi Arabic (NA). Therefore, the main aim of the current study is to analyze the average durations of the VOT values of the NA stops in onset positions, particularly the five plain stop consonants /b, d, g, t, k/. In addition, some of the earlier studies (e.g., Öğüt et al. 2006; Peng et al., 2014) have indicated that some other factors, such as place of articulation, gender, and vocalic contexts (e.g., short vs. long vowels), may affect the VOT values. Thus, the other aims of this study are to examine the possible effects of place of articulation, vocalic contexts, and gender on the VOT values of NA stops. The main research questions that the current study seeks to answer are stated below.

(i) What are the average durations of the VOT values of the NA stops /b, d, g, t, k, /?

(ii) Does the place of articulation have any effect on the VOT values of the NA stops?

(iii) Does gender play any role on the VOT values of the NA stops?

(iv) Does the vocalic context (viz., vowel length) affect the production of the VOTs of the NA stops?

Three hypotheses underlie this study. It is strongly predicted that the average VOT values of the stops produced further in the oral tract tend to be longer than those produced in front of the tract. It is also predicted that both gender and vowel length may play significant roles in the productions of the VOTs. In other words, we expect that female speakers produce longer VOT values than male speakers and that the stops in the environment of long vowels display longer VOT values than those in the context of short vowels.

The rest of the paper proceeds as follows. Section (2) provides short background on the notions of the VOT and the stop consonants in NA. Section (3) reviews the most relevant studies that have examined the VOTs of stop categories in several languages. Section (4) lays out the method employed in the current study and the procedures of data collection. Section (5) reports the findings of the current study. Section (6) discusses the main findings of the current study and links them with the past studies. Section (7) concludes the work and suggests avenues for future work.

2. Background

2.1 Voice Onset Time (VOT)

VOT is one of the reliable universal acoustic parameters in distinguishing the phonemic categories of stop consonants in various languages. The term VOT was first introduced by Lisker and Abramson (1964) in their influential study and defined it as the durational interval from the release of the stop consonant to the start of the voicing of the following vowel that reflects laryngeal vibrations. The VOT provides the listener with the required cue to which a plosive sound is being produced (Das & Hansen, 2004). Alghamdi (2006) points out that a speaker cannot have a conscious control to the duration of VOT “since VOT is a result of a complicated timing procedure where the glottis coordinates with supraglottal organs” (p.26). This duration has been calculated in milliseconds (ms) and classified into three categories based on the VOT continuum (Lisker & Abramson, 1964; Öğüt et al., 2006). The first category is the negative VOT or "voicing lead", which means that the vibration of the vocal cords begins before releasing the stop, as occurred in Italian voiced stops with a median measurement of -100 ms. The second category is the "short voicing lag" with a median measurement of 10 ms. This category indicates that both the release of the stop and voicing start simultaneously (e.g., English voiced stops /b, d, g/). The third category is the "long voicing lag" with a median measurement of 75 ms. In this category, the vibration starts after the release of the stop (e.g., English unvoiced stops /p, t, k/).

Despite these facts, the values of VOT cannot generally be observed in a fixed length frame. Peng et al. (2014) states that the VOT magnitude may be influenced by some circumstances, such as speaking rate, age, gender, place of articulation, vowel context, and tone; all of which contribute to the VOT variations across languages and make these factors have different phonetic cues in stop contrast. For example, the various Arabic dialects, like other languages, vary in their VOT values (AlDahri & Alotaibi, 2010). AlDahri (2012) claims that the VOT plays a major role in the perceptual discrimination of the phoneme pairs that occur in the same place of articulation. The VOT is also useful for word segmentation and dialectical or accented variations in speech patterns; thus, the variation in VOT values is language-specific (Khattab, 2002). The length of VOT undergoes a systematic variation across languages in which described as a short value ranging (0 – 35 ms) or described as an aspiration with VOT longer than 40 ms (Kulikov, 2021). For instance, the Hebrew stop consonants have VOT values ranging from 22 to 45 ms and -75 to -90 for /p, t, k/ and /b, d, g/, respectively (Raphael & Tobin, 1983). However, Korean, as reported in Lisker and Abramson (1964), measures the stops /p/ with (18 ms), /t/ with (25 ms), and /k/ with (47 ms) and their aspirated counterparts / p^h, t^h, k^h/ with (91 ms), (94 ms) and (126 ms), respectively. Although numerous studies (e.g., Jesry, 1996; Lisker, 1975) claim that the VOT is the main and reliable cue for differentiating the stops, Khattab (2002) argues that VOT has a language-specific variation.

2.2 Stop Consonants in Najdi Arabic (NA)

NA is a local Arabic dialect spoken in the central regions of Saudi Arabia. It is the native language of around eight million people (Aldwayan, 2008). The NA phonological system includes eight stop phonemes, as illustrated in Table (1). The sounds under investigation in this study are listed in bold.

Table 1. NA stops based on Alharbi & Aljutaily (2020)

	Bilabial	Labiodental	Dental	Alveolar(Emphatic)	Alveo-Palatal	Velar	Uvular	Pharyngeal	Glottal
NA Stops	**b**			**d** **t** (tˤ)		**g** **k**	q		ʔ

As pointed out by Newman (2008), “Arabic is one of only thirty-five languages within UPSID[2] to have stop phonemes in five different places of articulation” (p. 66). It can be observed that NA, like other Arabic dialects, does not have the voiceless bilabial stop /p/ in its phonemic inventory[3]. Næss (2008) states that /p/ exists in specific loanwords borrowed from Persian. Daniels and Kaye (1997) claim that the voiced bilabial stop /b/ is devoiced in Arabic when it is followed by a voiceless sound. For instance, the speakers of Arabic pronounce the voiced bilabial stop in the word /kabs/ ‘push’ as a voiceless bilabial stop. Unlike many languages, NA has two unique phoneme stops, namely the emphatic stop sound /tˤ/ and the voiceless uvular stop /q/. The latter stop /q/ is considered an unstable sound in many Arabic dialects. It could be realized as a voiced velar stop /g/, a glottal stop /ʔ/, or a uvular stop /q/ (Newman, 2008). Ingham (1994) reports that the glottal and uvular stops do not exist in NA except in words borrowed from classical Arabic, such as */qurʔa:n/* ‘Holy Muslim Book’ and */taʔaθar/* ‘affected’. The current study is concerned only with the NA plain stops /b, t, d, k, g/; the other stop categories are beyond the scope of the current study due to the fact that the stops /q, ʔ/ are unstable consonants and appear only in words borrowed from classical Arabic. The emphatic stop /tˤ/ is also excluded due to its secondary articulation which in turn may affect the production of VOT (AlDahri, 2012).

As for vowels, NA has three short and five long vowels. According to Ingham (1994), the short vowels are /i/, /a/, and /u/, while the long vowels are /iː/, /aː/, /uː/, /eː/, and /oː/. Watson (2002) points out that all modern dialects of Arabic have at least these three long vowels, /i:/, /a:/, and /u:/. In this study, we also examine the possible effect of vocalic length on the production of VOTs of the stop consonants, particularly in the environment of the vowels /a, u/ and their long counterparts /a:, u:/.

3. Literature Review

In the literature, there were numerous studies examined the VOT in various languages. For instance, Lisker and Abramson (1964) investigated the VOT of initial stops in a number of languages, including Dutch, Spanish, Hungarian, Tamil, Cantonese, English, Eastern Armenian, Thai, Korean, Hindi, and Marathi. They, in particular, examined how the VOTs of the binary oppositions of the stops in the languages under investigation can separate the voiced stops from their voiceless counterparts. In their study, the participants were asked to utter twice two sentences in their languages that include the stop in initial prevocalic positions. By using a wide band spectrogram, they measured the VOT of each stop, and then classified the target languages into three categories based on the number of stops occurred in each language: two-category language (e.g., English, Dutch, Spanish, Hungarian, Tamil, and Cantonese), three-category language (e.g., Eastern Armenian, Thai, and Korean), and four-category language (Hindi and Marathi). They also suggested three VOT categories depending on the VOT measurements. As described in subsection (2.1), the first category is the voicing lead, which indicates that the vibration of the vocal cords begins before releasing the stop with median value -100 ms, the second category is the long- lag, which means that the vibration starts after the release of the stop with median value 75 ms, and the third category is the short-lag, which indicates that both the release of the stop and voicing start simultaneously with median value 10 ms. Based on this categorization, languages, like English, Korean, and Eastern Arminian, followed all three VOT continuum patterns. For instance, the unaspirated English stops [p, t, k] display VOT values ranging from 0 to 25 ms (short-lag), the unaspirated voiced stops [b, d, g] show values from -125 to -75 ms (voicing lead), and the aspirated stops display values from 60 ms to 100 ms (long-lag). In contrast, Lisker and Abramson argued that the VOT alone may not be sufficient for categorizing the stop classes, at least in some languages (e.g., Hindi and Marathi). In other words, Hindi contains aspirated voiced stops, suggesting that it has four categories: voiced unaspirated, voiced aspirated, unvoiced unaspirated, and unvoiced aspirated. Another significant finding of their study was that Korean displays larger average values in all the three patterns (i.e., unaspirated, slightly aspirated, and heavily aspirated). Lastly, Lisker and Abramson did not report any effect of the place of articulation on the VOT values, even though the velar stops did display longer VOTs. They also argued that the vocalic context did not show any significant influence on the VOT values.

Öğüt et al. (2006) examined the average VOT values for the Turkish stop consonants /p, b, t, d, k, g/ in different vocalic contexts. The participants were asked to read Turkish words containing the target stops in an initial-syllable in different vocalic contexts. The words were presented in carrier sentences with three repetitions for each. The findings showed that Turkish, like English and Mandarin, demonstrated positive VOT values for the voiceless stops and negative values for the voiced counterparts. The average VOT values are (41, -66 ms) for the bilabial stops /p, b/, (50, -53 ms) for the alveolar stops /t, d/, and (69, -10 ms) for the velar stops /k, g/. The sound /k/ has the longest duration; this indicates the significant effect of the place of articulation on the VOT values. However, the findings reported no influence of the vocalic contexts or gender on the VOT values.

Peng et al. (2014) investigated the potential effect of gender on the VOT values in two varieties, namely Mandarin and Hakka. Given the facts that male and female speakers have different speaking rates and that female speakers are more likely to have a careful manner of production than male speakers, they expected a variation in length. Specifically, Peng et al. examined the VOT values of the voiceless stops [p, t, k] and their aspirated counterparts [p^h, t^h, k^h] positioned word-initially with different vocalic environments [i, u, a]. These voiceless stops are contrasted in Mandarin and Hakka by the aspiration. Forty-nine male and female native speakers from both varieties were asked to read words containing the target stops in a normal volume. In contrast to previous studies that used different measurement means for categorizing stops (e.g., obtaining the VOTs by measuring the temporal between the release of a stop and the beginning of a vowel first formant or second formant or by measuring the abrupt spike that signals the release and the beginning of the oscillating motion), Peng et al. used the Praat software to measure the VOT values with the primary reference to the waveform in the spectrogram. Although some studies (e.g., Öğüt et al., 2006; Ryalls et al., 2004) have claimed that gender did not show any significant influence on the VOTs, Peng et al. argued the contrary. This significant result is in line with the study of Whiteside and Marshall (2001), which also reported gender effect on the value of VOT. The findings of Peng et al.'s study showed that the male Mandarin and Hakka participants produced the unaspirated stops with mean VOT at 22 ms with longer duration than female participants at 18 ms; but a reverse realization happened with aspirated stops in which the female participants tended to have longer VOTs than the male participants. As for the effect of gender on the place of articulation, the findings showed a significant difference with only Mandarin labials. The male Mandarin speakers obtained a mean value with labials at 50 ms with SD (11.46) compared to the female speakers at 53 ms with SD (11.13). Peng et al. suggested a number of factors that may result in significant VOT differences between male and female speakers, including the speech materials (i.e., isolated words or sentences), the speech rate (i.e., males speak faster than females in spontaneous speech), the speaking style, and the length of the vocal cords. Along the same line, Alsamaani (2021) has also studied the gender differences in Swahili and found out that the speakers of the language implemented different strategies when producing the VOTs. L1 speakers have longer VOTs for males than females but with no significant difference. However, L2 speakers have significantly longer VOTs for females than males.

Within the Arabic context, numerous studies have examined the VOT patterns in various Arabic varieties. One of the earlier studies was conducted by AlAni (1970) on Iraqi Arabic (IA), which analyzed the durations of voicing and release in the production of the voiceless stops [t, k, q] and their voiced counterparts [b, d]. AlAni reported that the voiced bar in the spectrogram varies in its duration based on the phonetic context in which it occurs (initially, medially, or finally). That is, the voiced stops measure 130 to 150 ms word-initially, 50 to 60 word-medially, and 300 to 350 word-finally. As for the stop release, AlAni reported that the voiceless stops obtain longer duration compared to the voiced stops in which they aspirated with different length: 30 to 60 ms for the voiceless stops word-initially, 110 to 130 ms word-medially, and 300 to 350 ms word-finally.

Alghamdi (2006) also examined the VOTs of the Saudi Arabic (SA) voiceless stops /t, k, $t^ʕ$/. In his study, participants read a list of words containing the target stops in the carrier CV (target stop) VC with six repetitions where C is /z/ and V is /a/, such as /za (k, t, $t^ʕ$) az/. The most significant result of his study was that SA speakers produced both VOT and closure durations with different measurements. To illustrate, the SA speakers produced the stops with positive VOT values ranging from (21 to 78 ms) for /t/, (21 to 55 ms) for /k/, and (14 to 33 ms) for /$t^ʕ$/. On the other hand, they produced the closure duration with values ranging from (51 to 81 ms) for /t/, (59 to 84 ms) for /k/, and (62 to 93 ms) for /$t^ʕ$/. In general, the SA speakers demonstrated different measurements with the production of VOTs. However, Alghamdi's study discussed only the voiceless stops without taking their voiced counterparts into consideration as the current study does. It is useful to examine all the plain stops to provide a thorough analysis of the VOT features in SA so that we can categorize the types of VOTs (i.e., positive /negative VOTs) that SA, or more specifically NA, has.

Another study within the Arabic context was conducted by Mitleb (2001) to examine the VOT values of the Jordanian Arabic (JA) stops /t, d, k, g/. In his study, four male JA speakers were asked to read a list of words containing the target stops in initial syllables and in the environments of the short and long low vowels /a, a:/. The main result confirmed the conclusion obtained in other studies (e.g., Kulikov, 2021; Lisker & Abramson, 1964) in which the VOT values can be used to differentiate the voiced stops from their unvoiced counterparts. The results showed that the JA stops have different measurements depending on the following vowels. The voiceless stop /t/ measures (37 ms) with a short vowel and (46 ms) with a long vowel, while the voiced stop /d/ measures (10 ms) in the environment of a short vowel and (23 ms) with a long vowel. The velar stops /k/ and /g/ have respectively the VOT values as (59 ms) and (15 ms) in the contexts of the short vowel and (60 ms) and (20 ms) in the contexts of the long vowel. Given these results, Mitleb argued that the VOT values correlate to voice contrast and they vary depending on the length contrast of the following vowel. As for the effect of the place of articulation on the VOT values, his results were consistent with the prevailing concept in the literature and the VOT values in both vowel contexts (short vs. long) showed a significant difference across the place of the articulation and the values increased as the place of articulation goes further in the oral tract.

Furthermore, AlDahri and Alotaibi (2010) investigated the VOT values of the Modern Standard Arabic (MSA) stops /b, k/ and compared them with those reported in other languages (e.g., Dutch, Tamil, and English) and other Arabic dialects (e.g., JA and SA). In their study, sixty male and female participants were asked to produce eight words containing the target stops in medial position, preceded and followed by the low vowel /a/, with five repetitions. The most significant finding was that both of the MSA stops /b/ and /k/ obtained positive VOT values regardless of their voicing. The value of the MSA voiced bilabial stop /b/ ranges between 10 and 25 ms, whereas the voiceless velar stop /k/ ranges between 34 and 80 ms. This result contrasted with the results reported for other languages (e.g., English, Dutch, Tamil), which display negative VOT values, specifically with the production of voiced stops. The results also revealed that there were some variations in the production of the VOT of the MSA voiceless velar stop /k/. However, the values of VOT vary from language to language even among the dialects of a single language. For instance, in comparing the values that represent Arabic dialects (i.e., SA and JA dialects), they found a noticeable difference in which the stop /k/ measured (60 ms) for JA dialect, (42 ms) for SA dialect, and (52 ms) for MSA. Overall, this study is inadequate since it did not consider the other stop categories (e.g., /t, d, g/. Therefore, a comprehensive study that examines the VOT values of all Arabic stops is required to enrich the phonetic literature, and more specifically the Arabic literature.

Lastly, Kulikov (2021) examined the VOT values of the alveolar stops /t, d/ in the dialect of Qatari Arabic (QA). Eight female students at Qatar University, who are native QA speakers, were asked to read 25 words having the target stops in initial positions, with six repetitions for each word. The results showed that the voiceless stop /t/ has the longest VOT value (M = 51 ms), while the voiced stop /d/ has the value of (M = -45 ms). This result suggests that the VOT can be considered as the most invariant cue for identifying the voicing of alveolar stops. Nevertheless, the findings of Kulikov’s study cannot be generalized to the various Arabic varieties due to the small number and gender of the participants. The size of the vocal tract varies for males and females, thus, differences in acoustic values are predictable. Accordingly, we believe that an alternative study with a larger number of participants and with different genders may provide more precise findings on the acoustic values of VOTs in Arabic and its various dialects.

To summarize, this section discussed several relevant studies that have examined the VOTs values of stop categories in Arabic and other languages. It also discussed a number of studies that have investigated the possible effects of some factors on the production of VOTs, such as gender, vocalic contexts, and place of articulation.

4. Methodology

4.1 Participants

This acoustic study was based on 60 native speakers of NA: Male (n = 30) and Female (n = 30). All the participants were born and raised in the region of Qassim in Saudi Arabic. NA is their native language. All participants were aged 19-30 years old (median 22 years). They voluntarily participated in this experiment and none of them reported any speech or hearing impairments.

4.2 Stimuli and Procedure

The data were collected by creating a list of forty NA CVC or CVCVC meaningful words incorporated in simple carrier sentences having the target stop categories in word-initial positions. The target stops, /b, t, d, k, g/, were positioned before short vowels /a, u/ (e.g., /gabur/ ‘grave and /kum/ ‘sleeve’) and their long counterparts /a:, u:/ (e.g., /ga;ʕ/ ‘bottom’ and /ku:x/ ‘cottage’). See Appendix (A) for the entire list of words used in this study. The words were also listed in Arabic orthography.

The participants were chosen based on to two crucial criteria: speaking NA natively and living in the region of Qassim, which is part of the large region of Najd. Due to Covid-19, the participants were met both individually and virtually via Zoom and their productions were recorded by using Praat Recording at a sampling rate of 44.1 kHz.

Prior to reading the sentences that include the target words, the participants were given clear instructions on how to perform the task. Then, they were asked to read the carrier sentences (*say* ***(the word)*** *again*) casually as normal speech in a very quiet environment. Each sentence was repeated twice, and the tokens that include the same consonantal stops were presented randomly in which the participants had the same randomization of words.

4.3 Segmentation of VOTs and Acoustic Analysis

The study employed an acoustic analysis to analyze the VOTs, using the relevant acoustic measurement (i.e., VOT) to classify the consonantal stops of NA. To accomplish this analysis, we first recorded all the target stops produced by the participants by using Praat. Then, we divided the recording files for each participant into smaller files, each of which consists of a word with the target stop which was extracted from the carrier sentences, so as to create a TextGrid and start the acoustic analysis. The values of the acoustic correlates were taken automatically using a Praat script. In case that an unexpected measurement was occurred (i.e., longer VOT of the voiced stop than the voiceless stop or in case the measurement wasn't detected by the software and so on), we double checked the file and corrected it manually. Lastly, the values were exported to an Excel spreadsheet for statistical procedures. The total amount of token was 4800 (i.e., 5 stop consonants X 8 different words with the same target sound X 60 participants X 2 repetitions for each).

Praat was used to measure the acoustic correlates that are related to the analysis of VOTs of NA stops. These measurements included the VOT of each consonantal stop. The VOT was identified as the period from the release of the occlusion to the onset of the following vowel and measured in milliseconds (ms), as demonstrated in Figure (1).

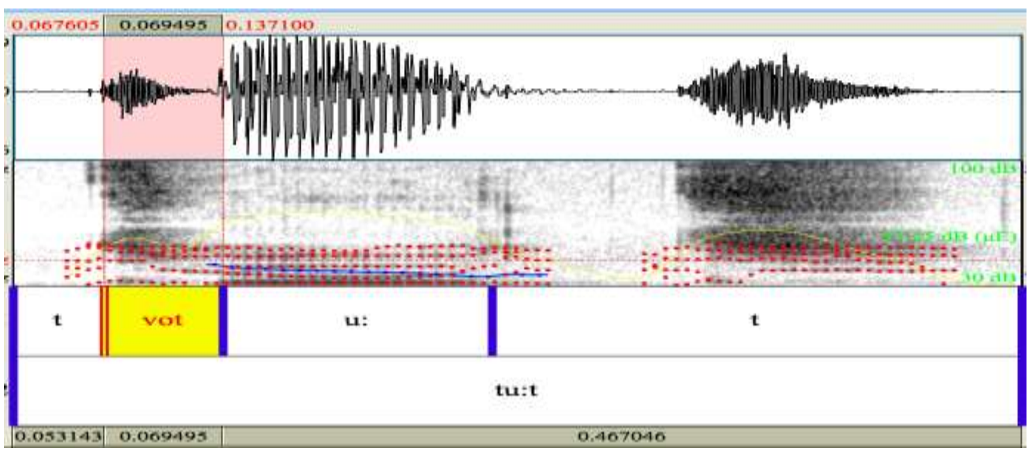

Figure 1. Spectrogram of the VOT of the stop /t/ in the word /*tu:t*/ 'berry' measured (69 ms)

4.4 Statistical Procedure

All statistics were conducted using IBM SPSS Statistics for Mac, Version 26.0. The statistical analysis was based on four 3-way repeated measures mixed between/within ANOVA tests with a significant level set as p <.05 to specify if there was a significant interaction between place of articulation, vowel length, and gender. For example, a test for the voiceless stops followed by the vocalic contexts /a, aa/ designed as 2x2x2, the voiced stops followed by the vocalic contexts /a, a:/ designed as 3x2x2, the voiceless stops followed by the vocalic contexts /u, uu/ designed as 2x2x2, and the voiced stops followed by the vocalic contexts /u, uu/ designed as 3x2x2. Each test investigated the main effects of: within-subject effects place of articulation (labial (if any), alveolar, and velar), vowel length (short and long), between-subject effects gender (male and female). We also tested the interaction effects of place of articulation *x* vowel length, place of articulation *x* gender, vowel length *x* gender, and place of articulation *x* vowel length *x* gender.

5. Findings

As explicitly stated in section (1), the main goals of the current study are to analyze the average durations of the VOT values of the NA stop categories /b, d, g, t, k/ and to examine the possible effects of vocalic contexts, place of articulation, and gender on the VOT values of these five categories. This section about findings is divided into two main subsections. Subsection (5.1) lays out the overall results of the VOT values of the NA voiceless stops in both vocalic contexts (short vowels /a, u/ and long vowels /a:, u:/). Subsection (5.2) presents the VOT values of the NA voiced stops in the vocalic contexts of (short vowels /a, u/ and long vowels /a:, u:/).

5.1 VOTs of the NA Voiceless Stops

The results demonstrated that both NA male and female speakers produced positive VOT values in realizing the initial voiceless stops. Table (2) displayed descriptive statistics for the mean VOT values of the voiceless stops for both genders in the contexts of (short vowels /a, u/ and long vowels /a:, u:/).

Table 2. Mean values and (standard deviation) of VOTs of voiceless stops in the contexts of short and long vowels by female and male participants

	Alveolar		Velar		Alveolar		Velar	
Voiceless stops	ta	ta:	ka	ka:	tu	tu:	ku	ku:
Female (*n*=30)	54 ms (.0099)	62 ms (.0095)	60 ms (.0106)	71 ms (.0119)	50 ms (.0111)	53 ms (.0113)	65 ms (.0111)	77 ms (.0138)
Male (*n*=30)	49 ms (.0069)	58 ms (.0090)	56 ms (.0081)	68 ms (.0081)	45 ms (.0071)	48 ms (.0076)	60 ms (.0071)	74 ms (.0099)

As indicated in Table (2), the mean VOT values of the voiceless stops for both genders increased in the contexts of long vowels compared to the contexts of short vowels. Moreover, the voiceless stops produced further in the oral tract (i.e., velar) tended to be longer in duration than those produced in the front of the oral tract (i.e., alveolar). Therefore, in performing the repeated measures ANOVA for determining the main effect and interaction between the independent variables (i.e., place of articulation, vowel length, and gender) on the dependent variable (i.e., VOT value), there was no a statistically significant interaction between the place of articulation, the vowel length, and the gender on the VOT values of voiceless stops in both vowel contexts, (F (1, 58) = 001^{b} , P = .979 and F (1, 58) = 356^{b}, P = .553, respectively). However, there was a significant interaction only between the place of articulation and vowel length on the VOTs in both vocalic contexts (i.e., (short vowels /a, u/ and long vowels /a:, u:/). Consider Table (3).

Table 3. Interaction effect between place of articulation and vowel length on VOTs of voiceless stops

		F	*df*	Sig.	Partial Eta Squared
Place & vowel length with /a,a:/	Wilks' Lambda	6.143^{b}	1.000	.016	.096
Place & vowel length with /u, u:/		178.723^{b}	1.000	.000	.755

Table (3) displayed a significant interaction between the place of articulation and vowel length on the VOTs of the voiceless stops in both vocalic contexts. Therefore, the place of articulation causes a significant difference on the VOT values with vowel length. Consider Table (4), which explains the means of the VOTs in the interaction effect.

Table 4. Means of the VOTs in the interaction between the place of articulation and vowel length for both genders

				95 % Confidence Interval	
Place of Articulation	Vowel length	Mean	Std. Error	Lower Bound	Upper Bound
Alveolar	Long	.061	.001	.058	.063
	Short	.052	.001	.050	.055
Velar	Long	.070	.001	.067	.073
	Short	.058	.001	.056	.061
Alveolar	Long	.051	.001	.048	.053
	Short	.051	.001	.048	.053
Velar	Long	.076	.002	.073	.079
	Short	.063	.001	.061	.065

As observed in Table (4), the velar stops had higher VOT values than the alveolar stops in both contexts. It appears that NA has longer VOTs for the voiceless stops when followed by long vowels /a: or u:/ than those with short vowels /a or u/ and has longer VOTs for velar stops than for alveolar stops. Figure (2) demonstrated the overall differences in the means of the VOTs of voiceless stops in the vocalic contexts and place of articulation by all participants.

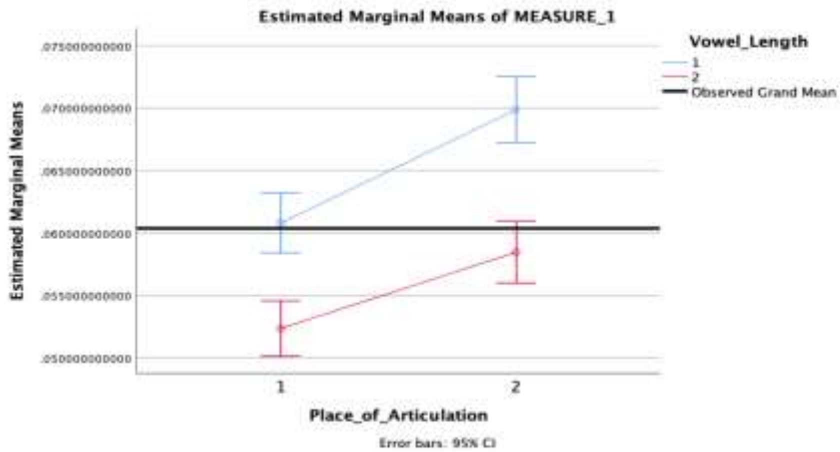

Figure 2. Means of VOTs of voiceless stops with long vowel (blue color) and short vowel (red color) and means of VOTs with alveolar (1) and velar (2)

As depicted in the figure, the top and bottom lines represented the long and short vowels, respectively. The values are longer with long vowels compared to the short ones and the lines rise as moved towards the velar stop.

As for gender influence on the VOT values, Table (5) illustrated the Tests of Between-Subjects Effects of the overall VOTs by female and male participants across the investigated voiceless stops in both vocalic contexts.

Table 5. Tests of Between-Subjects Effects for gender in the production of VOTs of voiceless stops

Tests of Between-Subjects Effects					
Source	*df*	Mean Square	F	Sig.	Partial Eta Squared
Gender /a,a:/	1	.001	3.801	.056	.062
Gender /u,u:/	1	.001	4.265	.043	.068

As shown in Table (5), gender affected the VOT values of the voiceless stops in the vocalic context of (short vowels /a, u/ and long vowels /a:, u:/). The means of the VOT values showed significant statistical differences of both genders. Table (2) above revealed the trend of an increase in the VOT values among the female participants compared to the male participants in the VOTs of the voiceless stops, indicating that gender did influence the VOTs of the voiceless stops in both vocalic contexts.

5.2 VOTs of the NA Voiced Stops

Table (6) demonstrated the means of the VOT values of the voiced stops in the contexts of (short vowels /a, u/ and long vowels /a:, u:/). for both female and male participants together with the standard deviation.

Table 6. Descriptive statistics for the VOTs of voiced stops in the contexts of short vowels /a, u/ and long vowels /a:, u:/ for all participants

	Labial		Alveolar		Velar	
Voiced stops	ba	ba:	da	da:	ga	ga:
Female (*n*=30)	-16 ms (.0030)	-20 ms (.0048)	-20 ms (.0036)	-24 ms (.0062)	-27 ms (.0052)	-31 ms (.0061)
Male (*n*=30)	-15 ms (.0024)	-19 ms (.0029)	-19 ms (.0040)	-24 ms (.0049)	-26 ms (.0049)	-31 ms (.0055)
	Labial		Alveolar		Velar	
Voiced stops	bu	bu:	du	du:	gu	gu:
Female (*n*=30)	-20 ms (.0041)	-24 ms (.0043)	-20 ms (.0043)	-27 ms (.0064)	-29 ms (.0056)	-34 ms (.0076)
Male (*n*=30)	-18 ms (.0027)	-22 ms (.0036)	-20 ms (.0040)	-25 ms (.0055)	-27 ms (.0054)	-35 ms (.0070)

Based on Table (6), it appears that the NA voiced stops are produced with shorter VOT values compared to their voiceless counterparts. The results also reported the same trend obtained with the voiceless stops in that the voiced labial stop has shorter VOT value than the voiced alveolar stop and the voiced alveolar stop is shorter than the voiced velar stop. Furthermore, the voiced stops with long vowels have longer VOT values than those with short vowels. Similar to English, the results showed that NA has negative VOT values (i.e., prevoiced) with the voiced stops; the voicing starts before releasing the stop, as demonstrated in figure (3).

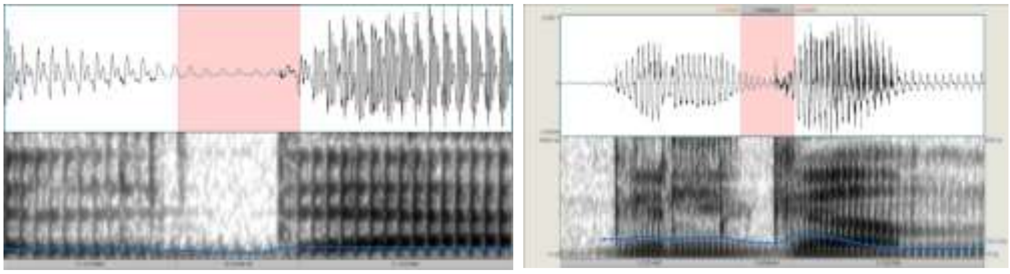

Figure 3. Spectrograms for prevoiced stops in NA: /d/ (left spectrogram) and /g/ (right spectrogram)

Repeated measures ANOVA was used to establish the main effect and interaction between the independent variables (i.e., place of articulation, vowel length, and gender) on the dependent variable (i.e., VOT value). Consider Table (7).

Table 7. Tests for the main effect and interaction of the variables of voiced stops in both vocalic contexts by all participants

Tests of Within-Subjects Effects					
Effect	Test	*df*	F	Sig.	Partial Eta. Squared
Place effect with /a,a:/	Greenhouse Geisser	1.693	224.048	.000	.794
Vowel length with /a,a:/	Sphericity Assumed	1	151.113	.000	.723
Place * Vowel length with /a,a:/	Sphericity Assumed	2	. 050	.951	.001
Place * Vowel length * Gender with /a,a:/	Wilks' Lambda	2.000	1.066[b]	.351	.036
Place effect with /u,u:/	Sphericity Assumed	2	142.211	.000	.710
Vowel length with /u,u:/	Sphericity Assumed	1	250.262	.000	.812
Place * Vowel length with /u,u:/	Sphericity Assumed	2	5.360	.006	.085
Place * Vowel length * Gender with /u,u:/	Wilks' Lambda	2.000	2.890[b]	.064	.092

Based on Table (7), the test showed no statistically significant interaction between the place of articulation, vowel length, and gender on the VOT values of the voiced stops in both vowel contexts (i.e., (short vowels /a, u/ and long vowels /a:, u:/). However, there existed a significant interaction between the place of articulation and vowel length only on the VOTs in the vocalic context of (short and long vowels /u, u:/). That is, there was an interaction between the place of articulation and the length of vowels in which the velar stops had higher VOT values than the bilabial and alveolar stops. The stops with the long vowels showed longer VOTs than those with the short vowels. As for the effect of the place of articulation or vowel length on the VOTs, the voiced stops in the vocalic contexts /a, a:/ showed a high significant effect on the VOT values across the places of articulation (i.e., bilabial, alveolar and velar) having the means (.018, .022, .029), respectively; while with the long and short vowel displayed the means (.025, .021), respectively.

With respect to the gender effect on the VOTs of the voiced stops, the Pairwise Comparisons of the overall VOTs by female and male participants among the voiced stops in both vocalic contexts didn't show any statistically significant effect of gender on the means of the VOT values, (P= .451, 95% C.I. = -.001, .002) for the voiced stops with the short and long vowels /a, a:/, and (P= .350, 95% C.I. = -.001, .003) for the voiced stops with the short and long vowels /u, u:/. The means of the VOT values of the voiced stops with the short and long vowels /a, a:/ was (.024 ms) for the female participants and (.023 ms) for the male participants, whereas the means of the VOT values of the voiced stops with the short and long vowels /u, u:/ was (.026 ms) for the female participants and (.025 ms) for the male participants. Since both genders had almost similar means, this result indicated that gender did not influences the VOTs of the voiced stops and consequently performed similarly. In the next section, we discussed the main findings reported in this section.

6. Discussion

The current study examined the average durations of the VOT values of the NA stops /b, d, g, t, k/ and the possible effects of the vocalic contexts, place of articulation, and gender on the VOT values of these stop categories. Following the categorization introduced by Lisker and Abramson (1964), it appears that NA belongs to the two-category language (like English, Spanish, and Dutch) in terms of the number of the stop categories occurred in the language, and thus has the two VOT continuum patterns: voicing lead and long lag. From the perspective of articulatory phonetics, the voiced and voiceless stops can be distinguished by the absence and presence of the vocal pulsing, and acoustically by the VOTs, which were considered reliable predictors in categorizing the NA voiced and voiceless stops. Confirming the conclusions obtained in earlier studies (e.g., Kulikov 2021; Lisker & Abramson 1964), all NA speakers realized the voiced and voiceless stops with different VOT patterns: the voiceless stops obtained positive values, whereas the voiced stops obtained negative values in which the voicing starts before the stop release, as demonstrated in figure (3). This finding is in line with numerous studies that reported similar conclusions in other languages (e.g., English and Mandarin (Lisker & Abramson, 1964), Lebanese Arabic (Khattab, 2002), and Turkish (Öğüt et al., 2006)), but in contrast to other earlier studies (e.g., Iraqi Arabic (Alani, 1970), Jordanian Arabic (Mitleb, 2001), and MSA (AlDahri & Alotaibi, 2010), which reported that the stops obtained positive VOT values regardless of their voicing. As pointed out in section (5), the NA voiceless stops described as aspirated with longer VOTs. The VOT durations for the voiceless stops /t, k/ range between (.45 ms - .65 ms) and (.48 ms - .77 ms) in the environments of short and long vowels, respectively. On the other hand, the VOTs of NA voiced stops /b, d, g/ range between (.15 ms - .29 ms) and (.19 ms - .35 ms) in the environments of short and long vowels, respectively. These VOT ranges differ from those in other Arabic varieties (e.g., JA, IA, MSA) as each variety obtained different values; this conclusion supports the fact that the VOT values undergo a systematic variation across languages and the variation in the VOT is dialect/language-specific (Khattab, 2002).

Regarding the effect of articulation place on the VOTs of the NA stops, the findings confirmed that both place of articulation and VOTs were related. Previous studies (e.g., Mitleb 2001; Öğüt et al., 2006) have reported that the VOT values increased as the place of articulation goes further in the oral tract. Similarly, our findings confirmed that there was a significant main effect of the place of articulation on the VOTs of the NA stops. The VOT was consistently longer for velars than for alveolars and labials. This cross-linguistic pattern may be attributed to some physiological/aerodynamic features. Brinca et al. (2016) state that "the volume of the cavity in front of the point of constriction is greater in velar stops than in bilabial and dental stops, and this causes a greater obstruction to the release of the pressure behind the velar stop" (p. 141). Therefore, our findings are incompatible with the conclusion reported in the study of Lisker and Abramson (1964) but are in line with the studies of Mitleb (2001) and Öğüt et al. (2006), which both reported that the bilabials have shorter VOTs than the alveolars and the alveolars have shorter VOTs than the velars.

With respect to the question of whether the vowel length (long vs. short) affects the VOTs of the NA stops, our findings indicated that vowel length was another factor playing a significant role in the VOT values of the NA stops. It was noticeable that there was a significant difference between the stops produced in the contexts of short vowels and the stops produced in the contexts of long vowels. In other words, the VOT values of the stops with long vowels were significantly longer than those of the stops with short vowels. We assume that the effect of vowel length on the VOTs of NA stops may result from the fact that vowel lengthening in NA is phonemic and consequently the NA speakers realized the VOTs differently depending on the vocalic environments. Generally speaking, these findings are compatible with the study of Mitleb (2001), which concluded that the VOTs vary due to the length contrast of the following vowels. Nonetheless, these findings are incompatible with the results reported in the studies of Lisker and Abramson (1964) and Öğüt et al. (2006), which indicated that the vocalic contexts did not have a significant effect on the VOTs.

Finally, one of the goals of this study was to test the effect of gender on the VOT values of the NA stops. Previous studies (e.g., Whiteside & Marshall 2001; Robb et al. 2005; Peng et al. 2014) have concluded the gender effect on the VOT values; female participants realized the VOTs with higher values than male participants, and this is due to the different vocal anatomy between males and females. Although our findings indicated that the overall performance showed the trend of an increase in the VOT values among the female participants compared to the male participants, the findings, more particularly, showed a significant main effect of gender only on the VOTs with the NA voiceless stops, but not with the voiced stops. Both genders realized the VOTs of the NA voiced stops similarly; the mean VOT values did not differ significantly between male and female participants. Thus, it could be concluded that results of the gender effect in the current study did not fully support our hypothesis, which states that female participants may realize longer VOT values than male participants. These findings about gender effect are inconsistent with numerous previous studies (e.g., Whiteside & Marshall 2001; Robb et al. 2005; Peng et al. 2014), which proved the gender effect on the VOT values due to physiological differences between males and females, but are consistent with the studies of Öğüt et al. (2006), Morris et al. (2008), and Ryalls et al. (2004), which all concluded that gender did not influence the VOTs of voiced stops at least in the current study.

7. Conclusions

This study examined the VOT values of the NA stops /b, d, g, t, k/ and the possible effects of the articulation place, gender, and vocalic contexts on their production. Sixty male and female native NA speakers participated in the study. They produced a list of forty NA sentences including the target stops in word-initial positions in carrier sentences. The results revealed that NA has two types of VOTs: positive VOTs with the voiceless stops and negative VOTs with the voiced stops. The former is a very common pattern across languages, whereas the latter has not been attested in the various Arabic varieties except in Lebanese Arabic. The results also showed that the VOT values of the NA stops follow the universal acoustic tendency in which they increase as long as the articulators move back to the oral cavity. The NA speakers produced the VOTs of the velar stops with longer durations than the alveolar stops and alveolars longer than the bilabials. It was shown that the VOTs with short vowels are significantly shorter than the VOTs with long vowels and that gender does not affect the VOTs of the NA voiced stops, but does affect the VOTs of their voiceless counterparts. The results obtained in the current study may contribute to the literature on acoustic phonetics by corroborating the assumptions that the VOT can be substantially considered a reliable universal acoustic parameter in discriminating the stop consonants and that other external factors, like place of articulation and vowel contexts, may play significant roles on the production of the VOTs of the consonantal stops.

Due to time and space limitations, three issues have been left for future research. First, there is a need for a further study that examines the VOT of the NA emphatic stop consonant /tˤ/. Phonologically speaking, this category has a secondary place of articulation (see e.g., AlAni 1970; AlDahri 2012; Alharbi and Alammar 2022; Davis 1995; Youssef 2014), which may make its VOT different from the VOTs of its plain counterparts. Second, there is a need for a comprehensive study that examines the VOTs of the NA stop categories in other positions (i.e., word-medially and word-finally) and that test the effect of all NA short and long vowels. Lastly, it will be useful to investigate the possible effect of age on the production of the VOTs of the NA stop consonants. Hopefully, the conclusions reached in this study will serve as a basis for these avenues of future work.

Acknowledgements

We would like to express our sincere thanks and appreciation to our colleague, Dr. Soliman Ismail, for his assistance with statistical consultation on our work. We would also like to thank all the participants, specifically the graduate students at the Department of English Language and Translation, for their active participation and cooperation in collecting the data for this study.

References

Al-Ani, S. (1970). *Arabic phonology: An acoustical and physiological investigation.* The Hague. https://doi.org/10.1515/9783110878769

AlDahri, S. (2012). *A study of voice onset time for Modern Standard Arabic and Classical Arabic* [Conference presentation]. The IEEE International Conference on Signal Processing, Communications and Computing, Hong Kong. https://doi.org/10.1109/ICSPCC.2012.6335582

AlDahri, S. (2012). *The effect of Arabic emphaticness on voice time onset (VOT)* [Conference presentation]. International Conference on Audio, Language and Image Processing, Shanghai, China. https://doi.org/10.1109/ICALIP.2012.6376655

AlDahri, S., & Alotaibi, Y. (2010). A crosslanguage survey of VOT values for stops (/d/, /t/) [Conference presentation]. The IEEE International Conference on Intelligent Computing and Intelligent Systems, Xiamen, China. https://doi.org/10.1109/ICICISYS.2010.5658744

Aldwayan, S. N. (2008). *The acquisition and processing of wh-movement by Najdi learners of English.* PhD Dissertation, University of Kansas.

Alghamdi, M. (2006). Voice print: Voice onset time as a model. *Arab Journal for Security Studies and Training, 21,* 89-118.

Alharbi, B., & Alammar, A. (2022). Emphasis spread in Qassimi Arabic within the underspecification theory. *World Journal of English Language*, *12*(1), 407-418. https://doi.org/10.5430/wjel.v12n1p407

Alharbi, B., & Aljutaily, M. (2020). On the perceptual accuracy of non-native phonemic contrasts: A case study of native Arabic speakers. *Journal of Language and Linguistic Studies, 16*(4), 2003-2023. https://doi.org/10.17263/jlls.851031

Alsamaani, M. (2021). *VOT and F0 in the production and perception of Swahili obstruents: From the island to the coast to the inland region.* Arts.

Brinca, L., Araujo, H., Nogueira, P., & Gil, C. (2016). Voice onset time characteristics of voiceless stops produced by children with European Portuguese as mother tongue. *Ampersand*, *3*, 137-142. https://doi.org/10.1016/j.amper.2016.06.006

Daniels, P., & Kaye, A. (1997). *Phonologies of Asia and Africa: Including the Caucasus.* Warsaw: Eisenbrauns.

Das, S., & Hansen, J. (2004). *Detection of voice onset time (VOT) for unvoiced stops (/p/, /t/, /k/) using the Teager energy operator (TEO) for automatic detection of accented English.* Proceedings of the 6th Nordic Signal Processing Symposium, 2004. NORSIG 2004., 344-347.

Davis, S. (1995). Emphasis spread in Arabic and grounded phonology. *Linguistic Inquiry, 26,* 465-498.

Flege, J., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech, 24*(2), 125-146. https://doi.org/10.1177/002383098102400202

Ingham, B. (1994). *Najdi Arabic: Central arabian.* Amsterdam, Philadelphia: John Benjamins. https://doi.org/10.1075/loall.1

Jesry, M. (1996). *Some cognitively controlled coarticulatory effects in Arabic and English with particular reference to VOT* [Doctoral dissertation, University of Essex].

Jessen, M. (1999). *Phonetics and phonology of tense and lax obstruents in German.* Amsterdam, Philadelphia: John Benjamins. https://doi.org/10.1075/sfsl.44

Jongman, A., Herd, W., Al-Masri, M., Sereno, J., & Combest, S. (2011). Acoustics and perception of emphasis in Urban Jordanian Arabic. *Journal of Phonetics, 39*(1), 85-95. https://doi.org/10.1016/j.wocn.2010.11.007

Kehoe, M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition, 7*(1), 71-88. https://doi.org/10.1017/S1366728904001282

Khattab, G. (2002). VOT production in English and Arabic bilingual and monolingual children. In D. Parkinson & E. Benmamoun (Eds.), *Perspectives on Arabic Linguistics XIII-XIV* (pp. 1–37). Amsterdam, Philadelphia: John Benjamins. https://doi.org/10.1075/cilt.230.03kha

Kulikov, V. (2021). Voice and emphasis in Arabic coronal stops: Evidence for phonological compensation. *Language and Speech,* 1-32. https://doi.org/10.1177/0023830920986821

Lisker, L. (1975). Is it VOT or a first-formant transition detector? *The Journal of the Acoustical Society of America, 57*(6), 1547-1551. https://doi.org/10.1121/1.380602

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical Measurements. *Word, 20*(3), 384-442. https://doi.org/10.1080/00437956.1964.11659830

Mitleb, F. (2001). *Voice onset time of Jordanian Arabic stops.* The 3rd International Conference on Arabic Language Processing (CITALA' 09), Rabat, Morocco.133-135. https://doi.org/10.1121/1.4744787

Morris, R. J., McCrea, C. R., & Herring, K. D. (2008). Voice onset time differences between adult males and females: Isolated syllables. *Journal of phonetics, 36*(2), 308-317. https://doi.org/10.1016/j.wocn.2007.06.003

Næss, U. 2008). *Gulf Pidgin Arabic: Individual strategies or structured variety* [Unpublished master's thesis]. University of Oslo.

Newman, D. (2002). The phonetic status of Arabic within the world's languages: The uniqueness of the lughat al-daad. *Antwerp Papers in Linguistics, 100,* 65-75.

Öğüt, F., Kiliç, M., Engin, E., & Midilli, R. (2006). Voice onset times for Turkish stop consonants. *Speech Communication, 48,* 1094-1099. https://doi.org/10.1016/j.specom.2006.02.003

Papanicolaou, A., Castillo, E., Breier, J., Davis, R., Simos, P., & Diehl, R. (2003). Differential brain activation patterns during perception of voice and tone onset time series: A MEG study. *Neuroimage, 18*(2), 448-459. https://doi.org/10.1016/S1053-8119(02)00020-4

Peng, J., Li-mei, C., & Chia-Cheng, L. (2014). Voice onset time of initial stops in Mandarin and Hakka: Effect of gender. *Taiwan Journal of Linguistics, 12,* 63-80.

Qafisheh, H. (1977). *A short reference grammar of Gulf Arabic.* Tucson: University of Arizona Press.

Raphael, L., & Tobin, Y. (1983). Perceptual and acoustic studies of voice onset time in Hebrew. In A. Cohen & M. van den Broeke (Eds.), *Abstracts of the Tenth International Congress of Phonetic Sciences* (pp. 117-134). Dordrecht: Foris Publications.

Robb, M., Gilbert, H., & Lerman, J. (2005). Influence of gender and environmental setting on voice onset time. *Folia Phoniatrica et Logopaedica, 57*(3), 125-133. https://doi.org/10.1159/000084133

Ryalls, J., Simon, M., & Thomason, J. (2004). Voice onset time production in older Caucasian and African-Americans. *Journal of Multilingual Communication Disorders, 2*(1), 61-67.

https://doi.org/10.1080/1476967031000090980

Shar, S., & Ingram, J. (2010). *Pharyngealization in Assiri Arabic: An acoustic analysis.* Paper presented at the 13th Australasian International Conference on Speech Science and Technology, Melbourne, Australia.

Watson, J. (2002). *The phonology and morphology of Arabic.* Oxford: Oxford University Press.

Whiteside, S., & Marshall, J. (2001). Developmental trends in voice onset time: Some evidence for sex differences. *Phonetica, 58*(3), 196-210. https://doi.org/10.1159/000056199

Youssef, I. (2014). Emphasis spread in Cairene Arabic: A reassessment. In O. Durand, A. Langone & G. Mion (Eds.), *Alf lahga wa lahga* (Proceedings of the 9th AIDA conference) (pp. 455-464). LIT Verlag.

Zsiga, E. C. (2013). *The sounds of language: An introduction to phonetics and phonology.* Malden, Oxford: Wiley-Blackwell.

Notes:

Note 1. The following abbreviations are used in this paper: F = female, IA = Iraqi Arabic, JA = Jordanian Arabic, M = male/median, ms = milliseconds, MSA = Modern Standard Arabic, NA = Najdi Arabic, SA = Saudi Arabic, QA = Qatari Arabic, VOT = Voice Onset Time.

Note 2. UPSID is a Phonological Segment Inventory Database that contains the sound inventories of 317 languages.

Note 3. Some works on Arabic phonology, such as Qafisheh (1977), consider the consonant /p/ as a phoneme in the inventory.

“Appendix”

Stimuli

Target Stop	Vocalic Context	Carrier Word	Transcription	Gloss
/b/	/a:/	باب	ba:b	‘door’
		بات	ba:t	‘slept’
	/a/	بَق	bag	‘mosquitos’
		بَرّ	bar	‘desert’
	/u:/	بوك	bu:k	‘wallet’
		بومه	bu:mah	‘owl’
	/u/	بُرْ	bur	‘wheat’
		بُنْ	bun	‘coffee’
/t/	/a:/	تاب	ta:b	‘repented’
		تاج	ta:dʒ	‘crown’
	/a/	تَمْ	tam	‘done’
		تَفْ	taf	‘spat’
	/u:/	توم	tu:m	‘twins’
		توت	tu:t	‘berry’
	/u/	تُفْ	tuf	‘spit’ (imper)
		تُسْتْ	tust	‘toast’
/d/	/a:/	داب	da:b	‘snake’
		داخ	da:x	‘dizzy’
	/a/	دَقْ	dag	‘rang’
		دَسْ	das	‘hid’
	/u:/	دوك	du:k	‘take’
		دود	du:d	‘worms’
	/u/	دُبْ	dub	‘bear’
		دُفْ	duf	‘push’ (imper)
/k/	/a:/	كاس	ka:s	‘glass’
		كان	ka:n	‘was’
	/a/	كَح	kaħ	‘coughed’
		كَع	kaʕ	‘suddenly stopped’
	/u:/	كوب	ku:b	‘cup’
		كوخ	ku:x	‘cottage’
	/u/	كُبْ	kub	‘pour’ (imper)
		كُمْ	kum	‘sleeve’
/g/	/a:/	قاع	ga:ʕ	‘bottom’
		قاس	ga:s	‘measured’
	/a/	قَزْ	gaz	‘stove’
		قَبْر	gabur	‘grave’
	/u:/	قُوْة	gu:wah	‘force’
		قُوت	gu:t	‘food’
	/u/	قُبَه	gubbah	‘dome’
		قُفا	gufa	‘back’