Language Related Research
E-ISSN: 2383-0816
https://lrr.modares.ac.ir
https://doi.org/10.52547/LRR.13.5.17
http://dorl.net/dor/20.1001.1.23223081.1401.0.0.104.6

Vol. 13, No. 5
pp. 477-510
November & December 2022

Received: 16 May 2021
Received in revised form: 24 July 2021
Accepted: 7 August 2021

Word-initial Voicing Contrast in Khuzestani Arabic Stops

Nawal Bahrani[1]

Abstract

This research generally aimed to gain a better understanding of the stop voicing contrast in Khuzestani Arabic (KhA) by focusing on VOT and vowel onset f0. The potential effects of voicing status, place of articulation, vocalic context, and gender on the word-initial VOT and f0 measurements were investigated. 15 females and 15 males were asked to repeat 31 authentic words three times out of context with no marked intonation pattern. The results showed significant differences between VOT and f0 values as a function of voicing status, among others. In addition word-initially, the statistical analysis indicated significant place-dependent and vocalic-related VOT variations in the context of VOICELESS stops, whereas for f0 the effects of gender, place of articulation, and vocalic context were significant both at the VOICED and VOICELESS levels. Word-initially, in KhA /b, d, g/ are produced with voicing lead, while their VOICELESS counterparts /p, t, k/ have long lag, and the two stops with no VOICED cognate /tˤ, q/ make use of short lag. Overall, KhA exhibits a two-way laryngeal contrast system, and similar to Swedish the two opposite ends of VOT continuum are utilized. Logistic regression models were executed separately for both genders to predict the levels of voicing status and emphaticness. The results revealed that in voicing contrast more weight was given to VOT compared to f0. Lastly, Pearson's correlation showed a significantly positive linear relationship between the two acoustic parameters in marking the voicing feature of VOICELESS plosives.

Keywords: voice onset time, fundamental frequency, voicing contrast, stop consonant, Khuzestani Arabic

---

[1] Corresponding Author: M.A. in General Linguistics Allameh Tabataba'i University; Tehran, Iran
Email: nawal.bahrani@yahoo.com , ORCID ID; https://orcid.org/0000000179123609

[DOR: 20.1001.1.23223081.1401.0.0.104.6]
[DOI: 10.52547/LRR.13.5.18]

1. Introduction

Phonemic voicing distinction between homorganic stops has been the focus of several studies aiming to determine its acoustic correlates and their language-specific patterns. One of the most important indicators of this underlying feature is voice onset time (VOT), which refers to the time interval between the stop release and the onset of quasi-periodicity reflecting laryngeal vibration (Lisker & Abramson, 1964, p. 565). Lisker and Abramson's (1964) pioneering cross-linguistic study revealed that in many languages such as English, the word-initial /b, d, g/ are predominantly articulated with no vocal fold vibration; hence, the phonological feature [±voice] does not suffice to distinguish these consonants from their VOICELESS[1] counterparts. Instead, the distinction is heavily based on VOT patterns, measured in milliseconds (ms). If phonation starts prior to stop release, VOT is negative; when voicing onset coincides with plosive release, VOT is zero; and if voicing appears with a delay after the release, VOT is positive. There are three VOT categories: 1) voicing lead (voicing starts before the release, and the consonant is voiced); 2) short lag (voicing is zero or occurs shortly after the release, and the consonant is voiceless unaspirated); 3) long lag (voicing occurs considerably after the release, and the consonant is voiceless aspirated) (Lisker & Abramson, 1964, p. 532).

World's languages employ different laryngeal distinctions. The majority of languages have a two-way voicing contrast. In this group, languages such as English (Lisker & Abramson, 1964), Persian (Bijankhan & Nourbakhsh, 2009), and German (Jessen & Ringen, 2002) are known as aspiration languages, and Fenno-Swedish (Ringen & Suomi, 2012), French (Tranel, 1998, p. 131), and Russian (Ringen & Kulikov, 2014) belong to the true voicing group. While a contrast between short lag and long lag is observed in the former group, in the latter the distinction is implemented by voicing lead and short lag. In other words, in aspiration languages, the laryngeal contrast is manifested by [spread glottis] whereas in the voicing group the feature [voice] is adequate. Interestingly, in Lisker and Abramson's (1964) study, none of the two-way voicing contrast systems made use of both extreme ends of VOT continuum, namely voicing lead and long lag. For a long time, it was believed that in no such a system both [voice] and [spread glottis] are used (e.g. Iverson & Salmons, 1995). Recent studies have shed light on

---

[1] In this article, following Khattab (2002), the capital forms VOICED and VOICELESS refer to the phonological voicing and the lower-case ones, voiced and voiceless, denote the actual phonetic realization.

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

the existence of this unusual typological category in Swedish (Helgason & Ringen, 2008), Turkish (Öğüt et al., 2006), Najdi Arabic (AL-Gamdi et al., 2019), and Qatari Arabic (Kulikov, 2020). In these languages, the VOICED stops are specified for [voice], and the VOICELESS group has the [spread glottis] feature. Further studies on different languages might reveal that this category is not as rare as it is claimed.

Thai (Lisker & Abramson, 1964), Eastern Armenian in Tehran (Amirian, 2017), and Yerevan Armenian (Seyfarth & Garellek, 2018) possess a three-way contrast by contrasting voicing lead, short lag, and long lag. Within this category, Korean is an interesting exception in that its three sets of homorganic stops, commonly described as reinforced/fortis/tense (p*[1], t*, k*), plain/lenis/lax (p, t, k), and aspirated (p^h, t^h, k^h), are word-initially voiceless with positive VOTs (Han & Weitzman, 1970; Cho et al., 2002). Yet, in intervocalic position lenis stops are voiced in a process known as intervocalic lenis stop voicing (Cho et al., 2002, p. 196). Lastly, there are languages with a four-way phonological contrast, including Hindi (e.g., /b, p, b^h, p^h/) and Marathi (Lisker & Abramson, 1964). In these languages the distinction between VOICED unaspirated and VOICED aspirated stops is rather based on the phonation type (Dutta, 2007; Shimizu, 1996, as cited in Abramson & Whalen, 2017, p. 77).

VOT can vary as a function of several factors. Some of the observations are as follows: 1) the further back the place of articulation of the stop, the longer the VOT (Peterson & Lehiste, 1960; Cho & Ladefoged, 1999); 2) the more extended the stop closure area, the higher the VOT value (Stevens et al., 1986); 3) generally stops show longer VOTs when preceded by high vowels as opposed to low ones (Klatt, 1975; Morris et al., 2008). Regarding the influence of gender, one consistent result is that women tend to have longer VOTs than men (Swartz, 1992; Robb et al., 2005). In some studies, no significant difference between both genders was reported (Morris et al., 2008). Some scholars, however, observed longer VOT durations for males (Oh, 2011).

While VOT is the most common and widely studied phonetic parameter of voicing contrast, the onset f0 of the following vowel can serve as an additional correlate (Haggard et al., 1970; Han & Weitzman, 1970; Cho et al., 2002; Wright & Shryock, 1993; Dutta, 2007). Generally, after VOICELESS stops f0 is higher than

---

[1] The fortis sounds of Korean language are presented using capital letters, an apostrophe, or asterisk (Shin et al., 2013, p. 35).

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

after VOICED ones (Haggard et al., 1970; Kirby & Ladd, 2016; House & Fairbank, 1953; Ohde, 1984). The same alternations were also noticed in Palestinian Arabic (Tamim, 2017). Based on Whalen et al. (1993) and Shultz et al. (2012), VOT is the main correlate of voicing distinction in English; while, in the speech of young Koreans (Kang & Guion, 2008; Silva, 2006; Wright, 2007; Kang, 2014) the role of f0 is more salient.

The experimental studies on voicing contrasts have produced a large body of data from different languages. Yet, the literature of Arabic dialects, provided in the next section, shows no such analysis conducted on Khuzestani Arabic (KhA), an under-documented minority language in south western Iran which has been in a long-term contact with Persian, the official language of Iran. This Arabic dialect has three[1] emphatics /ðˤ, tˤ, sˤ/. Similar to other Gelet dialects, all cases of OA (Old Arabic) /dˤ/ have merged into /ðˤ/. /q/ and /tˤ/ do not have a VOICED cognate, and /p/ is limited to the loan words only. The oral stops of KhA are presented in the table below.

Table 1
KhA Oral Stops (Leitner & Bahrani: forthcoming)

	bilabial	dental	velar	uvular
plain	p b	t d	k g	q
emphatic		tˤ		

The present study aimed to fill the descriptive gap related to the study of KhA laryngeal contrast by investigating VOT patterns and the role of f0. /q/ and /tˤ/, the two KhA plosives with no VOICED counterpart, were added for comparative and descriptive reasons. Word-initially the effects of voicing status (VOICED and VOICELESS), place of articulation (labial, plain dental, dental emphatic, velar, uvular), vocalic contexts (/iː, uː, aː, a/), and gender (male and female) on VOT and f0 measurements, along with the vowel:voice, place:voice, and gender:voice

---

[1] In Bahrani and Modarresi Ghavami (2019), [ɾˤ, lˤ] are given phonemic status. However, more investigation in this regard revealed that the presence of these two sounds is largely limited to the context of /a, aː/ and/or other emphatic/back phonemes, hence they are merely allophonic variations of /ɾ, l/.

interactions were also inspected. Lastly, the interaction between VOT and f0 in predicting the two levels of voicing status (VOICED vs. VOICELESS) and emphaticness (emphatic vs. non-emphatic) in the word-initial position and the correlation between VOT and f0 were examined. Besides attesting the features of this dialect, this study contributes to the growing body of research about stop voicing contrast typology and helps to make further cross-linguistic and cross-dialectal comparisons.

1.2. VOT in Arabic Dialects

Considering the vast number of local Arabic dialects and their sub-branches, one can say that few studies have dealt with laryngeal contrast in Arabic varieties. In this part, the results of ten Arabic dialects are reported: Lebanese Arabic (Yeni-Komshain et al., 1977), Saudi Arabic (Flege & Port, 1981), Jordian Arabic (Mitlab, 2001, as cited in Tamim, 2017, p. 6), Egyptian Arabic (Rifaat, 2003), Mosuli (Rahim & Kasim, 2009), Modern Standard Arabic (MSA) (AlDahri, 2013), Abha Arabic (Al Mawli, 2017), Palestinian Arabic (Tamim, 2017), Najdi Arabic (AL-Gamdi, et al., 2019), and Qatari Arabic (Kulikov, 2020).

In general, Arabic dialects have a two-way laryngeal contrast. In MSA the contrast is between short lag and long lag (AlDahri, 2013); while, in the other dialects these sounds are differentiated with negative and positive values. More specifically, Lebanese (Yeni-Komshian et al., 1977), Egyptian (Rifaat, 2003), and Palestinian Arabic (Tamim, 2017) contrast voicing lead and short lag, and in Qatari (Kulikov, 2020), Mosuli[1] (Rahim & Kasim, 2009), Abha (Al Mawli, 2017), and Najdi Arabic (AL-hamdi et al., 2019) prevoicing and long voice lag are utilized. For Saudi Arabic, the VOT values of VOICED stops were not reported by Flege and Port (1981). Based on their study, Saudi Arabic speakers produce the VOICELESS cognates with long lag. The emphatic /tˤ/ is articulated with a short voicing lag in Lebanese (Yeni-Komshian et al., 1977), MSA (AlDahri, 2013), Palestinian (Tamim, 2017), and Abha Arabic (Al Mawli, 2017). As the place of articulation of VOICELESS plosives in Lebanese (Yeni-Komshian et al., 1977) and Najdi Arabic (AL-Gamdi et al., 2019) moves back, VOT slightly increases. Moreover, in some data samples from Lebanese Arabic /d, dˤ/ were uttered by short lag resulting in a

---

[1] In this dialect the borrowed phoneme /p/ is produced with short lag (see Rahim & Kasim, 2009, p. 36, table 10).

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

slight overlap with the VOT distributions of /t, tˤ/ in the range of 0–30 ms.

The results of examining the effects of gender, emphaticness, vowel duration, place of articulation, and stress on VOT in Egyptian Arabic revealed significant influences by only stress and place of articulation (Rifaat, 2013). Also, the VOICELESS stops of Mosuli (Rahim & Kasim, 2009) and Najdi Arabic (AL-Gamdi et al., 2019) and both VOICED and VOICLESS stops of Palestinian (Tamim, 2017) and Jordanian Arabic (Mitleb, 2001, as cited in Tamim, 2017, p. 6) have longer VOT durations before high vowels as compared to low ones. Lastly, Abha Arabic female natives produced VOICED and VOICELESS consonants with longer VOTs, this is while /tˤ/ had longer VOTs in males' production (Al Mawli, 2017).

2. Methodology

2.1. Participants

To increase the generalizability power of the results, acoustic data were elicited from 15 female and 15 male KhA native informants aging between 20 and 43 (*M*= 27.2, *SD*= 6.4), from the middle class, and born and raised either in Abadan or Khorramshahr. Despite being fluent in Persian, the informants spoke primarily in KhA with their family, acquaintances, and the local people, and they had a full competence of KhA. Moreover, their knowledge of English was either non-existent or insufficient. They were not informed about the purpose of the experiment prior to the recording. At the time of experiment, the majority of participants, i.e. 23 individuals, were holders of high school diploma, five had academic degrees, and one participant was an undergraduate student. No subject was reported to have any history of speech, laryngeal, and hearing disorders.

2.2. Material and Recording

To pursue the research objectives, 31 (without taking into account the sporadically located fillers) authentic KhA words containing all KhA oral stops /b, p, d, t, tˤ, g, k, q/ word-initially (four disyllabic items and 27 monosyllabic words with ˈCV(ː)C structure) were selected.[1] In the majority of cases, stops formed minimal pairs with their homorganic VOICED/VOICELESS counterparts. To explore the effects of vocalic contexts on VOT and f0, the initial target consonants occurred before /iː/,

---

[1] The word lists are provided in the Appendix A.

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

/uː/, /aː/, and /a/. In this position, no example containing /p/ followed by /aː/ was found. The recordings were made in a quiet room, using Praat (Boersma & Weenik, 2020) software (version 6.1.11), with a sampling rate of 44100 Hz. The participants were tested individually and were instructed to read aloud each word three times at a normal pace and out of context with no marked intonation patterns. The words related to each place of articulation were presented in separate pages. Before reading, the participants took their time to familiarize themselves with the word lists and during the recording it was permitted to drink water, pause for around one or two minutes after reading each page, or to make corrections when mispronouncing a word.

2.3. Acoustic and Statistical Analysis

All tokens were measured manually in Praat, then the obtained VOT and onset f0 data points were written in an Excel sheet before transferring them into R (3.6.3). Among the 360 recorded tokens of /q/, 47 (13.1%) were either spirantized or sonorized; consequently, these non-plosive productions were discarded from the experiment. It is worth mentioning that, overall, four females and five males had no non-plosive /q/ utterances, only in the production of one female speaker all /q/ samples were either fricative or approximant, and the other speakers had between 1–10 non-plosive /q/ samples. In the end, a total of 2743 samples (= 31 words × 3 repetitions × 30 participants – 47) were analyzed. Waveforms were mainly utilized in measuring VOT; however, to ensure the accuracy of measurement landmarks spectrograms were often inspected. VOT was measured as the time interval between stop release and the start of quasi-periodicity indicating the start of voicing (figure 1). The mean f0, i.e., mean pitch, was calculated by identifying the first four complete glottal pulses at the vowel onset (figure 2).

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

Figure 1
Negative VOT of /d/ in /diːn/ (-94 ms), and Positive VOT of /t/ in /tiːn/ (76 ms)

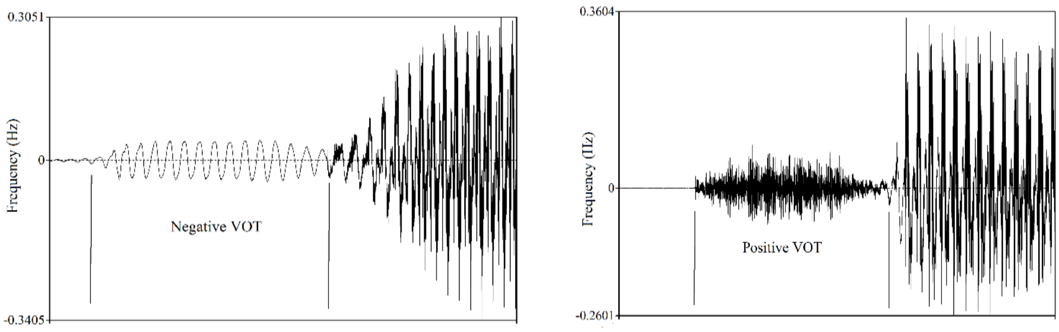
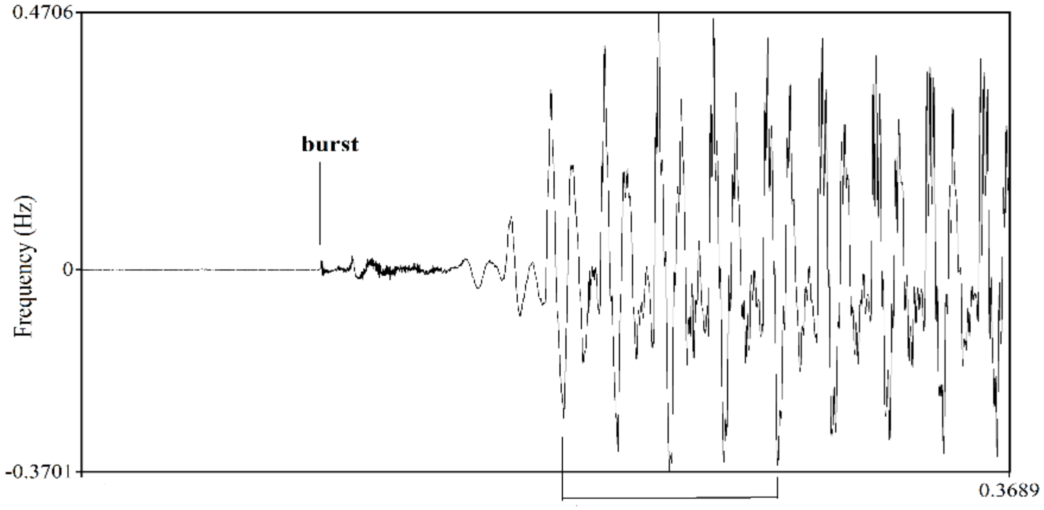


Figure 2
The First Four Complete Glottal Pulses at the Vowel Onset in /tˤiːn/



A number of advanced statistical analyses were performed on the data. The effects of voicing status, place of articulation, vocalic context, and gender on both VOT and f0 were examined using two linear regression mixed effects models run separately for each dependent variable. Since the general package used to fit the linear regression mixed effects models, i.e., lme4 (Bates et al., 2015), does not report *p*-values, the models were fitted to the data using the lmer () function in the lmerTest (Kuznestsova et al., 2017) package for R (3.6.3). This function computes the *p*-values via Satterthwaite's degrees of freedom. To test the interactions between the fixed factors in a linear mixed effects model, the data must be fully-crossed as a not fully-crossed data set results in large variance inflation factors (VIF) and multicollinearity among variables which in turn inflate the rate of type two error (i.e. the non-rejection of a false null hypothesis). Since this criterion could not be met in the interaction between voicing status and place of articulation, the three interactions, namely voice:place of articulation, voice:vocalic context, and voice:gender, could not be included in the two models mentioned above and were studied in separate models. More specifically, the interactions were split by dividing the word-initial data into VOICED and VOICELESS and running separate mixed effects models on each set. Then, for performing multiple pair-wise comparisons between the means of groups, post-hoc tests were run by the same package with the difflsmeans () function.

Also, two generalized mixed effects logistic regression models were run separately for males and females to predict the two levels of voicing status and emphaticness by word-initial VOT and f0 values with the glmer () function. Lastly, the strength and direction of the correlation and linear relationship between VOT and f0 in voicing distinction were tested by cor.test () function and the Pearson's correlation method. In all models the significance level of the test was set at 0.05. All p-values were rounded up to four decimals. The p-values smaller than 0.001 are expressed with a statement of inequality, i.e. $p< 0.001$.

3. Results

3.1. Voice Onset Time

Table 2 displays the number of tokens (N), mean, maximum (max), minimum (min), and the standard deviation (SD) of VOT (ms) values of each KhA stop in initial position. For VOICED plosives, the positive and negative values are shown separately. According to table 2, there were more tokens of aspiration for the VOICED velar stop than for the dental or labial. While both positive and negative VOTs were observed for initial VOICED stops, these consonants were predominantly produced with negative values and had a negative total mean. On the other hand, the initial VOICELESS consonants, with the exception of /q/, displayed only positive VOTs. In articulating /q/, three male informants produced four cases with the negative values of -54, -27, -37, and -31 ms. Generally, KhA homorganic stops are clearly distinguished from one another by their mean VOT values; however, there was a degree of VOT overlapping for all homorganic sets in initial position.

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

Table 2

VOT of KhA Word-Initial Stops

		mean	max	min	SD	N
/b/	negative	-81	-8	-191	31	327
	positive	12	37	0	10	33
	total	-73	37	-191	40	360
/p/	positive	55	149	11	25	270
/d/	negative	-86	-8	-245	34	325
	positive	16	37	0	7	35
	total	-76	37	-245	45	360
/t/	positive	63	128	24	21	360
/tˤ/	positive	16	46	0	7	360
/g/	negative	-79	18	-192	32	312
	positive	19	52	0	13	48
	total	-66	52	-192	45	360
/k/	positive	67	144	32	19	260
/q/	negative	-32	-27	-54	12	4
	positive	18	86	0	12	309
	total	18	86	-54	14	313

In word-initial position, the (total) maximum VOT value of each VOICED stop overlaps with the minimum VOT of its VOICELESS counterpart. This is similar to Lebanese dialect (Yeni-Komshian et al., 1977) which displayed overlaps in the ranges of the pair stops /d, t/ and /dˤ, tˤ/. According to Yeni-Komshian et al. (1977, pp. 40–41), the overlap effect was caused by all participants, and each participant showed at least one instance of overlapping. Word-initially, in KhA labial, dental, and velar places of articulation, the overlaps appeared both at the group and at the individual level with varying degrees among the speakers. More specifically, in the labial position, one female speaker, and in the velar position one male and two female participants showed overlapping values in their productions. In the production of /t, tˤ/, the overlapping was in the samples from three females and one male informant. The maximum value of /d/ and the minimum value of /tˤ/ overlapped in the speech of eight females and five males. Figure 3 shows the density plots of VOT values of KhA initial stops in labial, dental, and velar places of articulation. An inspection of table 2 and figure 3 reveals that the word-initial overlap zones are on the positive side of the VOT continuum, which suggests that it is the VOICED consonants that intrude into the ranges of their homorganic counterparts. Table 3 provides information about the range and percentage of each VOT overlap.

Figure 3
VOT Distributions of Initial Stops

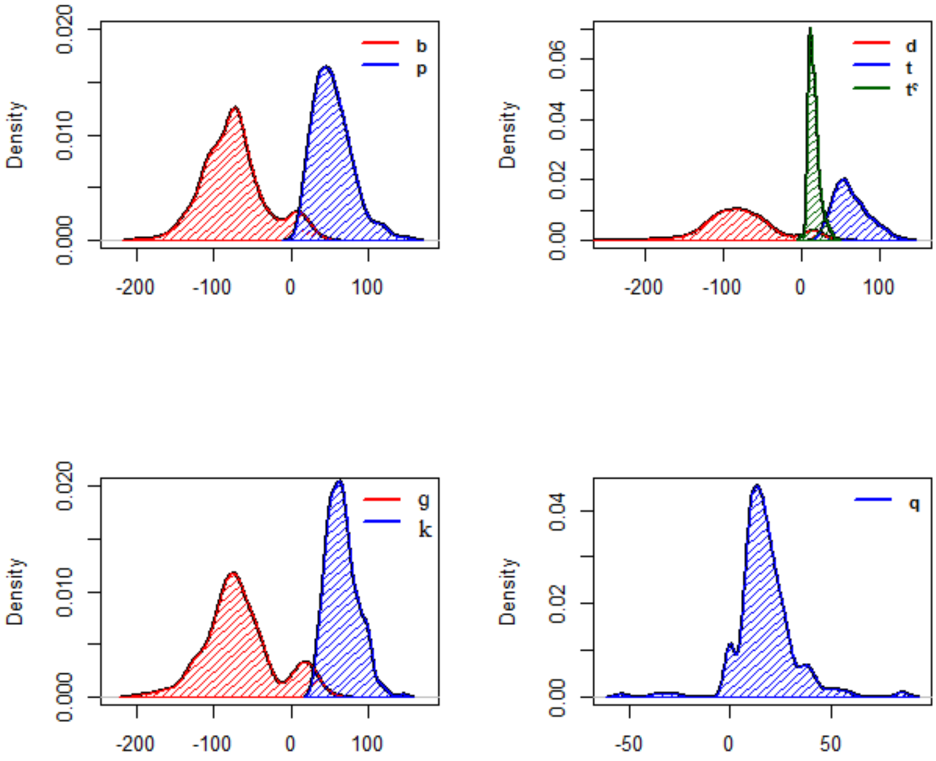


Table 3
The Range and Percentage of VOT Overlap

paired stops	range of overlap	percentage of overlap						
		b	p	d	t	tʕ	g	k
/b, p/	11–37	5.27	26.66					
/d, t/	24–37			1.11	8.05			
/t, tʕ/	24–46				23.88	13.88		
/d, tʕ/	0–37			9.27		98.88		
/g, k/	32–52						2.5	25.83

The overlapping of VOT values of homorganic stops has also been documented in a number of other languages, as a result of an internal change, e.g., Korean (Kang, 2014; Kang & Guion, 2008) or due to language contact or bilingualism, e.g., Fenno-Swedish (Ringen & Suomi, 2012) and Canadian French (CF) (Caramazza & Yeni-Komshian, 1974). The classic studies on Korean voicing contrast, conducted in the

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

1960s and 1970s, showed that speakers relied primarily on VOT patterns in voicing distinction (Hardcastle, 1973; Lisker & Abramson, 1964; Kim, 1965). In Lisker and Abramson (1964), the VOT means of Korean fortis, lenis, and aspirated were 12, 30, 104 ms, respectively. According to Hardcastle (1973, p. 265), the values of lenis stops were 3 to 5 times longer than those for fortis, and the VOTs of aspirated sounds were 2 to 3.5 times longer than lenis obstruents. However, recent studies on younger adults revealed that the VOT difference between lenis and aspirated has decreased significantly (Kang & Guion, 2006; Silva, 2006; Wright, 2007). According to Kang (2014) and Kang and Guion, (2008), the role of VOT in Korean voicing contrast is being replaced by tonal contrasts as younger natives tend to make use of f0 in signaling a distinction between lenis and aspirated. F0 values are relatively higher at the onset of the vowels preceded by aspirated and fortis obstruents than those at the onset of vowels following lenis obstruents (Cho et al., 2002).

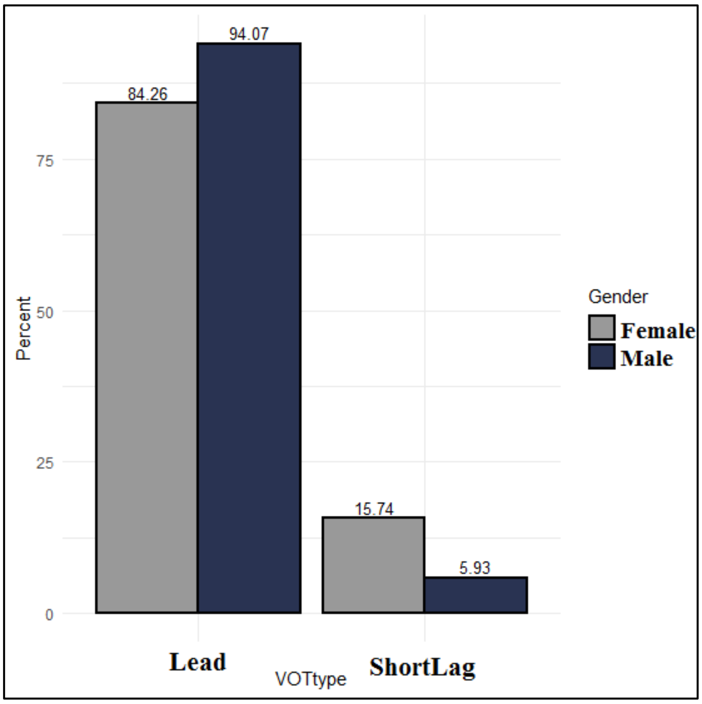
Fenno-Swedish (FS) is a Germanic language and a dialect of Swedish with a stop contrast between short lag and prevoicing. In studying the voicing contrast of FS and the potential effect of language contact and/or bilingualism on the laryngeal contrast patterns of this language, Ringen and Suomi (2012) recorded 12 FS informants who were also fluent in Finnish, a language with short lag and long lag contrast. In their experiment, some occurrences of /b, d, g/ had short-lag and overlapped with the range of /p, t, k/. This observation was introduced as an influence of Finnish. Canadian French is another language with a voicing contrast between prevoiced and short lag. This language is in contact with Canadian English (CE). Caramazza et al. (1973) reported that the CF subjects produced overlapping VOT distributions at each place of articulation as in some samples VOICED stops had short voicing lag. Caramazza and Yeni-Komshian (1974) attributed this to the extensive contact with CE. In KhA, 78% of the initial VOICED tokens were produced with prevoicing as opposed to 98% in Najdi Arabic (AL-Gamdi et al., 2019). While it cannot be confirmed yet, the relatively higher occurrence of short lag in the production of initial VOICED stops in KhA compared to Najdi Arabic might be the result of speakers growing up in a bilingual setting and acquiring Persian in which word-initially the contrast is between short lag and long lag (Bijankhan & Nourbakhsh, 2009). On the other hand, this can be only a language-internal feature. More studies, in particular, comparing KhA VOT with other Gelet varieties could reveal more information about this observation.

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

Figure 4

Percentage of Voicing Lead and Short Lag in the Articulation of Initial VOICED Stops According to Gender Differences



Based on figure 4, the percentage of short lag in the production of initial VOICED plosives was higher for females than males. Also, the greater occurrence of voicing lead in males' production compared to females' can be due to biological differences between their vocal tracts. The smaller volume of vocal tract in females leads to the supraglottal pressure adding up faster than in the case of males. This rapid change of pressure makes it hard to produce voicing lead (van Alphen & Smits, 2004, p. 459).

3.1.2. Factors Affecting VOT in Word-initial Position

To investigate the main effects of the four fixed factors, i.e. voicing status, vocalic context, place of articulation, and gender on VOT, a linear mixed effects model was built. In this model, by-speaker and by-word adjustment to intercept were treated as random effects. It should be noted that as each word was produced by both males and females, gender was included as a random slope for word. Also, voicing status and place of articulation were added as random slopes for participant. All the

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

predicators were categorical variables, as a result they were dummy-coded such that female, VOICED, labial, and /iː/ were the reference categories for gender, voicing status, place of articulation, and vocalic context respectively, and the intercept represented the grand mean. The results of Analysis of Variance (ANOVA) obtained from fitting the mixed effects model of VOT are summarized in table 4. Based on the results, there were statistically significant main effects of voicing status, *F*(1, 46.896)= 558.459, *p*<0.001, and place of articulation, *F*(4, 29.545)= 21.8656, *p*< 0.001, on VOT. The significant *p*-value of voicing status and the noticeable mean differences between the VOICED (*M*= -71.42, *SD*= 43.499) and VOICELESS (*M*=43.89, *SD*= 28.804) put emphasis on the effectiveness of VOT in differentiating between these two groups.

Table 4

The ANOVA Output of the Linear Mixed Effects Model for VOT

	Sum Sq	Mean Sq	Num DF	Den DF	F value	Pr (>F)	
voicing status	347377	347377	1	46.896	558.459	<0.001	***
vocalic context	2962	987	3	22.482	1.5873	0.2203	
gender	1838	1838	1	45.492	2.9544	0.0924	.
place of articulation	54404	13601	4	29.545	21.8656	<0.001	***

Note. Significance codes: "***"0.001; "**"0.01; "*"0.05, "."0.1.

To see how vocalic context, place of articulation, and gender interact with voicing status, two similar mixed effects models were performed separately on the data related to the VOICED and VOICELESS levels. The outputs revealed that none of the three fixed factors, i.e. vocalic context, *F*(3, 6.3295) = 0.465, *p* = 0.7167, place of articulation, *F*(2, 9.4566) = 0.647, *p*= 0.545, and gender, *F*(1, 28.6329)= 0.9172, *p* = 0.3462, had significant effects at the VOICED level, while at the VOICELESS level, only vocalic context, *F*(3, 13.468)= 4.8696, *p*= 0.0167, and place of articulation, *F*(4, 17.102) = 27.327, *p*< 0.001, showed significant effects.

To confirm where the differences occurred between groups in relation to the vocalic context and place of articulation effects, multiple comparisons with post hoc tests[1] were carried out. Based on the results for vocalic context, the only significant mean differences were between /iː/ (*M*= 51.65, *SD*= 33.23) and /aː/ (*M*= 36.83, *SD*=

---

[1] All post hoc test results are presented in the Appendix B.

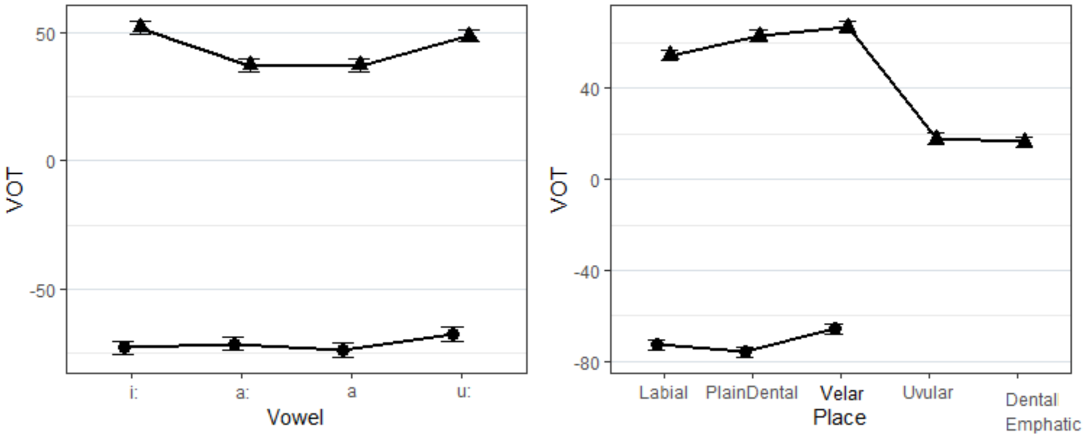
[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

22.5), *t*(13.5)= 2.9077, *p*= 0.0118, /i:/ (*M*= 51.65, *SD*= 33.23) and /a/ (*M*= 37.1, *SD*= 22.77), *t*(13.5)= 2.9519, *p*= 0.0109, /u:/ (*M*= 48.58, *SD*= 31.03) and /a:/ (*M*= 36.83, *SD*= 22.5), *t*(13.5)= 2.4194, *p*= 0.0304, and /u:/ (*M*= 48.58, *SD*= 31.03) and /a/ (*M*= 37.1, *SD*= 22.77), *t*(13.4)= 2.4231, *p*= 0.0302. Taken together, the significant and non-significant paired comparisons revealed that in the context of VOICLESS stops, vowel height had significant effects on the VOT measurements, while no significant VOT differences were generated due to vowel backness and vowel duration. It is clear from the left graph of figure 5 that the VOTs of VOICELESS stops before high vowels /i:, u:/ are longer than those in the context of the low ones /a:, a/. This confirms the previous findings regarding the systematic effect of vowel height on VOT values (e.g., Klatt, 1975; Morris et al., 2008; Esposito, 2002; Kaur, 2015). During the production of high vowels, the oral cavity is more obstructed than low vowels. This results in a more delay in air pressure drop that is essential for the occurrence of an adequate transglottal air pressure for the start of voicing, thereby the stops preceding high vowels will have longer voice lag (Kaur, 2015, p. 176).

Figure 5

The Effects of Vocalic Context and Place of Articulation on VOT at both VOICED (the Line at the bottom) and VOICELESS (the Line at the Top) Level



The multiple comparisons between the levels of place of articulation indicated that the VOT means did not differentiate between /p/ (*M* = 55, *SD*= 25) and /t/ (*M* = 63, *SD* = 21), *t*(16.6) = -1.9429, *p* = 0.0692, /t/ (*M* = 63, *SD* = 21) and /k/ (*M* = 67, *SD* = 19.4524), *t*(16.1) = -0.2763, *p* = 0.7859, /tˤ/ (*M* = 16, *SD* = 7, and /q/ (*M* = 18,

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

SD = 14), *t*(15.5) = 0.5379, *p* = 0.5983, while the other pairs were proved statistically significant. Yet, the visual examination provided in the right graph of figure 5 implies that, although insignificant, in KhA there was a systematic VOT increase from labial to dental and from dental to velar in line with the previous studies (e.g., Peterson & Lehiste, 1960). Cho and Ladefoged (1999, p. 209) provided multiple reasons for the occurrence of this phenomenon. According to their aerodynamic explanation, the small volume of supraglottal cavity behind a back constriction result in the air pressure behind the closure to be far greater than that behind a labial or alveolar constriction. Thus, it will take a longer time for this high pressure of air to drop at the release phase and allow the formation of a transglottal air pressure difference essential for initiating vocal fold vibration. Furthermore, compared to the labial and alveolar constrictions there is a larger body of air in front of a back constriction, and this mass of air must be moved out of the vocal tract in order for the compressed air behind the closure to be released. Obviously, a greater delay will occur in the start of voicing in the case of back stops.

In the uvular place of articulation, the mean VOT value of /q/ was remarkably lower. Similarly, in Iraqi Arabic (Al-Ani, 1970), /q/ was observed to have VOTs shorter than those for /k/. Ladefoged and Maddieson (1996, p. 36) found no significant difference with respect to the VOT mean of the VOICELESS velar and uvular in K’ekchi language. Based on these two studies, Ham (2001, p. 134) claimed that ‘although there is a marked decrease in VOT from velar to uvular stops in Arabic, this effect is not inextricably linked to some physiological aspect of uvular stop production’. On the other hand, according to Docherty (1992, as cited in Bijankhan & Nourbakhsh, 2009), the lower VOT mean of /q/ can be the result of languages choosing the simplest articulatory gestures when there is no voicing contrast in a certain place of articulation. This is while, Cho and Ladefoged (1999, pp. 221–222) observed little consistent differences between VOT means of velar and uvular VOICELESS stops in six languages. They suggested that the smaller contact area in the uvular region compared to the velar might be the reason for this group of stops showing lower VOT values in some languages. In addition, the emphatic consonant /tˤ/ in comparison with /t/ had a much lower VOT mean. This indicates that the secondary articulation of pharyngealization has a profound reducing impact on VOT.

3.2. F0 and Factors Affecting its Value in Word-initial Position

The study of initial VOT values in the previous section indicated the existence of overlapping regions. It is believed that overlapping zones in the range of the values of an acoustic cue marking a phonological contrast might lead to confusion in the perception if that cue is the only acoustic feature to mark the contrast (Choi, 2002, p. 3). Hence, the examination of other potential correlates to voicing contrast in KhA is vital. To test the role of word-initial f0 and the effect of a number of factors on its values, a linear regression mixed effects model was executed with voicing status, place of articulation, vocalic context, and gender as predicators, f0 as the response variable, and participant and word as random effects. Again, gender was included as a random slope for word and voicing status and place of articulation as random slopes for participant. All the predicators where dummy-coded. The output of the ANOVA is presented in the following table.

Table 5
The ANOVA Output of the Linear Mixed Effects Model for F0

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr (>F)	
voicing status	13248	13248	1	30.741	56.2156	< 0.001	***
vocalic context	15788	5263	3	21.746	22.3314	< 0.001	***
gender	39914	39914	1	34.352	169.372	< 0.001	***
place of articulation	8043	2011	4	37.281	8.5326	< 0.001	***

Note. Significance codes: "***"0.001; "**"0.01; "*"0.05, "."0.1.

Based on the table, all fixed factors had highly significant effects on f0. It was revealed that f0 is another denominator of laryngeal contrast as it was statistically influenced by the voicing category of the obstruents. The mean f0 in the context of VOICELESS stops (*M* =192.27, *SD* = 45.66) was greater than that in the context of VOICED series (*M* = 183.27, *SD* = 41.80). This model was followed by two mixed effects models to evaluate the effects of vocalic context, gender, and place of articulation at the VOICED and VOICELESS levels. The outputs indicated that in the context of VOICED stops vocalic context, *F*(3, 25.508) = 9.3723, *p*< 0.001, place of articulation, *F*(2, 22.26) = 11.9119, *p*< 0.001, and gender (females (*M* = 219.23, *SD* = 22.07), males (*M* = 147.31, *SD* = 20.49)), *F*(1, 30.66) = 154.078, *p*< 0.001) had significant effects on f0. The same patterns were observed in the context of VOICLESS plosives: vocalic context, *F*(3, 10.831) = 40.675, *p*< 0.001, place of

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

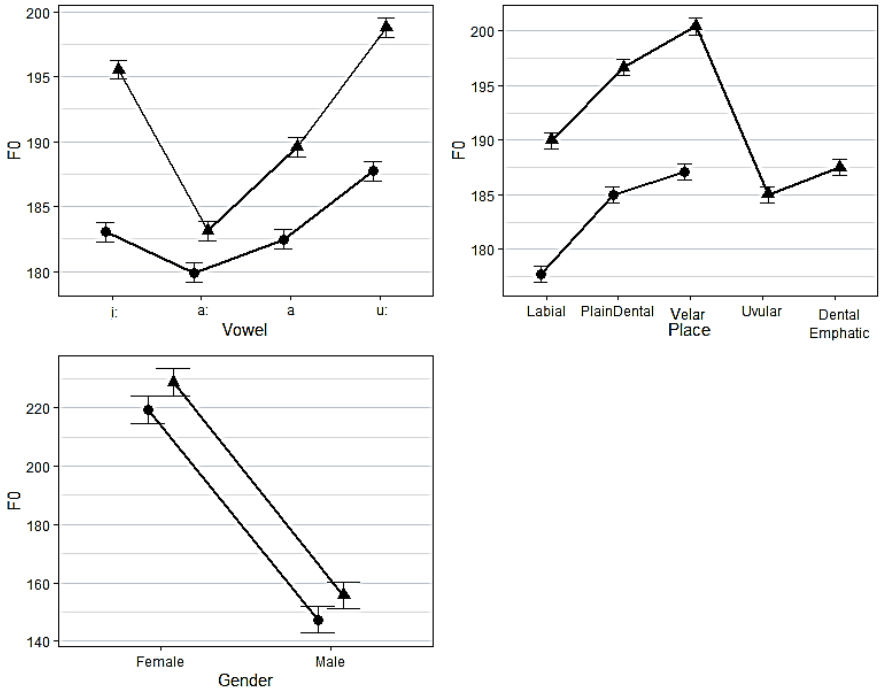
[DOI: 10.52547/LRR.13.5.18]

articulation, $F(4, 19.905) = 10.678$, $p < 0.001$, and gender (females ($M = 229.22$, $SD = 28.15$), males ($M = 155.63$, $SD = 25.87$), $F(1, 29.632) = 121.506$, $p < 0.001$.

According to the post hoc analysis of the VOICED level, only the differences between /iː/ ($M = 183.033$, $SD = 40.44$) and /a/ ($M = 182.44$, $SD = 42.06$), $t(25.5) = 0.8738$, $p = 0.3904$, /aː/ ($M = 182.5$, $SD = 44.47$) and /a/ ($M = 182.44$, $SD = 42.06$), $t(25.5) = -1.9203$, $p = 0.0661$, and /d/ ($M = 185.011$, $SD = 43.01$) and /g/ ($M = 187.114$, $SD = 41.54$), $t(20.2) = -1.749$, $p = 0.0955$, did not reach a significant level. On the other hand, in the context of VOICELESS stops, the mean difference between the two high vowels /iː/ ($M = 195.894$, $SD = 45.87$) and /uː/ ($M = 198.955$, $SD = 46.03$), $t(10.8) = -2.1695$, $p = 0.0532$, /p/ ($M = 177.69$, $SD = 40.35$) and /tˤ/ ($M = 187.52$, $SD = 43.86$), $t(25.3) = -0.2955$, $p = 0.7700$, /p/ ($M = 189.97$, $SD = 44.76$) and /q/ ($M = 200.41$, $SD = 45.96$), $t(28.9) = 0.5021$, $p = 0.6194$, /t/ ($M = 196.68$, $SD = 45.85$) and /k/ ($M = 187.11$, $SD = 41.54$), $t(20.4) = -2.019$, $p = 0.0568$, and /tˤ/ ($M = 187.52$, $SD = 43.86$) and /q/ ($M = 200.41$, $SD = 45.96$), $t(23.9) = 0.994$, $p = 0.3302$, were insignificant.

Figure 6

The Effects of Vocalic Context, Place of Articulation, and Gender on F0 at both VOICED (Line at the Bottom) and VOICELESS (Line at the Top) Level



[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

It has been widely attested that, all things being equal, intrinsically high vowels have greater f0s than of the low vowels (Whalen & Levitt, 1995). One of the hypotheses describing the intrinsic f0 variations due to vowel height is known as the tongue-pull hypothesis. Advocated by Ladefoged (1964a, 1964b) and Lehiste (1970), among others, it suggests that the raising of tongue during the production of high vowels pulls the larynx and consequently increases the tension of vocal folds which is directly linked to high f0 values (Ohala & Eukel, 1987). The effectiveness of this articulatory explanation was proved in Ohala and Eukel's (1987) study. The top left plot in figure 6 and the post hoc results demonstrate the same tendency for KhA /iː, uː/ to have higher onset f0 values than /aː, a/ after VOICELESS stops. Following VOICED plosives, however, the mean value of /iː/ was negligibly and insignificantly different from /a/. In the same context the f0 mean of /uː/ was statistically higher than both low vowels /a, aː/. Generally, the f0 differences between high and low vowels were noticeably greater when preceded by VOICELESS stops as compared to the context of initial VOICED plosives. In addition, there were significant differences between /uː/ and all three front vowels at the VOICED level. The f0 difference as a function of vowel backness was also documented in Belgian Dutch (Verhoeven & Van Hoof, 2007). With regard to interactions, according to figure 6, the pair of adjacent levels of vowels and places of articulation almost had the same direction in the context of VOICED and VOICELESS stops, as a result the interaction of these two independent factors with voicing status seems to be rather marginal or more possibly insignificant. Furthermore, for gender these two lines were completely parallel; due to this, no significant interaction between gender and voicing status is possible.

3.3. The Interaction between VOT and F0

This part deals with VOT and f0 interaction. Generalized mixed effects logistic regression models were performed to examine the cue weighting of these two correlates in the distinction between homorganic VOICED-VOICELESS stops, as well as the plain dental-emphatic dental contrast. The models were carried out separately for males and females. The acoustic variables were normalized using *z*-score transformation with the scale function so that the effect size of the two variables could be directly compared. The by-speaker (participant) and word random intercepts were added to the models. In the first two models, the binary response variable was the voicing status and the fixed effects predictors were VOT and f0. It should be mentioned that the data related to /tˤ/ and /q/, which lack a VOICED counterpart, were not included. The results of fitting the corresponding models for the voicing status as the binary variable are summarized in table 6. The obtained *p*-values showed that in the production of male participants, both VOT ($p< 0.001$) and f0 ($p = 0.0359$) are used in voicing distinction; nevertheless, the role of VOT was significantly more noticeable as it showed a higher absolute value of estimate. For females, the two predicators had similar significance levels ($p<0.001$), but the absolute value of estimate of VOT was greater than that of f0. Overall, it can be concluded that in the production of KhA male and female speakers VOT is the more prominent marker of voicing contrast. The logistic regression models of emphaticness (see table 7) suggested male-female differences with regard to the weight given to VOT and f0. VOT was the only significant correlate for males ($p = 0.0031$), while for females, none of the two predictors could predict the levels of emphaticness.

Table 6

Generalized Logistic Mixed Effects Models of VOICED-VOICELESS Contrast: VOICED is the Reference Category

Male	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-57.2	15.8	-3.629	< 0.001	***
VOT	156	47	3.314	< 0.001	***
F0	15	7.15	2.098	0.0359	*
Female	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-136	0.00279	-48860	< 0.001	***
VOT	242	0.00465	52033	< 0.001	***
F0	-13.4	0.00541	-2466	< 0.001	***

Note. Significance codes: "***"0.001; "**"0.01; "*"0.05, "."0.1.

Table 7

Generalized Logistic Mixed Effects Models of Emphaticness: Emphatic is the Reference Category

Male	**Estimate**	**Std. Error**	**z value**	**Pr(>|z|)**	
(Intercept)	37.35	17.48	2.136	0.0327	*
VOT	76.25	25.8	2.955	0.0031	**
F0	22.56	12.37	1.823	0.0683	.

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

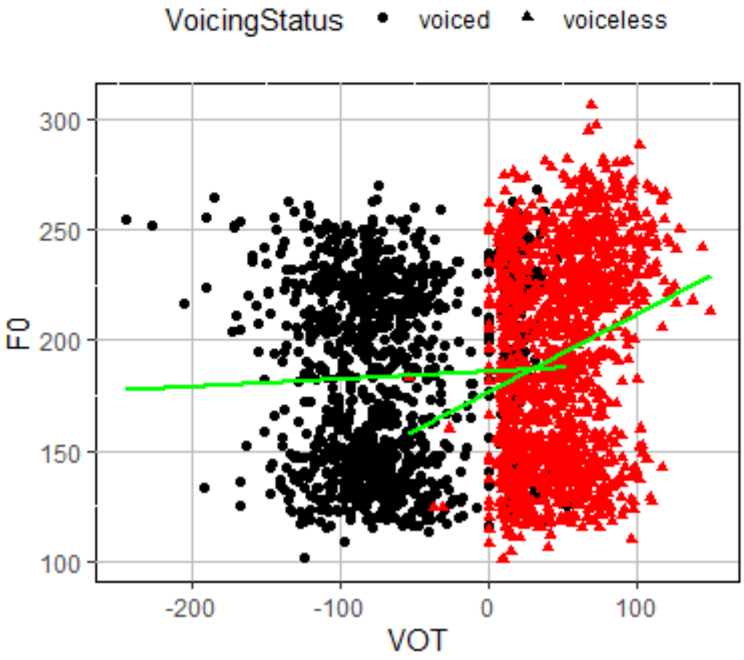
[DOI: 10.52547/LRR.13.5.18]

Male	Estimate	Std. Error	z value	Pr(>|z|)	
Female	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-7.11	9.038	-0.787	0.4314	
VOT	32.282	18.26	1.768	0.0771	.
F0	-1.665	9.724	-0.171	0.864	

Note. Significance codes: "***"0.001; "**"0.01; "*"0.05, "."0.1.

Figure 7

VOT and F0 Correlation



As a visualization of VOT and f0 interaction, a scatterplot showing separate slopes for VOICED and VOICELESS, was made (figure 7). As it can be seen, in producing VOICED plosives VOT and f0 were not considerably correlated, as the f0 values only slightly increased with the VOT changes towards the positive side of the x-axis. However, the VOICELESS category had a sharp positive slope indicating that in this category the two variables were highly correlated. In order to assess the relationship between the two numeric variables statistically, two Pearson's product-moment correlations were computed with the cor.test () function. The results demonstrated that the increases in VOT were positively and

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

significantly correlated with increases in f0 in the production of VOICELESS stops, $r(1661) = 0.2241564$, $p < 0.001$. However, in the case of VOICED plosives the obtained positive correlation was found insignificant, $r(1078) = 0.03547099$, $p = 0.2441$.

4. Conclusion

The current study mainly was aimed at assessing the phonetic make-up of laryngeal contrast between KhA stops. For this purpose, the role of VOT and the onset f0 of the following vowels were examined. In addition, the variations in the magnitude of VOT and f0 as a function of voicing status, vocalic context, place of articulation, and gender were investigated. The findings revealed that in this Arabic dialect, word-initially the plain VOICELESS sounds with a VOICED cognate /p, t, k/ are articulated with positive VOTs of around 50 ms; and the stops with no VOICED counterpart /tˤ, q/ have positive VOT values of roughly 20 ms. With this in mind we can argue that the former group possesses long voice lag, and the latter is characterized by short voice lag. Similar to a number of Arabic dialects (e.g., Lebanese (Yeni–Komshian et al., 1977), Egyptian (Rifaat, 2003), Palestinian (Tamim, 2017), and Qatari (Kulikov, 2020)), the VOICED series /b, d, g/ predominantly display voicing lead. Moreover, although the peak of VOT distributions of homorganic pairs were considerably apart along the VOT continuum, in initial position there was a degree of overlapping between homorganic sounds both at the group and at the individual level as a number of VOICED samples were made by short lag. This tendency can be attributed to the bilingualism and/or contact with Persian, or it can be simply a language-internal feature. Generally, KhA exhibits a two-way stop series that similar to Swedish (Helgason & Ringen, 2008), Najdi (AL-Gamdi et al., 2019), and Qatari Arabic (Kulikov, 2020) contrasts prevoicing with aspiration.

The linear regression mixed effects model demonstrated that voicing status and place of articulation had significant effects on VOT. However, only at the VOICELESS level, the effects of place of articulation and vocalic context were proved significant. In this context, the post hoc results indicated meaningful mean differences between high and low vowels. VOICELESS stops before high vowels had greater VOT means compared to those in the context of low ones. Also, there was no significant difference due to vowel duration or vowel backness. The pair-wise comparisons of places of articulation revealed no significant differences

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

between the means of /p/ and /t/, and /t/ and /k/. Though insignificant, the visual representation (figure 5) showed that in accordance with the related previous findings in Lebanese (Yeni-Komshain et al., 1977) and Najdi Arabic (AL-Gamdi et al., 2019) there was a systematic increase of VOT from labial to dental and from dental to velar. In the case of the uvular and dental emphatic plosives, the VOT mean was noticeably lower compared to the other VOICELESS consonants. The output of the linear mixed effects model of f0 indicated that the dependent variable was significantly influenced by voicing status, vocalic context, place of articulation, and gender. In addition, at the VOICED and VOICELESS level, the three studied fixed variables, i.e. vocalic context, place of articulation, and gender, had meaningful impacts on f0 values. Post hoc comparisons showed that after VOICELESS stops, there were meaningful differences due to vowel height. With regard to the effect of vowel height at the VOICED level, all differences, except /iː/ and /a/ were reported significant; however, the value of the high back vowel was statistically greater than the front ones. Based on the mixed effects logistic regression model, both VOT and f0 serve as voicing contrast denominators; nevertheless, the results displayed that VOT is the primary and f0 is the secondary phonetic correlate for both genders. The same model was run for emphaticness. The outputs indicated that while neither could predict the two levels of the response variable in females' samples, VOT was an effective predicator in males' production. Lastly, there was a positive linear relationship between the two phonetic descriptors. Yet this correlation was statistically meaningful only in the production of VOICELESS implosives.

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

References

Abramson, A. S., & Whalen, D. H. (2017). Voice onset time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, *63*, 75–86. https://doi.org/10.1016/j.wocn.2017.05.002.

Al Dahri, S. S. (2013). A study for the effect of the emphaticness and language and dialect for voice onset time (VOT) in Modern Standard Arabic (MSA). *Signal & Image Processing*, *4*(2), 11–29. http://arxiv.org/abs/1305.2680.

Al Malwi, I. M. (2017). *The effects of gender and age on voice onset time by Abha Arabic speakers* (publication no.10252690) [Master’s thesis, California State University, Long Beach]. ProQuest Information and Learning Company.

Al–Ani, S. H. (1970). *Arabic phonology: An acoustical and physiological investigation*. De Gruyter Mouton. https://doi.org/10.1515/9783110878769.

AL-Gamdi, N., Al-Tamimi, J., & Khattab, G. (2019). The acoustic properties of laryngeal contrast in Najdi Arabic initial stops. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (Eds.), *Proceedings of the 19th international congress of phonetic sciences* (pp. 2051–2055). Australasian Speech Science and Technology Association Inc.

Amirian, B. (2017). *Vizhegihā-ye sotshenakhti-ye hamkhānhā-ye ensedādi va ensāyeshi dar zabān-e Armani-ye Sharghi* [A study of acoustic features of stops and affricates in Eastern Armenian] [Unpublished Master’s thesis]. Allameh Tabataba’i University.

Bahrani, N., & Modarresi Ghavami, G. (2019). Khuzestani Arabic. *Journal of International Phonetic Association*. https://doi.org/10.1017/S0025100319000203.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. http://dx.doi.org/10.18637/jss.v067.i01.

Bijankhan, M., & Nourbakhsh, M. (2009). Voice onset time in Persian initial and intervocalic stop production. *Journal of International Phonetic Association*, *39*(3), 335–364. https://doi.org/10.1017/S0025100309990168.

Boersma, P., & Weenik, D. (2020). *Praat: Doing phonetics by computer* (version 6.1.11) [Computer software]. http://www.praat.org/.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of Acoustical Society of America*. *54*(2), 421–428. https://doi.org/10.1121/1.1913594.

Caramazza, A., & Yeni-Komshian, G. H. (1974). Voice onset time in two French dialects. *Journal of Phonetics*, *2*(3), 239–245. https://doi.org/10.1016/S0095-4470(19)31274-4.

Cho, T., Jun, S., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics*, *30*(2), 193–228. https://doi.org/10.1006/jpho.2001.0153.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, *27*(2), 207–229. https://doi.org/10.1006/jpho.1999.0094.

Choi, H. (2002). Acoustic cues for the Korean stop contrast- dialect variation. In T. A. Hall, B. Pompino, & M. Rochoń (Eds.), *Zap papers in linguistics* (*28*, pp. 1–12). Zentrum für Allgemeine sprachwissenchaft.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. De Gruyter Mouton. https://doi.org/10.1515/9783110872637.

Dutta, I. (2007). *Four-way stop constraints in Hindi: An acoustic study of voicing, fundamental frequency and spectral tilt* [Doctoral dissertation, University of Illinois at Urbana-Champaign]. Graduate Dissertations and Theses at Illinois. http://hdl.handle.net/2142/82651.

Esposito, A. (2002). On vowel height and consonantal voicing effects: data from Italian. *Phonetica*, *59*(4), 197–231. https://doi.org/10.1159/000068347.

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

Haggard, M. A. S., & Callow, M. (1970). Pitch as a voicing cue. *Journal of the Acoustical Society of America*, *47*(2), 613–617. https://doi.org/10.1121/1.1911936.

Ham, H. W. (2001). *Phonetics and phonological aspects of geminate timing*. Routledge. https://doi.org/10.4324/9781315023755.

Han, M. S., & Weitzman, R. S. (1970). Acoustic features of Korean /P, T, K/, /p, t, k/, and /ph, th. Kh/. *Phonetica*, *22*(2), 112–128. https://doi:10.1159/000259311.

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

Hardcastle, W. J. (1973). Some observations of the tense-lax distinction in initial stops in Korean. *Journal of Phonetics*, *1*(3), 263–271. https://doi.org/10.1016/s0095-4470(19)31390-7.

Helgason, P., & Ringen, C. (2008). Voicing and aspiration in Swedish stops. *Journal of Phonetics*, *36*(4), 607–628. https://doi.org/10.1016/j.wocn.2008.02.003.

House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America*, *25*(1), 105–113. https://doi.org/10.1121/1.1906982.

Iverson, G. K., & Salmons, J. C. (1995). Aspiration and laryngeal representation in Germanic. *Phonology*, *12*(3), 369–396. https://doi:10.1017/S0952675700002566.

Jessen, M., & Ringen, C. (2002). Laryngeal feature in German. *Phonology*, *19*(2), 189–218. https://doi:10.1017/S0952675702004311.

Kang, K., & Guion, S. G. (2006). Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. *Journal of Acoustical Society of America*. *119*, 1672–1683. https://doi.org/10.1121/ 1.2166607.

Kang, Y. (2014). Voice onset time merger and development of tonal contrast in Seoul Korean stops: a corpus study. *Journal of Phonetics*, *45*, 76–90. https://doi.org/10.1016/j.wocn.2014.03.005.

Kaur, J. (2015). Factors influencing voice onset time (VOT): Voice recognition. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, *3*(10), 174–179.

Khattab, G. (2002). VOT production in English and Arabic bilingual and monolingual children. In D. B. Parkinson & E. Benmamoun (Eds.), *Perspectives on Arabic linguistics: Papers from the annual symposium on Arabic linguistics*. (*13–14*, pp. 1–37). John Benjamins. https://doi.org/10.1075/cilt.230.03kha.

Kim, Ch. (1965). On the autonomy of the tensity in stop classification (with special reference to Korean stops. *Word*, *21*, 339–359. https://doi.org/10.1080/ 00437956.1965.11435434.

Kirby, J., & Ladd, R. (2016). Effects of obstruent voicing on vowel f0: Evidence from “true voicing” languages. *Journal of the Acoustical Society of America*, *140*(4), 2400–2411. https://doi.org/10.1121/1.4962445.

Klatt, D. H. (1975). Voice Onset Time, frication, and aspiration in word-initial

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

consonant clusters. *Journal of Speech and Hearing Research*, *18*(4), 686–706. https://doi.org/10.1044/jshr.1804.686.

Kulikov, V. (2020). Laryngeal contrast in Qatari Arabic: effect of speaking rate on VOT. *Phonetica*, *77*(3), 163–185. https://doi.org/10.1159/000497277.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. John Wiley and Sons Ltd.

Ladefoged, P. (1946a). Some possibilities in speech synthesis. *Language and Speech*, *7*(4), 205–214. https://doi.org/10.1177/002383096400700401.

Ladefoged, P. (1946b). *A phonetic study of West African languages: An auditory-instrumental survey*. Cambridge University press. https://doi:10.1017/S0022226700013001.

Lehiste, I. (1970). *Suprasegmentals*. MIT press.

Leitner, B., & Bahrani, N. (Forthcoming). Khuzestani Arabic. In Ch. G. Häberl (Ed.), *Non-Iranian languages in contact*. De Gruyter Mouton.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 527–565. https://doi.org/10.1080/00437956.1964.11659830.

Mitleb, F. M. (2001). Voice onset time of Jordanian Arabic stops. *Journal of the Acoustical Society of America*, *109*(5), 2427. https://doi.org/10.1121/1.4744787.

Morris, R. J., McCrea, Ch. 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.

Öğüt, F., Akif, K. M., Engin, E. Z., & Midilli, R. (2006). Voice onset time for Turkish stop consonants. *Speech Communication*, *48*(9), 1094–1099. https://doi.org/10.1016/j.specom.2006.02.003.

Oh, E. (2011). Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics*, *39*(1), 59–67. https://doi.org/10.1121/1.4783775.

Ohala, J. J., & Eukel, B. W. (1987). Explaining the intrinsic pitch of vowels. In R.

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

Chanon & L. Shockey (Eds.), *In honor of Ilse Lehiste* (pp. 207–215). De Gruyter Mouton. https://doi.org/10.1515/9783110886078.207.

Ohde, R. N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America, 75*(1), 224–230. https://doi.org/10.1121/1.390399.

Peterson, G. E. & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, *32*(6), 693–703.

Rahim, A. J., & Kasim, R. Z. (2009). A spectrographic study of voice onset time in Arabic. In *Journal of Education and Science*, *16*, 28–41.

Rifaat, Kh. (2003). Voice onset time in Egyptian Arabic: A case where phonological categories dominate. In *Proceedings of the 15th international congress of phonetic science*, 791–794. http://www.internationalphoneticassociation.org /icphs/ icphs2003.

Ringen, C., & Suomi, K. (2012). The Voicing Contrast in Fenno-Swedish Stops. *Journal of Phonetics*, *40*(3), 419–429. https://doi.org/10.1016/j.wocn.2012.02.010.

Ringen, C., & Kulikov, V. (2014). Voicing in Russian stops: cross-linguistic implications. *Journal of Slavic Linguistics*, *20*(2), 269–286. https://doi.org/10.1016/j.wocn.2012.02.010.

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

Seyfarth, S., & Garellek, M. (2018). Plosive voicing acoustics and voice quality in Yerevan Armenian. *Journal of Phonetics*, *71*, 425–450. https://doi.org/10.1016/j.wocn.2018.09.001.

Shimizu, K. (1996). *A cross-language study of voicing contrast of stop consonants in Asian languages*. Seibido Publishing co.

Shin, J., Jieun, K., & Cha, J. (2013). *The sounds of Korean*. Cambridge University press. https://doi.org/10.1017/CBO9781139342858.

Shultz, A. A., Francis, A. L., & Llanos, F. (2012). Differential cue weighting in perception and production of consonant voicing. *Journal of the Acoustical Society of America*. *132*(2), 95–101. https://doi.org/10.1121/1.4736711.

Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, *23*(2), 287–308. https://doi:10.1017/

S0952675706000911.

Stevens, K. N., Keyser, S. J., & Kawasaki, H. (1986). Toward a phonetic and phonological theory of redundant features. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 426–449). Lawrence Erlbaum Associates.

Swartz, B. L. (1992). Gender difference in voice onset time. *Perceptual and Motor Skills*, *75*(3), 983–992. https://doi.org/10.2466%2Fpms.1992.75.3.983.

Tamim, N. (2017). *Voicing contrast of stops in the Palestinian Arabic dialect* [Master's thesis, University Van Amsterdam]. University of Amsterdam Archive. http://www.fon.hum.uva.nl/archive/2017/2017-MA-NourTamim.pdf.

Tranel, B. (1998). *The sounds of French: An introduction.* Cambridge University Press. https://doi.org/10.1017/CBO9780511620645.

Van Alphen, P. M., & Smits, R. (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics 32*(4), 455–491. https://doi.org/10.1016/j.wocn.2004.05.001.

Verhoeven, J., & Van Hoof, S. (2007). Intrinsic vowel pitch in Dutch and Arabic. In *Proceedings of the 16th international congress of phonetic sciences* (pp. 1785–1788). Univ. des Saarlandes.

Whalen, D. H., Abramson, A. S., Lisker, L., & Mody, M. (1993). F0 gives voicing information even with unambiguous voice onset time. *Journal of the Acoustical Society of America*, *93*(4), 36–49. https://doi.org/10.1121/1.406678.

Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic f0 of vowels. *Phonetics*, *23*(3), 349–366. https://doi.org/10.1016/S0095-4470(95)80165-0.

Wright, J. D. (2007). *Laryngeal contrast in Seoul Korean* (Publication No. AAI3292100) [Doctoral dissertation, University of Pennsylvania]. ProQuest Information and Learning Company.

Wright, R. & Shryock, A. (1993). The effects of implosives on pitch in SiSwati. *Journal of the International Phonetic Association*, *23*(1), 16–23. https://doi:10.1017/S0025100300004734.

Yeni-Komshian, G. H., Caramazza, A., & Preston, M. S. (1977). A study of voicing in Lebanese Arabic. *Journal of Phonetics*, *5*, 35–48. https://doi.org/10.1016/S0095-4470(19)31112-X.

Appendix A

Test material (target stops are in bold)

/**b**iːd/	in hand	بید	/**d**all/	[he] guided	دل	/**g**uːl/	say (IMP. M)	گول
/**b**uːg/	steal (IMP.M)	بوگ	/**t**iːn/	fig	تین	/**g**aːl/	[he] said	گال
/**b**aːt/	(he) slept over/ sleep over (IMP.M.S)	بات	/**t**uːb/	repent (IMP.M)	توب	/**g**aʧ/	chalk	گچ
/**b**adɾ/ [ˈbadeɾ]	proper noun	بدر	/**t**aːb/	[he] repented	تاب	/**k**iːfa/ [ˈkiːfe]	bag	کیفه
/**p**iːp/	pipe	پیپ	/**t**all/	hill	تل	/**k**uːχ/	cottage	کوخ
/**p**uːk/	empty	پوک	/tˤiːn/	mud	طین	/**k**aːs/	Water glass	کاس
/**p**atu/ [ˈpatu]	blanket	پتو	/**tˤ**uːl/	length	طول	/**k**af/	foam	کف
/**d**iːn/	religion	دین	/**tˤ**aːb/ [tˤɑːb]	[he] got better	طاب	/**q**iːɾ/	tar	قیر
/**d**uːʃ/	shower	دوش	/**tˤ**all/	[he] peeped	طل	/**q**uːl/	monster	غول
/**d**aːʃ/	[he] entered	داش	/**g**iːɾa/ [ˈgiːɾe]	pin	گیره	/**q**aːli/	precious	غالی
						/**q**aʃ/	faint	غش

Appendix B

Table B1. Post hoc results of VOT mean differences of initial VOICELESS stops at different vocalic contexts

Mean difference Standard error	Vocalic contexts			
	/iː/	/uː/	/aː/	/a/
/iː/		2.1252 3.9754	12.5185 4.30528	11.7331 3.97482
/uː/	0.6016		10.3933 4.29585	9.6079 3.9652
/aː/	0.0118*	0.0304*		0.7854 4.29337
/a/	0.0109*	0.0302*	0.8576	

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

Table B2.

Post Hoc Results of VOT Mean Differences of Initial VOICELESS Stops in Relation to Different Places of Articulation

mean difference standard error	Initial VOICELESS stops				
	/p/	/t/	/tˤ/	/k/	/q/
/p/		11.677 6.00987	31.903 6.54554	13.149 5.91815	34.6109 6.67227
/t/	0.0692·		43.5797 5.62086	1.4719 5.32812	46.2876 6.09617
/tˤ/	< 0.001***	< 0.001***		45.052 5.31918	2.70789 5.03452
/k/	0.0403*	0.7859	< 0.001***		47.7596 5.84499
/q/	< 0.001***	< 0.001***	0.5983	< 0.001***	

Table B3.

Post Hoc Results of F0 Mean Differences of Initial VOICELESS Stops at Different Vocalic Contexts

Mean difference Standard error	Vocalic contexts			
	/i:/	/u:/	/a:/	/a/
/i:/		3.2513 1.49865	13.5635 1.62569	5.45755 1.49837
/u:/	0.0532·		16.8148 1.61975	8.70886 1.49237
/a:/	< 0.001***	< 0.001***		8.106 1.61811
/a/	0.004**	0.001***	< 0.001***	

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

Table B4.

Post Hoc Results of F0 Mean Differences of Initial VOICELESS Stops in Relation to Different Places of Articulation

mean difference standard error	Initial VOICELESS stops				
	/p/	/t/	/tˤ/	/k/	/q/
/p/		9.3245 2.22421	0.8436 2.8547	13.947 2.96302	1.71935 3.424
/t/	< 0.001***		8.4809 2.33165	4.623 2.28971	11.0438 3.07277
/tˤ/	0.7700	0.0015**		13.104 2.90813	2.56294 2.57851
/k/	< 0.001***	0.0568·	< 0.001***		15.6668 3.07011
/q/	< 0.001***	< 0.001***	0.5983	< 0.001***	

Table B5.

Post Hoc Results of F0 Mean Differences of Initial VOICED Stops at Different Vocalic Contexts

Mean difference Standard error	Vocalic contexts			
	/iː/	/uː/	/aː/	/a/
/iː/		3.1301 1.29138	3.60824 1.29138	1.12836 1.29138
/uː/	0.0228*		6.73833 1.29138	4.25845 1.29138
/aː/	0.0097**	< 0.001***		2.4799 1.29138
/a/	0.3904	0.0029**	0.0661·	

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

Table B6.

Post Hoc Results of F0 Mean Differences of Initial VOICED Stops in Relation to Different Places of Articulation

mean difference standard error	Initial VOICED stops		
	/b/	/d/	/g/
/b/		6.8056 1.93944	9.6162 1.97602
/d/	0.0018**		2.8106 1.60697
/g/	< 0.001***	0.09549·	

Table B7.

The Mean and SE Values of Stops and Vowels in the Post Hoc Analyses

	Category		VOT		F0	
			Mean	Standard Error	Mean	Standard Error
Voicing status* Vocalic context	VOICED	iː	-72.863	46.68911	183.033	40.43926
	VOICELESS	iː	51.6475	33.22959	195.894	45.87323
	VOICED	uː	-67.578	42.29846	187.741	43.13814
	VOICELESS	uː	48.5815	31.03336	198.955	46.02757
	VOICED	aː	-71.348	39.48183	179.87	41.36844
	VOICELESS	aː	36.8256	22.49778	182.5	44.47032
	VOICED	a	-73.889	45.15355	182.444	42.06247
	VOICELESS	a	37.1016	22.766	189.648	44.62131
Voicing status* Place of articulation	VOICED	labial	-72.694	40.31825	177.692	40.34637
	VOICELESS	labial	54.5407	24.52334	189.967	44.75537
	VOICED	plain dental	-75.886	44.59856	185.011	43.01117
	VOICELESS	plain dental	63.1972	21.17266	196.681	45.85493
	VOICELESS	dental emphatic	16.4306	6.94432	187.517	43.86044
	VOICED	velar	-65.678	44.91901	187.114	41.53544
	VOICELESS	velar	66.9833	19.45238	200.411	45.95605
	VOICELESS	uvular	17.5272	13.55142	185.304	46.23824

[DOR: 20.1001.1.23223081.1401.0.0.104.6]

[DOI: 10.52547/LRR.13.5.18]

Acknowledgments

The author is very grateful to Jalal Al-Tamimi, Henning Reetz, James Kirby, and Nicolai Pharao for their helpful comments on the statistical methods and all the speakers who participated in this study.

About the Author

Nawal Bahrani is a speaker of Khuzestani Arabic variety. She is interested in phonetics, phonology-interface, and phonology. In her research, she focuses on KhA and other understudied languages of Iran.