Laryngeal realism and the voicing contrast in Khuzestani Arabic stops

Nawal Bahrani

Faculty of Persian Literature and Foreign Languages, Allameh Tabataba'i University, Iran

nawal.bahrani@yahoo.com

Vladimir Kulikov

Department of English Literature and Linguistics, Qatar University, Doha, Qatar

vkulikov@qu.edu.qa

In LARYNGEAL REALISM (LR), laryngeal specification of stops is explained by direct maps of cues (e.g. VOT) onto privative phonological laryngeal features [voice] or [spread glottis]. Phonetic realization of the segments and speakers' 'control' (e.g. the degree of intervocalic voicing and speech rate manipulation effects on VOT duration) are used as diagnostics of phonological specification. Similar to some Arabic vernacular dialects (e.g. Qatari Arabic), Khuzestani Arabic in Iran presents a case where three voiced stops /b d g/ are in contrast with voiceless stops /p t k/, but two voiceless guttural plosives /tˤ q/ have no voiced homorganic counterparts. In this paper we examine the phonetic realization of voicing in these stops at word-initial and intervocalic position, as well as the effects of speech rate manipulation on VOT and closure voicing. The data came from 12 native speakers recorded in Khorramshahr, Iran. Our findings suggest an over-specified voicing system in this Arabic variety. We found that voiced /b d g/ were produced with voicing lead in initial position and complete closure voicing word medially, voiceless /p t k/ had long lag VOT, while guttural /tˤ q/ had short lag VOT. Speech rate manipulation revealed that only duration of (pre)voicing and duration of aspiration increased in slower speech. Also, f0, F1, and F2 were measured at vowel onset to evaluate the glottal state in production of stops. The results support the predictions of LR that voiced stops are specified by [voice], voiceless stops are [sg], while gutturals lack underlying specification for voice.

1 Introduction

In world's languages, stop consonants produced at the same place of articulation are often differentiated by a phonological contrast traditionally known as 'voicing contrast'. In this tradition, the contrasting segments are classified as either VOICED or VOICELESS. Acoustic correlates of this contrast include voice onset time (VOT), percent voicing (voicing ratio, VR), fundamental frequency (f0), frequencies of the first (F1) and second (F2) formant, constriction duration, or the length of the neighboring vowel. Among these, VOT (Lisker & Abramson 1964) is mainly considered the primary acoustic correlate of the voicing contrast in stop consonants in word-initial position. Laryngeal realism (LR) theory (Iverson & Salmons 1995, Honeybone 2005, Beckman et al. 2011, Beckman, Jessen & Ringen 2013) closely reflects VOT typology and directly maps privative phonological features [voice] or [spread glottis] on the VOT categories, such as voicing lead (prevoicing) or long voicing lag.

Journal of the International Phonetic Association (2024) 54/1
doi:10.1017/S002510032300004X

First published online 16 March 2023

The mapping between a VOT category and the corresponding phonological feature is direct and straightforward in word-initial position, but it becomes less obvious in other prosodic positions. For example, aspiration in voiceless stops is often reduced in intervocalic position (Lisker 1986) or even lost after [s] in English (Iverson & Salmons 1995). Thus, for word medial position, other acoustic correlates, such as duration and percent of closure voicing (Slis 1986, Beckman et al. 2013, Schwarz, Sonderegger & Goad 2019) or duration of the adjacent vowel (Lisker 1986), become more important to differentiate between voiced and voiceless stops in English.

To support direct mapping of the phonetic realization of the voiced and voiceless sounds onto phonological features, LR employs several diagnostics of speakers' 'control' (Schwarz et al. 2019), which include the effect of speech rate manipulation on VOT measurements. According to the literature (Solé 1992, Pind 1995, Kessinger & Blumstein 1997, Allen & Miller 1999) the duration of the phonetic correlates for the specified features increases as the rate of speech slows down. For the voicing contrast across languages, this typically results in asymmetric changes in VOT duration, when VOT increases only in a phonologically specified voiced or aspirated category, but not in an unspecified category (Beckman et al. 2011).

While studying the effects of rate manipulation on VOT in initial stops is a well-established procedure within the framework of LR, very few studies looked into voicing of word-medial stops, where acoustic correlates such as percent closure voicing or duration of a preceding vowel can be more salient than VOT (Lisker 1986). A recent study of voicing in Nepali (Schwarz et al. 2019) demonstrates that closure voicing in intervocalic stops also changes in response to rate manipulation in line with VOT patterns in initial stops. Therefore, a comprehensive study of a laryngeal contrast in a language in the framework of LR must include not only investigation of VOT categories but also the degree of voicing in intervocalic stops.

Although VOT is traditionally viewed as the main cue to VOICING across languages, recent studies suggest it can signal other phonological contrasts as well. In some vernacular Arabic dialects, e.g. Qatari Arabic (Kulikov 2020, 2022), voicing lead and long lag VOT are used in an over-specified laryngeal contrast similar to Swedish, but short lag VOT is consistently found in production of voiceless emphatic stop [tˤ]. This distribution of VOT categories makes short lag VOT a cue not to voicing but to a contrast in EMPHASIS, or pharyngealization (Kulikov 2022). To the best of our knowledge, there are no studies that investigated this case in depth.

It is of note, that some other Arabic dialects, e.g. Khuzestani Arabic (KhA; Bahrani 2022), also reveal a similar pattern. KhA is a Mesopotamian Arabic variety spoken in the southwest of Iran. The consonantal system of this language contains three voiced stops /b d g/ with voiceless counterparts /p t k/, and two guttural plosives /tˤ q/ which have no voiced counterparts (Bahrani & Modarresi Ghavami 2021, Leitner & Bahrani, forthcoming). Although in this variety the voicing contrast is between prevoicing and long lag VOT, the short lag pattern also exists as it is employed in the realization of /tˤ q/ (Bahrani 2022). Hence, this language is a convenient case to study the mapping of VOT on more than one phonological contrast in a language. In this paper, we are looking into the laryngeal feature specification in KhA based on LR and provide new evidence in support of a tight correspondence between the phonological features and their phonetic realization.

2 Background

2.1 Acoustic correlates of laryngeal contrast

As mentioned earlier, voice onset time (VOT) has the greatest acceptance as the primary phonetic attribute to voicing in utterance-initial plosives (Lisker & Abramson 1964). It is a temporal correlate referring to the time interval between the stop release and the start of

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

Table 1 Mapping of VOT on phonological categories across laryngeal contrasts in stops.

Contrast	Language	Voiced /b d g/	Voiceless /p t k/	Voiceless aspirated /pʰ tʰ kʰ/	Voiced aspirated /bʱ dʱ gʱ/
Two-way	French, Russian	voicing lead	short lag		
	English, German	short lag	long lag		
	Swedish	voicing lead	long lag		
Three-way	East Armenian, Thai	voicing lead	short lag	long lag	
Four-way	Marathi, Nepali	voicing lead	short lag	long lag	voicing lead, long lag

quasi-periodicity which indicates the presence of laryngeal vibration. If voicing onset occurs before the release, VOT is negative; when phonation happens after the release phase, VOT is positive; and if both vocal fold vibration and the stop release are simultaneous, VOT is considered zero. In utterance-initial position, languages use contrastive stops from these VOT categories: truly voiced stops with voicing lead/prevoicing; voiceless unaspirated stops with short lag VOT, and voiceless aspirated stops with long lag VOT.

Languages show variation in how the three categories of VOT correlate with phonological categories of VOICED and VOICELESS stops, as indicated in Table 1. In languages with a two-way laryngeal contrast, the contrast is often either between voicing lead and short lag, or between short lag and long lag, or between voicing lead and long lag. Languages employing the former contrast are referred to as TRUE VOICE languages, for example French (Tranel 1998: 131), Fenno-Swedish (Ringen & Suomi 2012), or Russian (Ringen & Kulikov 2012). Languages utilizing the second type are known as ASPIRATING languages, e.g. English (Lisker & Abramson 1964), German (Jessen & Ringen 2002), or Persian (Bijankhan & Nourbakhsh 2009). Finally, some languages utilize the two opposite ends of the VOT continuum, i.e., voicing lead and long voicing lag. Until recently, the existence of such a contrast was considered implausible (Iverson & Salmons 1995) or rare (Beckman et al. 2011) because it would require an OVER-SPECIFIED representation with both phonological features [voice] and [sg]. However, a growing number of studies reveal that this type of contrast is not uncommon across languages. Among languages exhibiting the contrast between prevoiced stops and stops with long lag VOT, Swedish (Helgason & Ringen 2008), Southern American English (Hunnicutt & Morris 2016), Najdi Arabic (AL-Gamdi, Al-Tamimi & Khattab 2019), and Qatari Arabic (Kulikov 2020) have been investigated within LR. Studies of stop voicing also suggest that this type of contrast may exist in Turkish (Öğüt et al. 2006), Ilami Kurdish (Abbaasian & Nourbakhsh 2015), and Sorani Kurdish (Ahmed 2019).

In addition, languages can possess either a three-way or a four-way laryngeal contrast. Languages with a three-way contrast, e.g. Eastern Armenian (Amirian 2017, Seyfarth & Garellek 2018), Hakha Chin (Lee & Harper Berkson 2019), Thai (Kessinger & Blumstein 1997), or Kurmanji Kurdish in Khorasan (Zirak 2014), utilize all three VOT categories. Languages with a four-way contrast, e.g. Hindi, Marathi (Lisker & Abramson 1964), Urdu (Hussain 2018), or Nepali (Schwarz et al. 2019) also utilize all three VOT categories, but the voiced aspirated series is articulated both with prevoicing and superimposed aspiration after stop release (Schwarz et al. 2019).

As a temporal cue, VOT has been shown to be sensitive to prosodic context. It tends to be more prominent in word-initial position than word-medially (e.g. Lisker & Abramson 1964, for English, but see Ringen & Kulikov 2012, who did not find difference in VOT between the two positions in Russian). Speech rate also affects realization of VOT so that duration of prevoicing and long lag VOT is increased in slow speech and decreased in fast speech (Kessinger & Blumstein 1997). Recent studies have shown that VOT (aspiration) in word-medial stops responds to changes in speech rate in the same fashion as in initial position (Schwarz et al. 2019).

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

The VOICED–VOICELESS distinction in word-medial stops has additional acoustic correlates: duration of glottal pulsing during closure (Lisker & Abramson 1964, Docherty 1992), percent closure voicing (Lisker 1986) and duration of an adjacent vowel (Chen 1970). Voiced stops are typically articulated with glottal pulsing during closure, although it is not uncommon for them to be voiced only for part of closure duration (Docherty 1992, Beckman et al. 2013). Although closure voicing in intervocalic stops is often viewed as a functional equivalent to voicing lead (prevoicing) in initial stops, their phonetic realization is not always identical. While onset of prevoicing typically occurs before the release so that glottal pulsing continues into the vowel, glottal pulsing in intervocalic stops can continue from the preceding vowel and cease in the middle of closure (Davidson 2016). As a result, a voiced stop can occasionally have a partially voiced closure and a voiceless release. It is of note that incomplete closure voicing is often found in voiced stops in aspirating languages (Docherty 1992).

Voiceless stops can also have a short voicing tail into the closure that continues from a preceding vowel, but they are typically voiceless for the most part of closure duration. The ratio of duration of glottal pulsing to closure duration (also voicing ratio, VR) is often used to evaluate the degree of closure voicing in intervocalic stops. Stops that are voiced for more than 50% of their duration are typically interpreted as voiced, and shorter voicing ratio is characteristic of voiceless stops (Slis 1986). A recent study of voicing in intervocalic stops in Nepali (Schwarz et al. 2019) reveals that closure voicing is also sensitive to speech rate manipulation. Glottal pulsing in phonologically voiced stops becomes longer in slow speech and shorter in fast speech to ensure that voicing continues during the entire closure.

There are several acoustic correlates whose role in the voiced–voiceless distinction is generally considered secondary. F0 (Haggard, Ambler & Callow 1970) and F1 (Liberman, Delattre & Cooper 1958) at the onset of the following vowel are two much discussed secondary correlates at this word position. Generally, voiced stops are typically associated with lower values of f0 (e.g. House & Fairbanks 1953, Ohde 1984) and F1 (e.g. Liberman et al. 1958, Summerfield & Haggard 1977) compared to the voiceless category. Although lower F1 is usually associated with voicing, the reasons behind this differ in true voice and aspirating languages. F1 lowering after phonologically voiced stops is usually attributed to the expansion of the vocal tract due to larynx lowering (Westbury 1983) or advancement of the tongue root (Westbury 1983, Kingston et al. 1997) aimed at creating rarefaction in the expanded supraglottal cavity in order to maintain effective glottal pulsing (Westbury & Keating 1986). Higher F1 after voiceless aspirated stops is a result of a delay in F1 transition (F1 cutback) after longer VOT (Stevens & Klatt 1974, Summerfield & Haggard 1977).

Researchers provided different phonetic explanations for f0 and F1 variation in voiced and voiceless stops. According to some (e.g. Ladefoged 1973, Hombert, Ohala & Ewan 1979, Löfqvist et al. 1989, Stevens 1998), the effect of voicing on f0 and F1 is an automatic consequence of the articulatory and/or aerodynamic conditions involved in voicing production, such as the tenseness of the vocal folds, the height of the larynx, or cutback of formant transition. Therefore, this variation is not directly controlled by the speakers. Other researchers have claimed in favor of an intentional and phonologically determined relationship between f0/F1 variations and VOT (Ohde 1984, Kingston & Diehl 1994, Dmitrieva et al. 2015). Furthermore, it was shown that the onset f0 and F1 enhance the perception of voicing in voiced stops (e.g. Liberman et al. 1958, Summerfield & Haggard 1977, Benkí 2001).

F2 transition is not often mentioned as a phonetic correlate of voicing, but some sources indicate that higher F2 is yet another acoustic aftermath of expansion of supraglottal cavity and advancement of the tongue root in voiced stops (Westbury 1983, Ahn 2018). This cue can be most noticeable in coronal stops, when voiced [d] is articulated closer to the dental area as the tongue is pushed forward as a result of expansion in the pharyngeal area (Bolla 1981, Ahn 2018). In a language that contrasts voiced or voiceless stops to their emphatic, or pharyngealized counterparts, F2 may become an important cue as it was shown to be lower in Arabic pharyngealized stops due to retraction of the tongue root (Ghazeli 1977, among others).

2.2 Traditional approach vs. laryngeal realism

Phonemic representation of laryngeal features among obstruents has been a topic of debate in phonological literature. Based on the physiological settings involved, Chomsky and Halle (1968, 328) defined four binary features to represent laryngeal contrasts in world's languages: [±tense], [±voice], [±heightened subglottal pressure], and [± glottal constriction]. Halle & Stevens (1971) proposed a different set of four binary features, namely [±spread glottis], [±constricted glottis], [±stiff vocal folds], and [±slack vocal folds]. While the two proposals could explain cross-linguistic differences in voicing and aspiration, they clearly lacked 'simplicity', one of the fundamental principles in the generative phonology. Later scholars mainly employed [±voice] (Keating 1984, Kingston & Diehl 1994) to explain voicing patterns across languages. The approach used in all these models is known as the 'traditional approach' (TA) (Honeybone 2005).

In TA, voicing contrast is displayed with the help of binary phonological features denoting the presence or absence of a feature with two values which have equal status. The phonetic realization of laryngeal contrast in terms of VOT (prevoicing or aspiration) is a function of language-specific rules of phonetic implementation. In Keating's (1984: 291) model, for instance, the binary feature [±voice] is phonetically implemented with three phonetic categories: {voiced} meaning fully voiced, {voiceless aspirated}, and {voiceless unaspirated}. As a result, voiced series is specified by [+voice] both in true voice and aspirating languages despite the fact that they are implemented by different VOT categories. In other words, VOT in TA is a phonetic detail not specified phonologically.

An alternative view to laryngeal representation which has recently received some significant attention is known as laryngeal realism (LR) (Honeybone 2005, Beckman et al. 2011, Beckman et al. 2013). Three types of evidence are usually considered to justify feature specification in LR: the phonetic realization of the segments in word initial position in terms of VOT patterns, diagnostics of speakers' control observed as effects of speech rate on VOT duration and the degree of intervocalic voicing, and phonological markedness and patterning of the segments (Schwarz et al. 2019).

LR is considered a phonetically-informed framework that employs privative laryngeal feature [voice] for voicing lead and [sg] for long lag VOT. The most common VOT pattern, short lag category, is claimed to be unmarked, or phonologically unspecified (Iverson & Salmons 1995).[1] This set of features directly encodes VOT typology in the word-initial position, where phonetic correlates are maximally contrastive (Schwarz et al. 2019), and it is assumed to account for most common VOT patterns among languages,[2] as shown in Table 2.

Cross-linguistically, it has been shown that temporal phonetic correlates mirroring phonological specifications tend to have longer duration in slower speech, while the unspecified categories remain unchanged. This behavior was found for oral and nasal vowels (Solé 1992), long and short vowels (Pind 1995), pre- and (post)aspiration (Pind 1995, Kessinger & Blumstein 1997, Allen & Miller 1999), and voicing (Kessinger & Blumstein 1997, Beckman et al. 2011, Schwarz et al. 2019), among others. LR argues that speech rate manipulation only influences segments specified with [voice] and [sg], but not phonologically unspecified segments (Beckman et al. 2011, Morris 2018, Schwarz et al. 2019, Kulikov 2020). This

[1] But see Vaux & Samuels (2005) and Kirby (2018) for an argument against the unmarkedness of short lag.

[2] An anonymous reviewer pointed out that some languages, e.g. Khmer contrast 'voiced' and 'voiceless' stops that show no or little difference in VOT (Kirby 2018).

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

Table 2 Laryngeal representation of four types of stops according to laryngeal realism.

Contrast	/b d g/	/p t k/	/p^h t^h k^h/	/b^h d^h g^h/
Two-way	[voice]	[ø]		
		[ø]	[sg]	
	[voice]		[sg]	
Three-way	[voice]	[ø]	[sg]	
Four-way	[sg]	[ø]	[sg]	[voice], [sg]

effect was first found both in languages with two-way and three-way laryngeal systems. But recently, Schwarz et al. (2019) demonstrated that in the production of voiced aspirated stops in Nepali, a language with a four-way voicing contrast, both voicing lead and long lag become longer when lowering speech rate.

The degree of voicing in intervocalic stops is another diagnostic of speakers' control over realization of laryngeal features (Beckman et al. 2013). The [voice]-specified voiced stops in true voice languages typically show ACTIVE intervocalic voicing being produced with a fully voiced closure (Jansen 2004, Ringen & Kulikov 2012). The requirement to actively maintain glottal pulsing in [voice]-specified stops may explain the fact that speakers tend to produce them with fully voiced closure both in slow and fast speech. But phonologically unspecified voiced stops in aspirating languages display PASSIVE voicing next to a sonorant segment (Jansen 2004). As a result, unspecified stops in these languages show variable or incomplete glottal pulsing being voiced only 62% of the time (Beckman et al. 2013).

In contrast, both [sg]-specified and unspecified voiceless stops block intervocalic voicing by displaying voicing only in 10–30% of the closure. As production of the intervocalic aspirated voiceless stops requires the significant opening in the glottis, the blocking of voicing from the preceding vowel is indeed expected. But the absence of passive voicing in unspecified voiceless stops in true voice languages is still an unanswered question in LR. Several explanations have been suggested. In line with generative formalism of Chomsky & Halle (1968), Beckman et al. (2013) argue that passive voicing in unspecified voiceless stops is blocked as a result of a language-specific rule that turns a privative feature value into a numerical value in phonetics. Thus, [voice]-specified stops become [9voice], but unspecified voiceless stops become [1voice], which ensures their ACTIVE DEVOICING (Jansen 2004). Voiced stops in aspirating languages lack specification for [voice], therefore they do not get a numerical value for this feature and can be passively and variably voiced in phonetics. Alternatively, in line with Kessinger & Blumstein (1997), Schwarz et al. (2019) argue that passive voicing in unspecified voiceless stops is avoided for perceptual reasons.

While LR adequately explains the typology of laryngeal contrasts in languages on the basis of VOT categories, it cannot fully account for the f0 patterns in voiceless stops in different laryngeal contrasts. In a recent study of f0 in French and Italian, true voice languages, Kirby & Ladd (2018) argue that f0 is raised after voiceless stops in these languages in the same fashion as it is raised in American English, an aspirating language. They claim that this situation is somewhat problematic for LR because different phonological specification of voiceless stops in true voice and aspirating languages should correlate with different and clear-cut acoustic realization. It is possible, however, that similar realization of f0 is the result of mere absence of glottal pulsing in the voiceless category in each of these languages. In a situation when the contrast is predominantly ensured by VOT as a primary cue to voicing, the role of f0 as a secondary cue becomes less important (e.g. van Alphen & Smits 2004). Speakers may variably use f0 to enhance the contrast in voiceless stops rather than target specific contrastive values. But should a language have more than one voiceless category, the difference in f0 might emerge. Studies of voicing in Khmer, Vietnamese and Thai reveal that voiceless unaspirated stops are often produced with slightly lower f0 than voiceless aspirated stops, signaling the contrast between the two otherwise voiceless categories (Kirby 2018).

2.3 Voicing contrast in Arabic varieties

Arabic dialects have a two-way contrast between voiced and voiceless stops; however, its realization varies from one dialect to another (Table 3). Lebanese, Egyptian, and Palestinian Arabic are examples of true voice languages (Yeni-Komshian, Caramazza & Preston 1977, Rifaat 2003, Tamim 2017). Jordanian Arabic (Khattab, Al-Tamimi & Heselwood 2006), in contrast, reveals a pattern typical for aspirating languages. In addition, it appears that the supposedly rare laryngeal contrast between voicing lead and long lag VOT is not uncommon among Arabic varieties. Based on the existing literature, five varieties spoken in the eastern part of the Arab world (and probably more dialects in this region) possess this type of contrast (shown as bold italics in Table 3): Mosuli Arabic (Rahim & Kasim 2009), Abha Arabic (Al Malwi 2017), Najdi Arabic (AL-Gamdi et al. 2019), Qatari Arabic (Kulikov 2020, Kulikov, Mohsenzadeh & Syam, published online 2 November 2021), and Khuzestani Arabic (KhA) (Bahrani 2022). Kulikov (2020) showed that similar to Swedish (Beckman et al. 2011), the duration of prevoicing and long lag in Qatari Arabic increases when speech rate is lowered.

Table 3 Word-initial VOT patterns in some Arabic varieties. Bold italics indicate the contrast between voicing lead and long lag VOT.

Arabic dialect	Voiced /b d g/	Voiceless /p t k/
Egyptian, Lebanese, Palestinian	voicing lead	short lag
Jordanian	short lag	long lag
Mosuli, Abha, Najdi, Qatari	***voicing lead***	***long lag***
Khuzestani (this study)	***voicing lead***	***long lag***

In addition to voiced and voiceless stops /b d t k/, most Arabic dialects have voiced and voiceless ‘emphatic’, or pharyngealized, plosives /dˤ/ and /tˤ/, produced with secondary construction in the posterior area, and the uvular stop /q/, which has primary constriction in the same posterior area (Ghazeli 1977). It is of note that these stops have different or irregular VOT patterns in many Arabic dialects and are often excluded from the analysis of the voicing contrast in these languages (e.g. Olson & Hayes-Harb 2019). For example, VOT in emphatic /tˤ/ is typically shorter than in plain /t/ (Khattab et al. 2006, Alzoubi 2016), and uvular /q/ does not have a voiced counterpart or it merged to a glottal stop /ʔ/ in many dialects of Levant or changed to /g/ in most of the Arabic words in eastern varieties. In many Gulf and Mesopotamian dialects, /dˤ/ is missing due to a merger with /ðˤ/, making /tˤ/ the only voiceless stop category with short lag VOT. Kulikov et al. (published online 2 November 2021) argue that the short lag VOT of /tˤ/ in Qatari Arabic is not just a mechanical consequence of pharyngealization because the latter does not spread to the adjacent long lag VOT. Rather, short lag is a phonological requirement necessary to distinguish the voiceless emphatic stop /tˤ/ from its plain counterpart /t/.

Similar to Qatari Arabic, KhA has voiced and voiceless plain stops /b d g t k/ and voiceless emphatic /tˤ/. In addition, it also has voiceless /p/ in non-Arabic words, and voiceless /q/ in both Arabic and non-Arabic words (Bahrani & Modarresi Ghavami 2021). KhA stop system is shown in Table 4. Voiced stops are consistently produced with voicing lead; plain voiceless stops are aspirated, and both emphatic /tˤ/ and uvular /q/ are voiceless unaspirated and do not have voiced counterparts. It is of note that coronal emphatics and uvulars are often reported to share some phonological specifications, e.g. feature [pharyngeal] (McCarthy 1994: 202) or [guttural] (Watson 2002: 38). To capture this generalization, we will refer to this group of stops with a post-velar constriction as GUTTURAL in line with Watson.

Therefore, KhA provides a convenient case to examine the behavior of three VOT categories in the same language in a situation when short lag VOT is associated with another phonological contrast in a language. To the best of our knowledge, no study investigated

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

Table 4 Stop consonants in Khuzestani Arabic.

	Bilabial	Coronal	Velar	Coronal emphatic	Uvular
Voiceless	/p/	/t/	/k/	/tˤ/	/q/
Voiced	/b/	/d/	/g/		

these guttural stops in relation to speech rate manipulation and the degree of intervocalic voicing. LR predicts that GUTTURAL /tˤ/ and /q/ in KhA are unspecified for the laryngeal feature in phonology and should not respond to speech rate manipulation. VOT in VOICED and VOICELESS stops should show such response and increase as speech rate slows. In addition, we analyze and test the predictions of LR for word-medial stops. Intervocalic voiced stops should have fully voiced closure, but both groups of voiceless stops in intervocalic position are expected to have voiceless closure.

3 Voicing in initial and medial stops

3.1 Method

3.1.1 Participants

Seven female and five male speakers participated in the study. They were born to middle class families and raised either in Abadan or Khorramshahr. They had either non-existent or insufficient knowledge of any other language excluding Persian, which is the lingua franca language of the Iranian community. Their age was between 20 years and 39 years ($M = 29$). They did not report any speech or hearing impairment and were not informed about the purpose of the experiment.

3.1.2 Materials

We evaluated laryngeal state in stops in word initial and word medial positions. For the word-initial position, the stimuli were 62 short Arabic words and two non-words with initial voiced and voiceless stops ($n = 8$) at four places of articulation: bilabial, coronal (alveolar/dental), velar, and uvular. Voiceless coronal stops were either plain /t/ or emphatic /tˤ/, which was articulated with a secondary constriction in the posterior area. This yielded to five contrasts in place of articulation, but only three of them were possible in voiced stops. Each stop was produced before four vowels: /a/, /aː/, /iː/, or /uː/. The complete list of target words is shown in the appendix Table A1. Table 5 exemplifies stimuli before long /aː/.

Table 5 Examples of the stimuli in word initial position.

Place	Voiced	Voiceless
Bilabial	/baːb/	/pɑːs/[a]
Coronal	/daːχ/	/taːʤ/
Coronal emphatic		/tˤaːɾ/
Velar	/gaːl/	/kaːfi/
Uvular		/qaːb/

[a]The vowel was more retracted in loan words [pɑːs] and [pɑːrk].

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

Table 6 Examples of the stimuli in word medial intervocalic position.

Place	Voiced	Voiceless
Bilabial	/χabaɾ/	/kapaɾ/
Coronal	/madad/	/ħatam/
Coronal emphatic		/maṭˤaɾ/
Velar	/ɾagasˤ/	/ħakam/
Uvular		/ʔaqal/

For the word-medial position, the stimuli were disyllabic Arabic words ($n = 16$) with intervocalic voiced and voiceless stops ($n = 8$) at five contrastive places of articulation. Each stop was produced between low vowels /a/ or /aː/. The complete list of target words is shown in the appendix Table A2. Table 6 exemplifies stimuli before short /a/. The vowel following the target stop was invariably unstressed.

3.1.3 Procedure

The recordings were made in a quiet room using a Sony ICD-PX440 recorder (320 kbps, 44,100 Hz). Target words were presented to the participants in Arabic orthography. The participants pronounced (read) each target word in a carrier phrase [gaːl faːɾes . . . mɑɾteːn] 'Fares said . . . two times' at two speaking rates, slow and fast, which is an adopted practice in studies of rate effects on VOT (e.g. Kessinger & Blumstein 1997, Beckman et al. 2011). In the slow rate condition, the participants were instructed to pronounce the phrase at a comfortable tempo. In the fast rate condition, the participants were asked to pronounce the phrase as fast as they could but not at the expense of clarity. They were instructed to speak as if they were going to say something important to a person who is about to leave the room.

3.1.4 Acoustic analysis

The recorded materials were evaluated for naturalness by one of the authors, a native speaker of Arabic. Ninety tokens (3.2% of the recorded items) were discarded due to mispronunciation ($n = 41$) or non-plosive realization of uvular stops ($n = 49$), as /q/ is optionally realized as a voiced fricative in KhA (Bahrani & Modarresi Ghavami 2021). A total of 3030 word-initial tokens and 730 word-medial tokens were submitted to acoustic analysis.

The segment boundaries were set manually in PRAAT (Boersma & Weenink 2021). The segment preceding word initial target stop was voiceless [s] to ensure there is no carry-over of glottal pulsing from a preceding segment. VOT was measured as timing between the stop release and the onset of voicing. Both waveforms and spectrograms were used to identify the beginning of glottal pulses. F0, F1 and F2 were measured from LPC spectra obtained with a 25 ms Hamming window at vowel onset. The amount of stop closure voicing was evaluated using absolute values of voicing duration and relative proportion of voicing during closure (hence, VR, or voicing ratio). The onset of stop closure was marked at the point of cessation of F2 and significant drop of periodic energy. The offset of voicing was marked at the point of cessation of glottal pulsing. The landmarks for acoustic measurements are summarized in Figure 1.

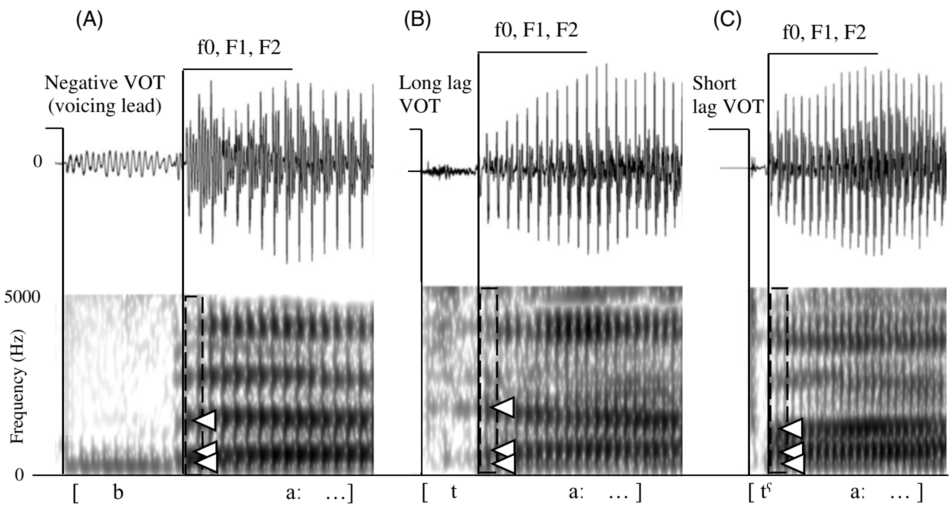

Figure 1 Examples of acoustic measurements: (A) negative VOT (voicing lead) in [baːt] 'slept', (B) long lag positive VOT in [tʰaːb] 'repented', (C) short lag positive VOT in [tˤaːɾ] 'flew'. F0, F1 and F2 were measured at vowel onset.

3.1.5 Data analysis

The acoustic data were submitted to several linear mixed effects models using the *lmer* package (Bates et al. 2015) in R (R Core Team 2021). Each acoustic cue was used as a dependent variable in a separate mixed-effects model. Fixed effects in the model were independent variables whose effect is investigated (e.g. stop class or place of articulation). When a fixed effect had more than two levels, it was first evaluated using a Log Likelihood (chi-square) test by comparing the model fit with and without the factor. Random effects in the model were sources of variance due to random selection of a subset of population (e.g. speakers or words).

Following Barr et al. (2013), we started selecting the optimal model with the most saturated one that included both random intercept and random slopes. Random intercept is a mean difference between each speaker or word; random slope explains additional variation in a fixed effect in relation to a given random effect. For example, the effect of place of articulation may vary from one speaker to another due to individual differences. Similarly, the effect of speech rate may vary from one word to another due to number of segments in a word. When adding some effects did not improve the model's performance, the simpler model was selected for the benefit of better convergence (Matuschek et al. 2017). The *p*-values for factor levels were calculated using the *lmerTest* package (Kuznetsova, Brockhoff & Christensen 2017).

3.2 Results I: Initial stops

3.2.1 Phonetic context

Before analysing acoustic properties of stops and effect of speech rate, we looked into acoustics of vowels. Since KhA is predominantly a vernacular dialect, we wanted to make sure speakers produced vowels in the reading tasks as intended. The summary of the vocalic cues is given in Table 7.

The results showed that the vowels were produced as intended. Formant values were consistent with vowel qualities reported in Bahrani & Modarresi Ghavami (2021). Duration of long high vowels /iː/ and /uː/ was shorter than that of low /aː/, following cross-linguistic tendencies (e.g. Peterson & Lehiste 1962). Duration of short /a/ was 65% shorter compared to long /aː/.

Table 7 Means and standard deviations (in parentheses) for major vocalic cues in KhA vowels.

Vowel		Duration (in ms)	F1 (in Hz)	F2 (in Hz)
/a/		116 (42)	545 (105)	1904 (315)
	Emphatic		672 (131)	1418 (198)
/aː/		178 (53)	612 (108)	1648 (341)
	Emphatic		691 (95)	1310 (130)
/iː/		151 (55)	360 (52)	2507 (320)
	Emphatic		471 (83)	1963 (280)
/uː/		150 (53)	399 (58)	1048 (286)
	Emphatic		457 (66)	948 (161)

The four vowels were also distinct in formant frequencies. As expected, long /aː/ was realized as a low central vowel, long /iː/ – as high front vowel, long /uː/ as a high back vowel, and short /a/ as a mid front vowel. All vowels were considerably lowered and retracted next to emphatic coronal and uvular stops revealing higher F1 and lower F2.

3.2.2 Word duration

Next, we analyzed duration of words to make sure speakers produced the desired difference in the two speech rate conditions. We used total word duration as a proxy of speech rate. It was analyzed in a linear mixed effects model with the following equation:

$$\text{Word\_duration} \sim 1 + \text{Voicing} + \text{SpRate} + (\text{Voicing} \times \text{SpRate}) + (1 + \text{Voicing} + \text{SpRate}|\text{Speaker})$$

Adding gender as a between-subject fixed effect did not improve the model ($p = .794$). The results are summarized in Table 8, and effects are plotted in Figure 2.

Table 8 Summary of fixed effects in a linear model examining word duration.

Factor	Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
(Intercept)		226.0	12.2	18.60	< .001	***
Rate	Slow	127.4	1.8	72.65	< .001	***
Voicing	Voiceless	10.9	7.0	1.55	.125	
	Slow: Voiceless	5.7	2.5	2.28	.023	*

$* = p < .05$; $*** = p < .001$

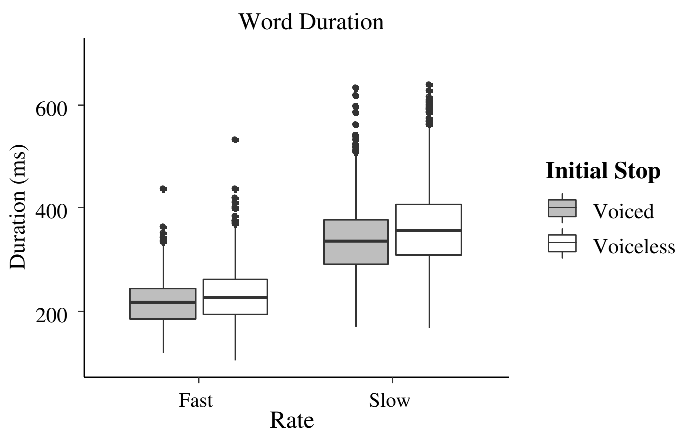

Figure 2 Boxplots of word durations as a function of initial stop voicing and rate condition.

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

The model revealed the effect of rate condition. Words in the slow condition were on average 127 ms longer than in the fast condition. The effect of voicing was not obtained, but the interaction with rate revealed that words with initial voiceless stops were 6 ms longer than words with initial voiced stops in the slow rate condition.

3.2.3 VOT

Next, we examined and analyzed VOT in initial stops. Observation of VOT distributions in each stop category (see Figure 3) revealed three types of VOT in the data that corresponded to three types of VOT commonly found in world's languages (Lisker & Abramson 1964). The majority of phonologically voiced stops /b d g/ were produced with negative VOT, or voicing lead, ranging from −200 ms to 0 ms. However, 7% of voiced stops ($n = 42$) were produced without prevoicing and had short lag positive VOT ranging from 5 ms to 35 ms. Voiceless stops /p t k/ were largely produced with positive VOT ranging from 11 ms to 128 ms, which we define as long lag, and guttural stops /tˤ q/ were produced with positive VOT ranging from 2 ms to 40 ms, which we define as short lag.[3] Table 9 summarizes the means

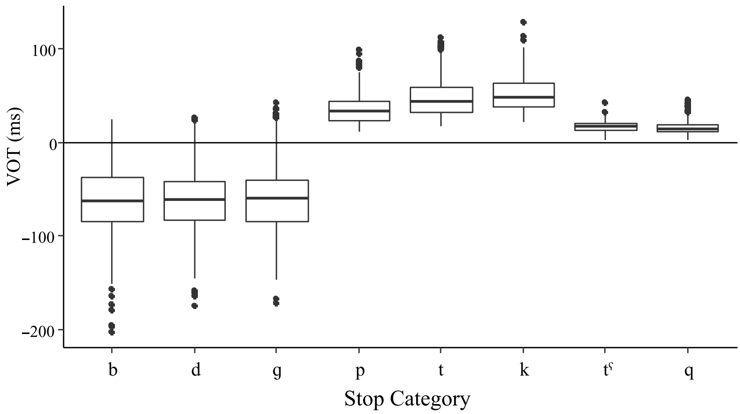

Figure 3 Boxplots of VOT in initial position across stop classes and speech rates.

Table 9 Summary of VOT durations in initial stops (in ms).

Stop class	*n*	% prevoiced	Mean	SD
Voiced /b d g/	1147	93%	−60	38
prevoiced	1063		−67	32
Voiceless /p t k/	1145	0%	45	19
Guttural /tˤ q/	738	0%	16	7

[3] The anonymous reviewer asked if the voiceless tokens with VOT as short as 11 ms should be called long lag. It is true that the short lag and long lag categories reveal some overlap within the range between 11 ms and 30 ms, which is larger in fast speech. This is not uncommon in aspirating languages. For example, Allen & Miller (1999) report a similar overlap in English. We believe that the nature of a category is defined by the mean and modal values rather than by outliers. Crucially, the majority of the voiceless stops were produced with VOT longer than 30 ms ($M = 45$ ms), and the majority of guttural stops had VOT shorter than 30 ms ($M = 16$ ms). We take this difference as a sufficient empirical ground to distinguish between long lag and short lag VOT categories in voiceless and guttural stops. Also, further analysis revealed that the two types of VOTs behaved differently in response to speech rate manipulation.

and standard deviations for the three types of VOT. Therefore, for subsequent analysis we divided all stops into three stop classes: (i) VOICED /b d g/, (ii) VOICELESS /p t k/, and (iii) GUTTURAL /tˤ q/.

These observations were confirmed in a liner mixed effects model with the following equation:

$$\text{VOT} \sim 1 + \text{StopClass} + \text{SpRate} + (\text{StopClass} \times \text{SpRate}) + (1 + \text{StopClass}+ \text{SpRate}|\text{Speaker}) + (1|\text{Word})$$

Stop class levels were coded as contrasts using Helmert coding (Davis 2010), in which voiced stops (coded 2/3) were compared to all phonetically voiceless stops (each class coded −1/3) at level 1, and guttural stops /tˤ q/ (coded −0.5) were compared to voiceless stops /p t k/ (coded 0.5) at level 2. The model is summarized in Table 10.

Table 10 Summary of fixed effects in a *lme* model examining VOT in initial stops.

Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
(Intercept)	2.7	1.7	1.58	.128	
Voiced vs. Voiceless	−72.1	5.6	−12.87	< .001	***
Voiceless vs. Guttural	21.1	3.1	6.83	< .001	***
Slow	5.1	1.4	3.62	.004	**
Slow rate: Voiced vs. Voiceless	38.8	1.6	24.76	< .001	***
Slow rate: Guttural vs. Voiceless	14.7	2.0	7.51	< .001	***

** = $p < .01$; *** = $p < .001$

We found an effect of stop class: each class was different from each other. The coefficient for voiced stops was negative indicating prevoicing, and positive VOT in the voiceless stops was significantly longer than in guttural stops. Importantly, stop class interacted with speech rate, revealing that each type of VOT reacted to speech rate manipulation differently. The slope was steeper for voiced stops ($\beta = 39$ ms, $p < .001$) than for all voiceless stops, and it was also steeper for voiceless stops ($\beta = 15$ ms, $p < .001$) as compared to guttural stops.

In order to explore the stop class-by-rate interaction in depth, we ran separate *lme* models for each stop class with the follow formula:

$$\text{VOT} \sim 1 + \text{Place} + \text{SpRate}+ (1 + \text{Place} + \text{SpRate}|\text{Speaker}) + (1|\text{Word})$$

The model included the effect of place of articulation (bilabial, coronal, velar, for voiced and voiceless stops; coronal, uvular, for guttural stops). Places of articulation were compared using backward difference coding (Davis 2010), in which coronal place was compared to bilabial place at level 1, and velar place were compared to coronal place at level 2. Tables 11–13 summarize the models. The effects are plotted in Figure 4.

For voiced stops, only a strong effect of rate was found (Table 11). Prevoicing (negative VOT) was on average 31 ms longer in the slow rate condition. No significant difference was found between places of articulation.

For voiceless stops, the effect of speech rate was also significant, but its magnitude was smaller (Table 12). Long lag positive VOT was on average 15 ms longer in the slow rate condition. Effect of place of articulation was also obtained, revealing that long lag VOT was 11 mm longer in coronal stops and velar stops than in bilabial stops.

For guttural stops, no effect of speech rate was found (Table 13). Short lag positive VOT in emphatic coronal and uvular stops did not change in response to rate manipulation. The effect of place of articulation was significant but very small, with a negligible differences of 1.3 ms.

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

Table 11 Summary of fixed effects in a lme model examining VOT in voiced stops.

Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
(Intercept)	−45.4	3.9	−11.46	< .001	***
Slow	−30.9	3.4	−9.39	< .001	***
Coronal vs. Bilabial	4.4	2.5	1.77	.104	
Velar vs. Coronal	1.4	2.3	0.76	.466	

*** = $p < .001$

Table 12 Summary of fixed effects in a lme model examining VOT in voiceless stops.

Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
(Intercept)	37.4	2.0	18.63	< .001	***
Slow	15.2	1.7	9.08	< .001	***
Coronal vs. Bilabial	11.0	3.3	3.31	.003	**
Velar vs. Coronal	−2.8	3.5	−0.81	.426	

** = $p < .01$; *** = $p < .001$

Table 13 Summary of fixed effects in a lme model examining VOT in guttural stops.

Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
(Intercept)	16.1	0.8	21.44	< .001	***
Slow	0.5	0.5	0.96	.359	
Uvular	−1.3	0.6	−2.10	.05	*

* = $p < .05$; *** = $p < .001$

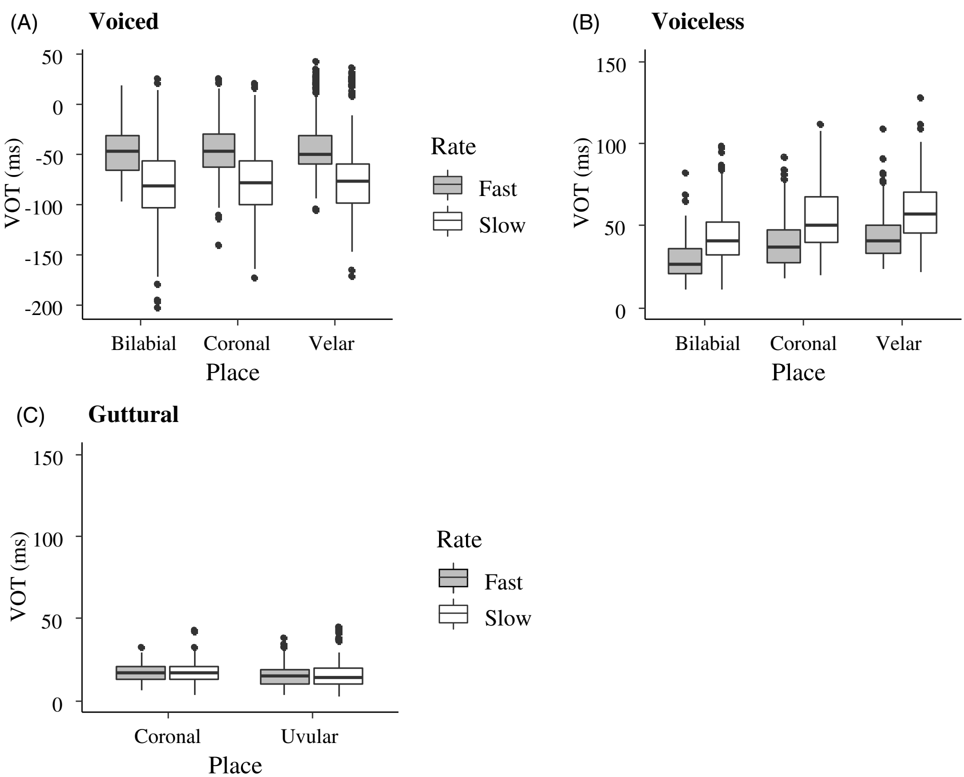

Figure 4 VOT in voiced, voiceless and guttural initial stops in slow and fast rate conditions broken down by place of articulation.

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

3.2.4 VOT and speech rate

The analysis revealed that the three types of VOT react to manipulation with speech rate differently. In line with previous studies of rate effects on VOT across languages (e.g. Kessinger & Blumstein 1997, Beckman et al. 2011, Schwarz et al. 2019, Kulikov 2020), we found a decrease in duration of negative VOTs and long lag VOTs in fast rate condition. Importantly, the decrease in duration of prevoicing and aspiration was not driven by changes of VOT values in outliers but rather it affected the whole distributions, as shown in Figure 5. The mode of negative VOTs shifted from −80 ms in slow speech to −50 ms in fast speech; the mode of long lag positive VOTs shifted from 45 ms in slow speech to 30 ms in fast speech. In contrast, no change in the range or modal values was found for the distribution of short lag positive VOTs. They were virtually unaffected by speech rate manipulation.

Our next analysis looked into a continuous relationship between VOT and speech rate. Recall that the relationship between VOT and speech rate is diagnostic to phonological features of contrast in a language. When word duration decreases in fast speech rate, speakers also decrease VOT values in initial stops, as shown in Figure 6. However, according to Beckman et al. (2011), this decrease affects only temporal cues that are correlates of contrastive phonological features. Under Laryngeal Realism, prevoicing manifests [voice] and long lag positive VOT manifests [spread glottis]. Short lag positive VOT is assumed to have no corresponding laryngeal feature.

To confirm empirical observations about VOT and continuous speech rate, a linear mixed effects model with the following formula was fitted to the data:

$$\begin{aligned}\text{VOT} \sim 1 + \text{StopClass} + \text{SylRate} + (\text{StopClass} \times \text{SylRate}) + (1 + \text{StopClass} + \\ \text{SylRate}|\text{Speaker}) + (1|\text{Word})\end{aligned}$$

It used number of syllables per second as a proxi to continuous speech rate. Stop classes were coded to compare contrasts between short lag VOTs in guttural stops (reference category, 0) to negative VOTs in voiced stops (−1) and to long lag positive VOTs in voiceless stops (1). Speaker and word were used as random intercepts; stop class was added as a random slope for speaker. The results of the model are summarized in Table 14.

The effect of speech rate was not obtained ($p = .276$), indicating that short lag VOTs in guttural stops did not decrease as speech became faster. The effect of stop class was significant ($p < .001$) indicating that duration of VOT was longer in voiced and voiceless stops than in guttural stops. The positive coefficient for voiceless stops indicated that they were produced with long lag VOT. The negative coefficient for voiced stops indicated that they were produced with robust prevoicing. Absolute duration of VOT was longer in voiced

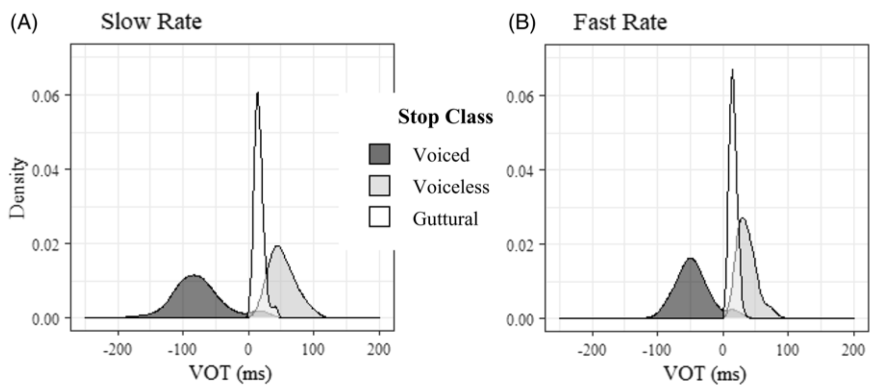

Figure 5 Shift in VOT distributions of voiced and voiceless initial stops in response to speech rate manipulation. No shift occurred in guttural stops.

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

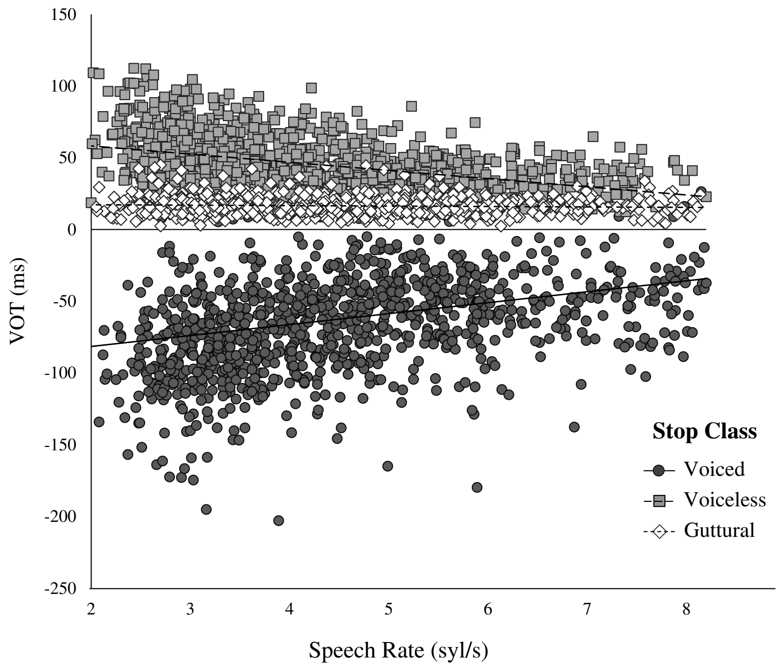

Figure 6 Effect of speech rate on VOT for stops in initial position.

Table 14 Summary of fixed effects in a linear model examining relationship between VOT and speech rate.

Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
(Intercept)	19.1	3.8	4.98	< .001	***
Voiced	−134.8	6.5	−20.83	< .001	***
Voiceless	54.5	5.2	10.47	< .001	***
Speech rate	−0.6	0.5	−1.09	0.276	
Voiced: Speech rate	12.7	0.7	17.03	< .001	***
Voiceless: Speech rate	−6.4	0.7	−8.38	< .001	***

*** = $p < .001$

stops than in voiceless stops. Significant interactions between speech rate and stop class ($p < .001$) revealed that the effects of speech rate were present in voiced stops and voiceless stops. Voicing duration in voiced stops had a steeper slope than duration of positive VOT in voiceless stops.

3.2.5 Spectral cues (f0, F1, F2)

Finally, we analyzed secondary, spectral cues to voicing in initial stops: f0, F1 and F2. This was performed to evaluate the glottal state for each category of VOT in a series of linear mixed effects models with the formula:

$$\text{Cue} \sim 1 + \text{StopClass} + \text{SpRate} + \text{Gender} + (\text{StopClass} \times \text{SpRate}) + (1 + \text{StopClass} + \text{SpRate}|\text{Speaker}) + (1|\text{Word})$$

The models examining F1 and F2 also included vowel as a fixed factor, but the differences in formant frequencies between vowels were predicted and are not reported here. All models

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

Table 15 Summary of fixed effects in a linear model examining f0.

Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
(Intercept)	236.8	7.3	32.42	< .001	***
Voiced vs. Voiceless	−5.9	1.0	−6.14	< .001	***
Voiceless vs. Guttural	3.1	1.5	2.04	.048	*
Slow	−12.1	0.6	−21.33	< .001	***

* $= p < .05$; *** $= p < .001$

Table 16 Summary of fixed effects in a linear model examining F1.

Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
(Intercept)	565.0	11.8	47.74	< .001	***
Voiced vs. Voiceless	−83.8	8.7	−9.60	< .001	***
Voiceless vs. Guttural	−84.0	13.8	−6.09	< .001	***
Slow	−24.0	1.9	−12.37	< .001	***

*** $= p < .001$

Table 17 Summary of fixed effects in a linear model examining F2.

Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
(Intercept)	1728.9	71.8	24.01	< .001	***
Voiced vs. Voiceless	−327.2	44.1	−5.30	< .001	***
Voiceless vs. Guttural	121.4	69.8	3.07	.008	**
Slow	4.4	6.5	0.67	.501	
Voiced: Slow	−43.6	14.7	−2.97	.003	**
Guttural: Slow	46.3	14.5	3.20	.001	**

** $= p < .01$; *** $= p < .001$

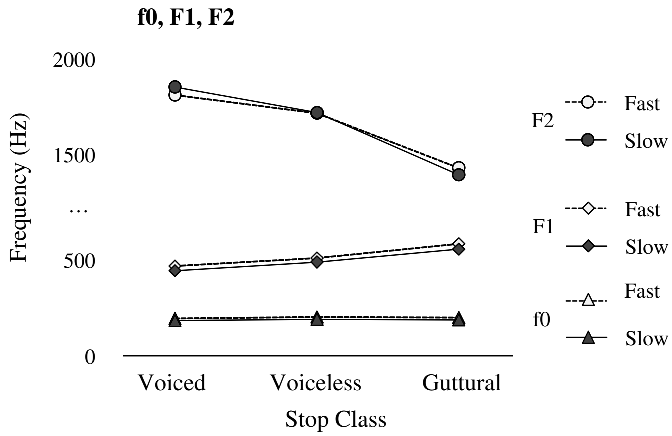

Figure 7 Effects of stop class and speech rate on spectral cues (f0, F1, F2) in initial stops.

also included gender as a fixed between-subject effect. It was predictably significant ($p <$ .001) for all spectral cues indicating lower frequencies in men. Importantly, it did not interact with other factors suggesting that both genders performed in a unified fashion; therefore, we do not discuss it in this paper. The results are summarized in Tables 15–17. The effects are plotted in Figure 7.

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

For f0 (Table 15), we found a significant effect of stop class. F0 was 6 Hz lower after voiced stops than after voiceless stops, and 3 Hz higher after voiceless stops than after guttural stops. The effect of speech rate was also significant revealing that pitch was lower by 12 Hz in slow rate condition.

For F1 (Table 16), the effect of stop class was significant. F1 was 84 Hz lower after voiced stops than after voiceless stops, and 84 Hz lower after voiceless stops compared to guttural stops.

For F2 (Table 17), the effect of stop class was also significant. F2 was 327 Hz lower after voiceless stops than after voiced stops, and 121 Hz higher after voiceless stops compared to guttural stops. The effect of speech rate was not significant, but interaction with stop class revealed that F2 significantly increased by 44 Hz in slow speech after voiced stops, but it decreased by 46 Hz in slow speech after guttural stops.

3.2.6 Interim summary

The results showed that phonologically voiced stops /b d g/ were produced with phonetic voicing. The analysis of VOT revealed that these stops were predominantly prevoiced in initial position. The analysis of spectral cues showed that the glottal state was consistent with voicing: both f0 and F1 were lower suggesting the larynx was lowered to facilitate vibration of the vocal folds. Phonologically voiceless stops /p t k/ were produced as voiceless aspirated. They had long-lag positive VOT and higher f0 and F1. The guttural stops /tˤ q/ were produced as voiceless unaspirated. They had short-lag positive VOT and a glottal state consistent with phonetic voicelessness (higher f0 and F1). It is of interest that F2 also correlated with voicing in KhA stops. Not only lower F2 predictably indicated emphatic phonation, but higher F2 was consistent with voicing.

The results suggest that the three types of VOT responded differently to speech rate manipulation. Prevoicing in voiced stops and aspiration in voiceless stops gradually increased

Table 18 Summary of fixed effects in a linear model examining word duration.

Level	Estimate	*SE*	*t*-value	Pr(>|t|)	Sig.
(Intercept)	304.0	22.8	13.34	< .001	***
Slow	109.0	11.4	9.59	< .001	***
Voiceless	34.3	25.6	1.34	.201	

*** = $p < .001$

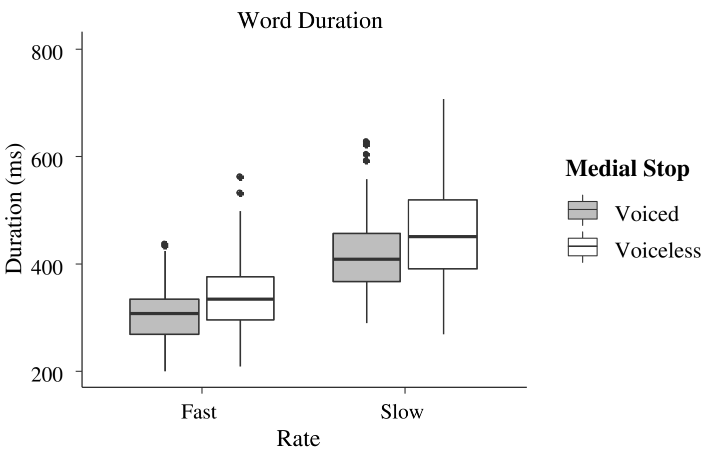

Figure 8 Mean word duration as a function of medial stop voicing and rate condition.

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

as speech became slower. But short-lag VOT in guttural stops was not affected by rate and remained stable across rate conditions.

3.3 Results II: Medial stops

Using the routine established in the analysis of initial stops, we first examined whether speech rate manipulation produced a desired effect on duration of target words. The results are summarized in Table 18, and effects are plotted in Figure 8. The data were fitted into a linear mixed effects model with rate condition (fast, slow) and phonetic stop voicing (voiceless, voiced) as fixed effect, speaker and word as random intercepts, and speech rate and stop voicing as random slopes for speaker.

The model revealed the effect of rate condition. Words in the slow condition were on average 109 ms longer than in the fast condition. The effect of voicing was not obtained. Words with medial voiceless stops were slightly longer than words with voiced stops, but this difference did not reach significance level.

3.3.1 VOT and closure voicing

Next, we examined distributions of VOTs in voiceless stops. We found the same types of VOT as in initial stops (Table 19A, Figure 9A). Voiceless /p t k/ were produced with long lag positive VOT averaging at 41 ms, which was very similar to what we found in initial stops ($M = 45$ ms). Guttural stops /tˤ q/ were produced with short-lag positive VOT averaging at 16 ms, which was also virtually identical to the type of VOT in guttural stops in initial position ($M = 16$ ms).

Table 19 Summary of VOT and closure voicing durations in medial stops (in ms).

A. VOT	*n*	% prevoiced	Mean	SD	
Voiceless /p t k/	283	0%	41	13	
Guttural /tˤ q/	160	0%	16	6	
B. Closure voicing	***n***	**% fully voiced**	**Mean**	**SD**	**Voicing ratio**
Voiced (b d g)	288	94%	55	15	.98
fully voiced	270		56	14	1.0
Voiceless /p t k/	283	0%	12	4	.17
Guttural /tˤ q/	160	0%	11	5	.16

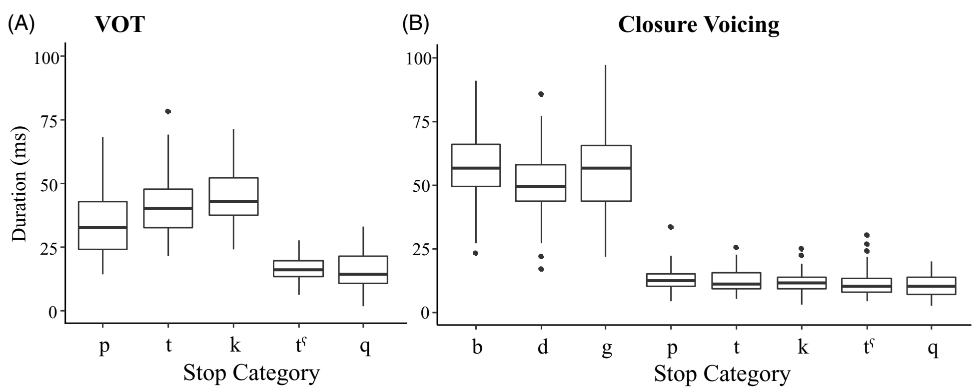

Figure 9 Boxplots of (A) VOT and (B) closure voicing in medial stops.

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

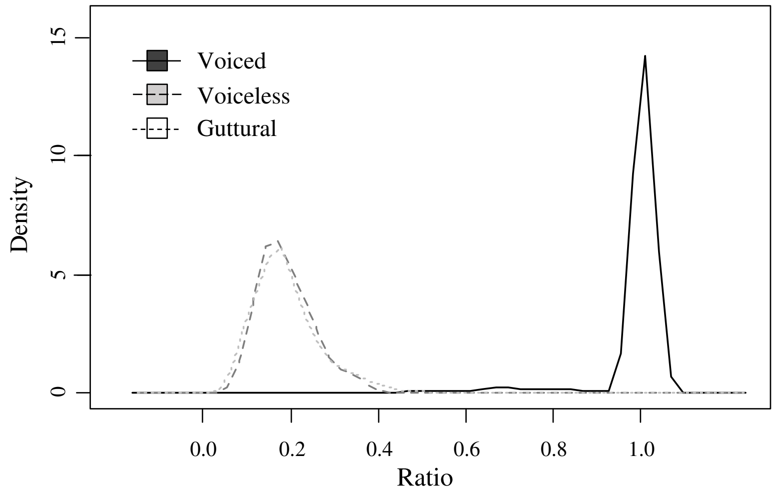

Figure 10 Proportion of closure voicing in three classes of medial stops.

It was not possible to identify VOT for voiced stops due to the continuous nature of stop voicing in intervocalic position. Although the majority (94%) of voiced /b d g/ were fully voiced, the glottal pulsing started in the preceding vowel and could not be measured in the same way as prevoicing in initial stops. We found that glottal pulsing in voiced stops continued throughout the entire closure averaging at 56 ms (Table 19B), and only in rare occasions (6% of cases) it ceased before the release. The ratio of voicing duration to the duration of closure (voicing ratio, VR) in voiced stops was 98%. Phonetically voiceless stops (both plain and guttural), in contrast, were articulated with closure that was essentially voiceless. It was voiced only for a small part, with a very short voicing tail that ended 12 ms after the onset of stop closure (Table 19B, Figure 9B). VR was 17% in voiceless stops and 16% in guttural stops. Figure 10 exemplifies the differences in closure voicing between voiced and voiceless stops.

3.3.2 Effect of speech rate

Next, we compared distribution of VOT and closure voicing of medial stops in the two speech rates. We found the same tendency as in initial stops. Longer duration in slow speech was found only for long lag VOT of voiceless stops but not for short lag VOT of guttural stops (Figure 11). The change affected the whole distribution, shifting the maximal and modal values from 60 ms and 35 ms in fast speech to 78 ms and 40 ms in slow speech. Similarly, duration of closure voicing in voiced stops was longer at slow rate in order to maintain it throughout the entire closure, but duration of a short voicing tail in all phonetically voiceless (both plain and guttural) stops remained stable across rates (Figure 12). Again, this change in voicing duration in voiced stops affected the whole distribution. The maximal and modal values changed from 96 ms and 60 ms in fast speech to 119 ms and 50 ms in slow speech.

We evaluated these observations in a series of linear mixed effects models with the formula:

$$\text{Cue} \sim 1 + \text{StopClass} + \text{SpRate} + (\text{StopClass} \times \text{SpRate}) + (1 + \text{StopClass} + \text{SpRate}|\text{Speaker}) + (1|\text{Word})\,.$$

Stop classes were coded using Helmert coding (Davis 2010) for contrasts between voiced stops (reference category) and all voiceless stops at level 1, and between voiceless and

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

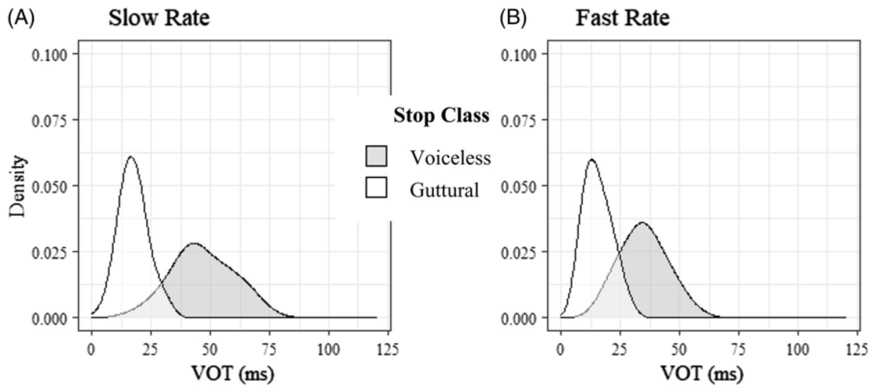

Figure 11 Shift in distributions of VOT in voiceless medial stops in response to speech rate manipulation. No shift occurred in the guttural stops.

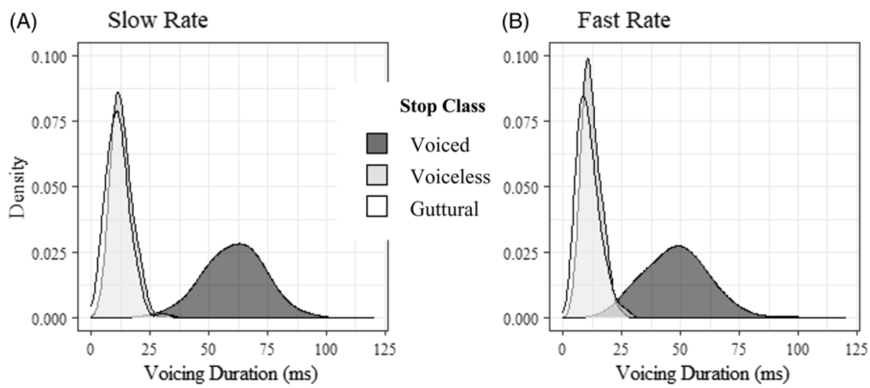

Figure 12 Shift in distributions of closure voicing in voiced medial stops in response to speech rate manipulation. No shift occurred in the voiceless and guttural stops.

guttural stops at level 2. Separate models were fitted to evaluate VOT in voiceless stops, and absolute duration of closure voicing in all medial stops.

For VOT (Table 20), we found significant effects of stop class and speech rate and interaction. VOT was 14 ms shorter in guttural stops that in voiceless stops. VOT increased by 7 ms in slow rate condition, but interaction revealed that this change was found only in voiceless stop. There was no increase in VOT guttural stops.

For closure voicing, we also found significant effects of stop class and speech rate, and interaction (Table 21). Duration of voicing in voiceless stop closure was by 37 ms shorter than in voiced closure, but there was no difference between duration of closure voicing in voiceless and in guttural stops. Speakers produced longer closure voicing in slow speech, but the interaction revealed the change affected only voiced stops. The negative coefficient indicated that increase in voicing duration did not occur in voiceless or guttural stops.

Similar to initial stops, the relationship between VOT/closure voicing in medial stops and speech rate was continuous as it appeared to be driven by duration of a word (Figure 13).

To confirm this observation, two separate linear mixed effects models were fitted to the data using the following formula:

$$\text{Cue} \sim 1 + \text{StopClass} + \text{SylRate} + (\text{StopClass} \times \text{SylRate}) + (1 + \text{StopClass} + \text{SylRate}|\text{Speaker}) + (1|\text{Word})$$

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

Table 20 Summary of fixed effects in a lme model examining VOT in medial stops.

Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
(Intercept)	24.1	1.7	14.03	< .001	***
Guttural	−13.7	2.2	−6.08	< .001	***
Slow	7.1	1.4	5.18	< .001	***
Slow: Guttural	−7.0	1.0	−7.07	< .001	***

*** = $p < .001$

Table 21 Summary of fixed effects in a lme model examining duration of closure voicing in medial stops.

Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
(Intercept)	24.0	1.3	18.73	< .001	***
Voiceless	−36.7	2.1	−17.33	< .001	***
Guttural	−1.2	1.9	−0.64	.531	
Slow	4.7	0.7	6.73	< .001	***
Slow: Voiceless	−12.2	1.2	−10.11	< .001	***

*** = $p < .001$

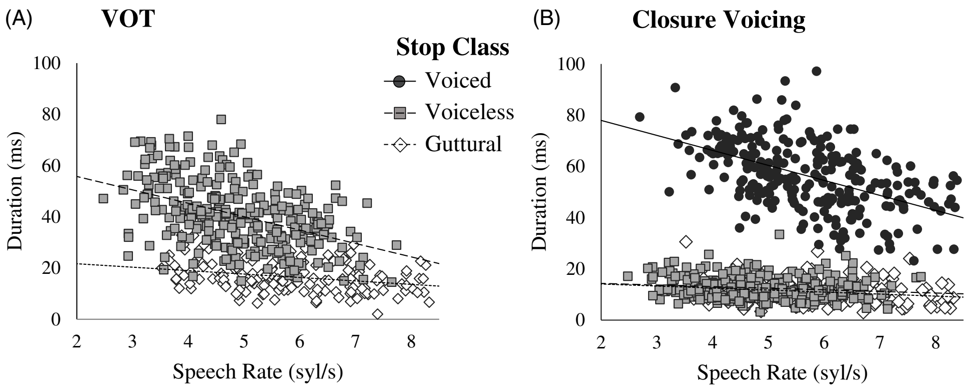

Figure 13 Effects of stop class and speech rate on spectral cues (f0, F1, F2) in medial stops.

The model used speech rate (number of syllables per second) as a continuous fixed effect (covariate). A smaller number of syllables per second indicated slower speech; a higher number was indicative of faster speech. Stop classes were coded by reverse Helmert coding (Davis 2010) to compare contrasts between short voicing tails in guttural stops and in voiceless stops at level 1 and to fully voiced closure in voiced stops at level 2. The model evaluating VOT was run on a subset of data that included only voiceless stops. The results of the models are summarized in Table 22.

For VOT (Table 22A), the effect of stop class was significant ($p < .0001$) indicating that VOT was significantly longer in voiceless stops. The effect of speech rate was not obtained ($p = .287$), indicating that short lag VOTs in guttural stops did not increase significantly as speech became slower. But a significant interaction between speech rate and stop class ($p < .001$) revealed that the effect of speech rate was present in voiceless stops. The coefficient was negative indicating that VOT increased as speech rate became slower.

For closure voicing (Table 22B), the effect of stop class was also significant ($p < .001$). Duration of closure voicing in voiceless stops was no different than in guttural stops, but it was significantly longer in voiced stops. The effect of speech rate was not obtained ($p = .110$), indicating that short voicing tails in guttural and voiceless stops did not increase significantly

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

Table 22 Summary of fixed effects in a linear model examining relationship between closure voicing and speech rate in medial stops.

Cue	Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
A. VOT	(Intercept)	45.9	2.6	17.93	< .001	***
	Voiceless	29.9	3.4	8.81	< .001	***
	Speech rate	−3.7	2.4	−1.07	.287	
	Voiceless: Speech rate	−13.9	1.4	−9.47	< .001	***
B. Closure voicing	(Intercept)	15.9	2.6	6.03	< .001	***
	Voiceless	3.5	4.3	0.81	.417	
	Voiced	78.1	3.9	19.79	< .001	***
	Speech rate	−0.7	0.4	−1.60	.110	
	Voiced: Speech rate	−16.1	2.3	−10.38	< .001	***

*** = $p < .001$

as speech became slower. But a significant interaction between speech rate and stop class ($p < .001$) revealed that the effect of speech rate was present in voiced stops. The coefficient was negative indicating that voicing duration increased as speech rate became lower.

3.3.3 Spectral cues (f0, F1, F2)

Finally, we analyzed secondary, spectral cues to voicing (f0, F1 and F2) in medial stops in order to evaluate the glottal state for each stop class in a series of linear mixed effects models using the formula:

$$\text{Cue} \sim 1 + \text{StopClass} + \text{SpRate} + (\text{StopClass} \times \text{SpRate}) + (1 + \text{StopClass} + \text{SpRate}|\text{Speaker}) + (1|\text{Word})$$

The results are summarized in Table 23. The effect of gender was also obtained for all cues, but it is not discussed here. Quite predictably, frequencies were significantly higher for female speakers ($p < .001$). The effects of stop class and speech rate are plotted in Figure 14.

For f0 (Table 23A), we found a significant effect of stop class. F0 was 6 Hz lower after voiced stops than after all voiceless stops. No difference in pitch was found between guttural and voiceless stops ($p = .374$). The effect of speech rate was not obtained.

Table 23 Summary of fixed effects in a model examining f0, F1, and F2 after medial stops. Only significant interactions are reported.

Cue	Level	Estimate	SE	*t*-value	Pr(>|t|)	Sig.
A. f0	(Intercept)	238.9	8.8	27.08	< .001	***
	Voiceless	5.5	2.4	2.30	.030	*
B. F1	(Intercept)	677.3	20.6	32.91	< .001	***
	Voiceless	95.3	17.9	5.32	< .001	***
	Guttural	37.3	17.2	2.17	.041	*
	Slow	−24.7	8.9	−2.77	.018	*
C. F2	(Intercept)	1590.2	61.8	25.74	< .001	***
	Voiceless	−194.1	104.2	−1.86	.084	
	Slow	24.2	12.3	1.96	.066	.
	Slow: Guttural	−59.1	27.4	−2.16	.031	*

* = $p < .05$; *** = $p < .001$

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

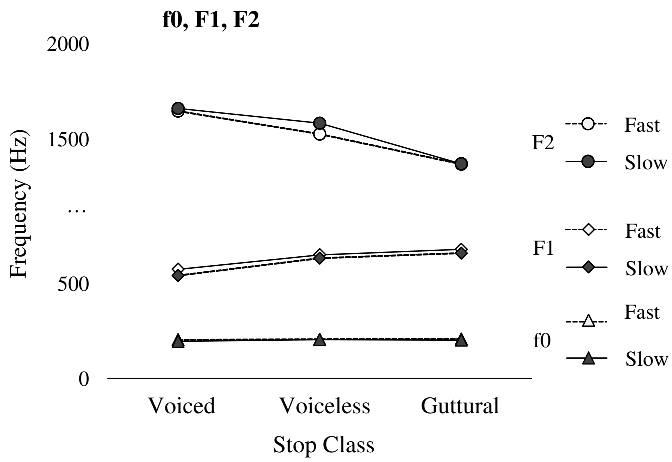

Figure 14 Effects of stop class and speech rate on spectral cues (f0, F1, F2) in medial stops.

For F1 (Table 23B), the effect of stop class was also significant. F1 was 95 Hz higher after voiceless stops than after voiced stops, and 37 Hz higher after guttural stops than after voiceless stops. The effect of speech rate was not obtained.

For F2 (Table 23C), the effect of stop class was marginally significant. F2 was 194 Hz lower after voiceless stops than after voiced stops. A 142 Hz decrease in F2 after guttural stops was not significant ($p = .249$). The effect of speech rate was only marginally significant, with a slight increase by 24 Hz in slow speech. The negative interaction coefficient indicated that the increase in slow speech was cancelled in guttural stops.

3.3.4 Summary

The analysis of medial stops revealed the same relationship between VOT and stop class as in word initial stops. Phonologically voiced stops /b d g/ were produced with phonetic voicing. They were articulated with voiced closure such as voicing started before the release, and the glottal state for these stops was consistent with phonation. F0 and F1 were lower after voiced stops indicating laryngeal adjustments to facilitate vibration of the vocal folds. Also, F2 was higher after medial voiced stops suggesting this might be an important cue to voicing in this dialect. But this effect was smaller in medial stops than in word-initial stops as the former did not reveal changes in F2 in response to rate manipulation.

Voiceless /p t k/ were produced with phonetic voicelessness. They long-lag positive VOT and had essentially voiceless closure with a short voicing tail from the preceding vowel. Guttural /tˤ q/ were also produced with phonetic voicelessness. They had short-lag positive VOT and voiceless closure. Both stop classes were articulated with a glottal state consistent with voicelessness: they had higher f0 and F1. F2 was predictably lower after guttural stops indicating the effect of tongue retraction due to pharyngealization.

In line with predictions of LR, VOT in voiceless stops and closure voicing in voiced stops were sensitive to speech rate manipulation in medial position. In slow speech, speakers demonstrated strong tendency to increase both duration of aspiration and duration of closure voicing to maintain it throughout the entire closure.

4 Discussion and conclusion

In this paper, we set to examine stop laryngeal contrast in KhA within the framework of LR based on two types of evidence: the phonetic realization of the sounds and diagnostics of

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

speakers' 'control', namely the effect of speech rate manipulation on VOTs and the degree of intervocalic voicing. We observed a complex pattern, in which two VOT categories – voicing lead and long lag – were mapped on the phonological contrast in VOICING, and the short lag category was consistently associated with PHARYNGEALIZATION. This pattern is very similar to the pattern reported in Kulikov (2022) for Qatari Arabic.

Word-initially, 93% of the phonologically voiced stops /b d g/ were produced with voicing lead. This ratio is very close to Swedish (Beckman et al. 2011) and Russian (Ringen & Kulikov 2012), in which 100% and 97% of the voiced stops were PREVOICED, respectively. The analysis of spectral cues showed that the glottal state was consistent with voicing: both f0 and F1 were lower and F2 was higher suggesting that speakers employed articulatory gestures to expand the supraglottal cavity in order to facilitate vibration of the vocal folds. Phonologically voiceless stops /p t k/ were produced as VOICELESS ASPIRATED. They had long lag positive VOT and higher f0 and F1. The guttural stops /tˤ q/ were produced as VOICELESS UNASPIRATED plosives. They had short-lag positive VOT and a glottal state consistent with phonetic voicelessness (i.e. higher f0 and F1). In addition, F2 was predictably lower in guttural stops as a result of tongue retraction (Ghazeli 1977).

Word-medially, we evaluated VOT for both plain and guttural voiceless stops and the degree of intervocalic voicing for all stop consonants. Voiceless /p t k/ and /tˤ q/ showed the same VOT patterns as in the word-initial position: voiceless stops /p t k/ had long lag VOT, but VOT was short lag in guttural /tˤ q/. Voiced stops were predominantly prevoiced, as 94% of the cases had a fully voiced closure. On the other hand, voiceless and guttural stops largely blocked the spread of voicing in the closure. They showed a small voicing tail of around 12 ms that continued from the previous sonorant segment. We evaluated the glottal state of the stops by measuring f0 and F1. Similar to the initial position, voiceless stops had higher f0 and F1 compared to the voiced stops.

The results of the current study largely support predictions of LR in terms of speakers' control of duration of temporal cues. In word-initial position, the three types of VOT responded differently to speech rate manipulation. Duration of voicing lead in voiced stops and long lag in voiceless stops gradually increased as speech became slower replicating the patterns previously reported for languages with prevoicing and/or aspiration (e.g. Kessinger & Blumstein 1997, Allen & Miller 1999, Magloire & Green 1999, Beckman et al. 2011, Kulikov 2020, among many). But short-lag VOT in guttural stops was not affected by rate and remained relatively stable across rate conditions. The fact that gutturals in KhA are consistently realized with short-lag VOT, which stays stable across rate manipulation, can be interpreted as a language specific requirement of this dialect, which links this VOT category with the contrast in pharyngealization.

In addition, we found that the pattern of response to speech rate manipulation in duration of closure voicing in word-medial voiced and voiceless stops mirrors the pattern of response in VOT of word-initial stops. As predicted by LR (e.g. Schwarz et al. 2019, for Nepali), long lag VOT in voiceless and closure voicing in voiced stops increased in slow speech while short lag in guttural stops and the short voicing tails of both phonetically voiceless categories did not change in response to rate manipulation.

Finally, our results demonstrate that the three classes of stops in KhA differed in spectral properties. Although differences in f0, F1, and F2 between voiced and voiceless stops have been previously reported in the literature and are expected, some findings were surprising. Quite predicably, f0 was consistently lower after voiced stops, which mirrors the cross-linguistic pattern (e.g. Westbury 1983, Lisker 1986, Kingston & Diehl 1994, among others). But unlike studies that report no or little difference in f0 between voiceless unaspirated and voiceless aspirated stops across languages (e.g. Kirby & Ladd 2018), our study demonstrates that long lag VOT in voiceless stops is aligned with higher f0 values than short lag VOT along the lines of Kirby (2018), who found slightly lower f0 in voiceless unaspirated stops compared to aspirated stops in Khmer, Thai and Vietnamese (languages with a three-way contrast). It is of note that this difference was observed only in word-initial position in both studies, being largely neutralized word- or phrase-medially. These findings probably suggest that higher f0 is typically associated with voicelessness in production, but it can be enhanced in prominent positions. The reason why differences in f0 between unaspirated and aspirated stops are not maintained in all tokens is not clear so far, but we believe it is probably due to the fact that it is not a primary cue to phonological voicing and, thus, it does not have to be obligatorily mapped on the phonological feature that specifies voiceless stops.

We also found consistent lowering of F1 in voiced stops in KhA, which is also a typical pattern across languages (e.g. Summerfield & Haggard 1977, Westbury 1983, Kingston & Diehl 1994). This is a typical aftermath of the expansion of the supraglottal cavity in order to create rarefaction and thus reduce supraglottal pressure to facilitate vibration of the vocal folds. F1 in voiceless stops, on the other hand, was considerably higher, but unlike languages like English, where F1 would be expected to be higher in voiceless aspirated stops than in voiceless unaspirated stops as a result of F1 cutback in the former (Stevens & Klatt 1974), voiceless aspirated stops in KhA were produced with lower F1 than unaspirated guttural stops. Higher F1 after guttural stops in Arabic, however, is an expected result of tongue root retraction (Ghazeli 1977). We argue that this finding provides evidence for non-laryngeal mapping of F1 in KhA gutturals.

Next, we found that F2 is a more important cue to voicing in KhA than in other languages reported in the literature. Whereas previous studies (e.g. Bolla 1981, Westbury 1983, Ahn 2018) demonstrated that F2 can be raised in production of voiced stops due to advancement of the tongue root – another strategy to expand the supraglottal cavity and facilitate glottal pulsing – the results of our study show that speakers use this strategy quite consistently. It is of interest, that F2 is also an important cue to distinguish between voiceless and guttural stops in KhA. F2 is lowered in production of Arabic gutturals because of a specific articulatory gesture of tongue root retraction into the pharyngeal area (Ghazeli 1977). Quite predictably, vowels following both guttural stops in our study had lower F2, indicating back articulation. In line with Kulikov (2022), we consider low F2 to be the main acoustic correlate for guttural stops in KhA.

Finally, the results of our study demonstrate that the spectral correlates of voicing and emphasis in KhA were also sensitive to manipulation of speech rate. But it is on note that the response patterns were different for f0/F1 and for F2. F0 and F1 were slightly lower in slow speech, but the adjustment of the larynx to the tempo affected realization of all stops in the same fashion. Changes in F2 in response to manipulation of speech rate, in contrast, were found only for voiced and guttural stops. Both increase in F2 for voiced and decrease in F2 for guttural stops were more prominent in slow speech than in fast speech, suggesting selective accommodation of the vocal tract according to phonological specifications of segments. While the greater drop in F2 in slow speech is, in fact, expected due to direct mapping of this cue to the feature [guttural], significant increase in F2 in word-initial voiced stops suggests that tongue advancement to ensure expansion of the pharyngeal cavity in order to effectively maintain glottal pulsing is also an important strategy in this dialect.

The mapping of the acoustic correlates of the three classes of KhA stops is summarized in Table 24. In line with predictions of LR, we argue that voiced and voiceless stops are specified with [voice] and [sg], respectively, while for guttural stops this laryngeal feature is unspecified. The unaspirated series in KhA, in contrast, is specified with the contrastive feature [guttural], which explains retraction of the tongue and subsequent backing of the neighboring vowel in Arabic (Watson 2002).

Crucially, although KhA uses three VOT categories – voicing lead, short lag, and long lag – the laryngeal contrast in this language is still between voiced and voiceless aspirated stops, similar to the contrast in Swedish (Beckman et al. 2011), Southern American English (Hunnicutt & Morris 2016), or Qatari Arabic (Kulikov 2020). The three distinct categories of VOT in KhA do not indicate a three-way contrast like in Eastern Armenian or Thai (Lisker & Abramson 1964). The unaspirated series of stops in these languages is another LARYNGEAL

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

Table 24 Mapping of acoustic correlates and phonological features in KhA stops.

Acoustic correlate	Stop class: Feature:	/b d g/ [voice]	/p t k/ [sg]	/$t^ʕ$ q/ [guttural]
VOT		voicing lead	long lag	short lag
f0		low	high	high
F1		low	high	very high
F2		high		very low

category. In contrast, unaspirated stops in KhA differ from voiced and voiceless stops in the NON-LARYNGEAL phonological feature [guttural].

To conclude, laryngeal systems contrasting voicing lead and long lag may not be consistent with the principle of 'economical representation' in phonology (Chomsky & Halle 1968), but they appear to be common in world's languages. In addition, VOT can signal not only a contrast in voicing but other phonological contrasts as well. Future studies should reveal whether this mapping in a language is stable or it indicates a change in progress.

Acknowledgements

We gratefully thank the editors of the *Journal of International Phonetic Association* and the anonymous reviewers for their insightful comments and suggestions that helped in improving this paper. We would also like to show our gratitude to all the participants for their assistance with data collection during the COVID-19 pandemic.

Appendix. Word list

Table A1 Initial position.

Place	Word	Orthography	Gloss
Bilabial	biːʕ	بيع	sell (IMP, 2MSG)
	biːd	بيد	in hand
	piːʧ	پيچ	screw
	piːp	پيپ	pipe
	baːb	باب	door
	baːt	بات	(he) slept over
	pɑːrk	پارک	park
	pɑːs	پاس	pass
	buːza	بوزه	muzzle
	buːg	بوگ	steal (IMP, 2MSG)
	puːk	پوک	empty
	puːʃa	پوشه	file
	bard	برد	cold (N)
	badr	بدر	full moon
	patu	پتو	blanket
	pad	پد	pad

Table A1 Continued.

Place	Word	Orthography	Gloss
Dental	diːn	دین	religion
	diːʧ	دیچ	rooster
	tiːm	تیم	team
	tiːn	تین	fig
	daːɾ	دار	room
	daːχ	داخ	(he) got confused
	taːb	تاب	(he) repented
	taːʤ	تاج	crown
	duːɾ	دور	rooms
	duːʃ	دوش	shower
	tuːb	توب	repent (IMP, 2MSG)
	tuːt	توت	berry
	dazz	دز	(he) sent
	daff	دف	a type of musical instrument
	tall	تل	hill
	tag	تگ	tag
Velar	giːra	گیره	pin
	giːg	گیگ	gigabyte
	kiːʃ	کیش	check (in chess)
	kiːst	کیست	cyst
	gaːl	گال	(he) said
	gaːʕ	گاع	land
	kaːfi	کافی	enough
	kaːʃi	کاشی	tile
	guːm	گوم	stand up (IMP, 2MSG)
	guːl	گول	say (IMP, 2MSG)
	kuːb	کوب	cup (for drinking)
	kuːχ	کوخ	cottage
	gaʧ	گچ	chalk
	gadd	گد	measurement
	kaʃk	کشک	a type of diary product
	kaf	کف	foam
Emphatic	tˤiːn	طین	mud
	tˤiːb	طیب	get better (IMP, 2MSG)
	tˤaːb	طاب	(he) got better
	tˤaːɾ	طار	(he) flied
	tˤuːl	طول	height
	tˤuːs	طوس	bowls

https://doi.org/10.1017/S002510032300004X Published online by Cambridge University Press

Table A1 Continued.

Place	Word	Orthography	Gloss
	tˤall	طل	(he) peeped
	tˤalab	طلب	request
Uvular	qiːɾ	قير	tar
	qiːf	قيف	funnel
	qaːb	غاب	(he) got absent
	qaːli	غالی	expensive
	quːl	غول	monster
	quːri	قوری	teapot
	qaʃ	غش	faint
	qand	قند	sugar cube

Table A2 Medial position.

Place	Word	Orthography	Gloss
Bilabial	χabaɾ	خبر	a piece of news
	ʔabad	ابد	never
	kapaɾ	كپر	a traditional house
	sepaɾ	سپر	shield
Dental	ʔadab	ادب	politeness
	madad	مدد	help
	ħatam	حتم	(it) necessitated
	χaːtam	خاتم	the last of
Velar	ɾagasˤ	رگص	(he) danced
	baːga	باگه	bouquet
	ħakam	حكم	judge
	ðakaɾ	ذكر	male
Guttural	batˤal	بطل	hero
	matˤaɾ	مطر	rain
	ʔaqal	اقل	less
	maqaɾ	مقر	military station

References

Abbaasian, Fatemeh & Mandana Nourbakhsh. 2015. Ilam Kurdish dialect type determined by measuring the VOT. *Journal of Western Iranian Languages and Dialects* 2(8), 79–99.

Ahmed, Zhwan O. 2019. *The application of English theories to Sorani phonology*. Ph.D. dissertation, Durham University.

Ahn, Suzy. 2018. The role of tongue position in laryngeal contrasts: An ultrasound study of English and Brazilian Portuguese. *Journal of Phonetics* 71, 451–467.

Al Malwi, Ibrahim M. 2017. The effects of gender and age on voice onset time by Abha Arabic speakers. Master's thesis, California State University, Long Beach.

AL-Gamdi, Nief, Jalal Al-Tamimi & Ghada Khattab. 2019. The acoustic properties of laryngeal contrast in Najdi Arabic initial stops. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences* (ICPhS XIX), 2051–2055. Australasian Speech Science and Technology Association Inc.

Allen, J. Sean & Joanne L. Miller. 1999. Effects of syllable-initial voicing and speaking rate on the temporal characteristics of monosyllabic words. *The Journal of the Acoustical Society of America* 106, 2031–2039.

Alzoubi, Abdulaziz. 2016. A sociophonetic study of the effect of gender on emphatic–plain contrast in Jordanian Arabic. *The Journal of the Acoustical Society of America* 140, 3113–3113.

Amirian, Beaina. 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]. Master's thesis, Allameh Tabataba'i University.

Bahrani, Nawal. 2022. Word-initial voicing contrast in Khuzestani Arabic stops. *Journal of Language Related Research* 13, 477–510.

Bahrani, Nawal & Golnaz Modarresi Ghavami. 2021. Khuzestani Arabic. *Journal of the International Phonetic Association* 51, 299–313.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68, 255–278.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software* 67, 1–48.

Beckman, Jill, Pétur Helgason, Bob McMurray & Catherine Ringen. 2011. Rate effects on Swedish VOT: Evidence for phonological overspecification. *Journal of Phonetics* 39, 39–49.

Beckman, Jill, Michael Jessen & Catherine Ringen. 2013. Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics* 49, 259–284.

Benkí, José R. 2001. Place of articulation and first formant transition pattern both affect perception of voicing in English. *Journal of Phonetics* 29, 1–22.

Bijankhan, Mahmood & Mandana Nourbakhsh. 2009. Voice onset time in Persian initial and intervocalic stop production. *Journal of the International Phonetic Association* 39, 335–364.

Boersma, Paul & David Weenink. 2021. Praat: Doing phonetics by computer (Version 6.1.50). https://www.praat.org/.

Bolla, Kalman. 1981. *A conspectus of Russian speech sounds*. Cologne: Bölau.

Chen, Mathew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22, 129–159.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper.

Davidson, Lisa. 2016. Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics* 54, 35–50.

Davis, Matthew J. 2010. Contrast coding in multiple regression analysis: Strengths, weaknesses, and utility of popular coding structures. *Journal of Data Science* 8, 61–73.

Dmitrieva, Olga, Fermando Liano, Amanda A. Shultz & Alexander L. Francis. 2015. Phonological status, not voice onset time, determines the acoustic realization of onset f0 as a secondary voicing cue in Spanish and English. *Journal of Phonetics* 49, 77–95.

Docherty, Gerard J. 1992. *The timing of voicing in British English obstruents*. Berlin: Foris.

Ghazeli, Salem. 1977. *Back consonants and backing coarticulation in Arabic*. Ph.D. dissertation, The University of Texas at Austin.

Haggard, Mark, Stephen Ambler & Mo Callow. 1970. Pitch as a voicing cue. *The Journal of the Acoustical Society of America* 47(2), 613–617.

Halle, Morris & Kenneth N. Stevens. 1971. A note on laryngeal features. *Quarterly Progress Report, Research Laboratory of Electronics, MIT* 101, 198–211.

Helgason, Pétur & Catherine Ringen. 2008. Voicing and aspiration in Swedish stops. *Journal of Phonetics* 36, 607–628.

Hombert, Jean-Marie, John J. Ohala & William G. Ewan. 1979. Phonetic explanations for the development of tones. *Language* 55, 37–58.

Honeybone, Patrick. 2005. Diachronic evidence in segmental phonology: The case of obstruent laryngeal specifications. In Marc van Oostendorp & Jeroen van de Weijer (eds.), *The internal organization of phonological segments*, 319–354. Berlin: Mouton de Gruyter.

House, Arthur S. & Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America* 25, 105–113.

Hunnicutt, Leigh & Paul A. Morris. 2016. Prevoicing and aspiration in Southern American English. *University of Pennsylvania Working Papers in Linguistics* 22(1), Article 24.
Hussain, Qandeel. 2018. A typological study of voice onset time (VOT) in Indo-Iranian languages. *Journal of Phonetics* 71, 284–305.
Iverson, Gregory K. & Joseph C. Salmons. 1995. Aspiration and laryngeal representation in Germanic. *Phonology* 12, 369–396.
Jansen, Wouter. 2004. *Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch.* Ph.D. dissertation, University of Groningen.
Jessen, Michael & Catherine Ringen. 2002. Laryngeal feature in German. *Phonology* 19, 189–218.
Keating, Patricia A. 1984. Phonetic and phonological representation of stop consonant voicing. *Language* 60, 286–319.
Kessinger Rachel H. & Sheila E. Blumstein. 1997. Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics* 25, 143–168.
Khattab, Ghada, Feda Al-Tamimi & Barry Heselwood. 2006. Acoustic and auditory differences in the /t/–/ṭ/ opposition in male and female speakers of Jordanian Arabic. In Sami Boudelaa (ed.), *Perspectives on Arabic linguistics XVI*, 131–160. Amsterdam & Philadelphia, PA: John Benjamins.
Kingston, John & Randy L. Diehl. 1994. Phonetic knowledge. *Language* 70, 419–454.
Kingston, John, Neil A. Macmillan, Laura Walsh Dickey, Rachel Thorburn & Christine Bartels. 1997. Integrality in the perception of tongue root position and voice quality in vowels. *The Journal of the Acoustical Society of America* 101, 1696–1709.
Kirby, James. 2018. Onset pitch perturbations and the cross-linguistic implementation of voicing: Evidence from tonal and non-tonal languages. *Journal of Phonetics* 71, 326–354.
Kirby, James & D. Robert Ladd. 2018. Effects of obstruent voicing on vowel f0: Implications for laryngeal realism. *Yearbook of the Poznań Linguistic Meeting* 4, 213–235.
Kulikov, Vladimir. 2020. Laryngeal contrast in Qatari Arabic: Effect of speaking rate on VOT. *Phonetica* 77, 163–185.
Kulikov, Vladimir. 2022. Voice and emphasis in Arabic coronal stops: Evidence for phonological compensation. *Language and Speech* 65(1), 73–104.
Kulikov, Vladimir, Fatemeh M. Mohsenzadeh & Rawand M. Syam. Effects of emphasis spread on VOT in coronal stops in Qatari Arabic. *Journal of the International Phonetic Association*, doi:10.1017/S0025100321000256. Published online by Cambridge University Press, 2 November 2021.
Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. *lmerTest* package: Tests in linear mixed effects models. *Journal of Statistical Software* 82, 1–26.
Ladefoged, Peter. 1973. Features of the larynx. *Journal of Phonetics* 1, 73–83.
Lee, Seung Suk & Kelly Harper Berkson. Speech rate effects on VOT in a three-category language: Evidence from Hakha Chin. *Indiana Working Papers in South Asian Languages and Cultures* 1(1), doi:10.14434/iwpsalc2019.v1i1.27453. Published online by Indiana University, 13 July 2019.
Leitner, Bettina & Nawal Bahrani. Forthcoming. Khuzestani Arabic. In Charles George Häberl (ed.), *Language diversity in Iran: New texts and perspectives from non-Iranian languages* (The Companion of Iranian Languages and Linguistics (CILL)).
Liberman, Alvin M., Pierre C. Delattre & Franklin S. Cooper. 1958. Some cues for the distinction between voiced and voiceless stops in initial position. *Language and Speech* 1, 153–167.
Lisker, Leigh. 1986. "Voicing" in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29, 3–11.
Lisker, Leigh & Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20(3), 384–422.
Löfqvist, Andres, Thomas Baer, Nancy S. McGarr & Robin Seider Story. 1989. The cricothyroid muscle in voicing control. *The Journal of the Acoustical Society of America* 85, 1314–1321.
Magloire, Joël & Kerry P. Green. 1999. A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish. *Phonetica* 56 158–185.
Matuschek, Hannes, Reinhold Kliegl, Shravan Vasishth, Harald Baayen & Douglas Bates. 2017. Balancing Type I error and power in linear mixed models. *Journal of Memory and Language* 94, 305–315.

McCarthy, John J. 1994. The phonetics and phonology of Semitic pharyngeal. In Patricia A. Keating (ed.), *Phonological structure and phonetic form: Papers in Laboratory Phonetics III*, 191–233. Cambridge: Cambridge University Press.
Morris, Paul A. 2018. Rate effects on Southern American English VOT. *Proceedings of the Linguistic Society of America* 3(60), 1–10.
Ohde, Ralph N. 1984. Fundamental frequency as an acoustic correlate of stop consonant voicing. *The Journal of the Acoustical Society of America* 75(1), 224–230.
Öğüt, Fatih, Mehmet Akif Kiliç, Erkan Zeki Engin & Raşit Midilli. 2006. Voice onset time for Turkish stop consonants. *Speech Communication* 48(9), 1094–1099.
Olson, D. Eve & Rachel Hayes-Harb. 2019. Voice onset time in Arabic and English stop consonants. *Al-'Arabiyya* 52, 29–48.
Peterson, Gordon E. & Ilse Lehiste. 1962. Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America* 32, 693–703.
Pind, Jörgen. 1995. Speaking rate, voice-onset time, and quality: The search for higher-order invariants for two Icelandic speech cues. *Attention, Perception & Psychophysics* 57, 291–304.
R Core Team. 2021. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.
Rahim, Abbas Jawdat & Ziyad Rakan Kasim. 2009. A spectrographic study of voice onset time in Arabic. *Journal of Education and Science* 16, 28–41.
Rifaat, Khaled. 2003. Voice onset time in Egyptian Arabic: A case where phonological categories dominate. *Proceedings of the 15th International Congress of Phonetic Science* (ICPhS XV), 791–794.
Ringen, Catherine & Vladimir Kulikov. 2012. Voicing in Russian stops: Cross-linguistic implications. *Journal of Slavic Linguistics* 20(2), 269–286.
Ringen, Catherine & Kari Suomi. 2012. The voicing contrast in Fenno-Swedish stops. *Journal of Phonetics* 40(3), 419–429.
Schwarz, Martha, Morgan Sonderegger & Healther Goad. 2019. Realization and representation of Nepali laryngeal contrasts: Voiced aspirated and laryngeal realism. *Journal of Phonetics* 73, 113–127.
Seyfarth, Scott & Marc Garellek. 2018. Plosive voicing acoustics and voice quality in Yerevan Armenian. *Journal of Phonetics* 71, 425–450.
Slis, Iman H. 1986. Assimilation of voice in Dutch as a function of stress, word boundaries, and sex of speaker and listener. *Journal of Phonetics* 14, 311–326.
Solé, Maria-Josep. 1992. Phonetic and phonological processes: The case of nasalization. *Language and Speech* 35(1–2), 29–43.
Stevens, Kenneth. 1998. *Acoustic phonetics*. Cambridge, MA: MIT Press.
Stevens, Kenneth & Dennis Klatt. 1974. Role of formant transitions in the voiced–voiceless distinction for stops. *Journal of Acoustic Society of America* 55, 653–659.
Summerfield, Quentin & Mark Haggard. 1977. On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *The Journal of the Acoustical Society of America* 62, 435–448.
Tamim, Nour. 2017. Voicing contrast of stops in the Palestinian Arabic dialect. Master's thesis, University Van Amsterdam.
Tranel, Bernard. 1998. *The sounds of French: An introduction*. Cambridge: Cambridge University Press.
van Alphen, Petra & Roel Smits. 2004. Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics* 32, 455–491.
Vaux, Bert & Bridget Samuels. 2005. Laryngeal markedness and aspiration. *Phonology* 22, 395–436.
Watson, Janet C. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.
Westbury, John R. 1983. Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *The Journal of the Acoustical Society of America* 73, 1322–1336.
Westbury, John R. & Patricia A. Keating. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22, 145–166.
Yeni-Komshian, Grace H., Alfonso Caramazza & Malcolm S. Preston. 1977. A study of voicing in Lebanese Arabic. *Journal of Phonetics* 5, 35–48.
Zirak, Mahjoob. 2014. *A case study of the Khorasani variety of Kurmanji language*. Ph.D. dissertation, Hiroshima University.