Voicing Contrast in Najdi Arabic Stops: Implications for Laryngeal Realism

Nief Aied Al-Gamdi

Thesis Submitted for the Degree of Doctor of Philosophy (Integrated)

School of Education, Communication and Language Sciences.

Newcastle University

August 2022

Abstract

The present study investigates the phonetic and phonological aspects of the voicing contrast in stops in Najdi Arabic, a dialect that has been found to contrast prevoiced and aspirated stops. This study discusses the implications of the acoustic correlates of Voiceless and Voiced stops for the phonological representation of the voicing contrast in this variety and examines the connection between the acoustic signal and the distinctive features that specify the opposition by employing the types of evidence proposed in the realm of laryngeal realism. These types of evidence include the manifestation of acoustic correlates of stops in various positions, speech rate effect on aspiration and prevoicing, and the Voiceless and Voiced stops' behaviour in stop-stop clusters across word boundary in terms of regressive voicing assimilation.

The manifestation of the acoustic correlates of Voiceless and Voiced stops shows that Voiceless stops are aspirated in the examined positions whereas Voiced stops show robust prevoicing in utterance-initial and utterance-medial contexts. The acoustic correlates also show that Voiceless stops are robustly accompanied by longer closure, longer burst, higher F0 and F1 onset, and lower burst intensity. Voiced stops, on the other hand, are robustly accompanied by shorter closure (utterance-medially), shorter burst, lower F0 and F1, and higher burst intensity. ***Speech rate*** affects both aspiration and prevoicing in Voiceless and Voiced stops, respectively. Prevoicing and aspiration are lengthened in normal speech rate in comparison to fast speech rate. ***Stop-stop cluster*** results show that both Voiceless and Voiced stops trigger some (de)voicing in the preceding member of the cluster. The acoustic analysis reveals that Voiceless stops show voicing assimilation in F0/F1 and burst intensity but not in voicing in the closure. For Voiced stops, the results show a degree of devoicing in their closure but not in F0/F1 and burst intensity.

The results suggest that Voiceless and Voiced stops in Najdi Arabic have features from both aspirating and voicing languages. This claim is supported by the three types of evidence implemented in this study. The assumption that both Voiceless and Voiced stops are specified implicates that the voicing contrast in Najdi Arabic is overspecified in the phonology with two features, [spread glottis] and [voice]. Applying the numeric values of phonetic distinctive features proposed by Beckman et al. (2013), on the scale of 1 to 9, the present study claims that Voiced stops in Najdi Arabic are specified with [9 voice] while Voiceless stops are specified with [8 spread glottis], mainly because of the existence of moderate aspiration in utterance-initial Voiceless stops and the robust prevoicing found in utterance-initial and utterance-medial Voiced stops (1 means inactive, 9 means highly active). The phonological repercussions for the proposed overspecification in the voicing contrast in

Najdi Arabic are discussed with a specific focus on the inclusion of such a patterning in theoretical models of voicing.

Acknowledgements

First, I would like to express my gratitude to Allah for giving me strength and hope. Praise be to Allah.

I would like to thank my supervisors, Jalal Al-Tamimi and Ghada Khattab, for their enormous support and guidance throughout this amazing journey. They provided me with the tools needed to help me gain success and fulfil the requirements of the present study to pursue my academic goals. I also thank Daniel Duncan and Gary Taylor-Raebel for their help and advice during the annual progressions.

I thank the staff at King Saud University for their help in recruiting and recording the participants of the present study. I would like to express a special thank you to Samia Al-Gamdi, who helped me record the female participants at King Saud University.

I also would like to offer my gratitude to my family. Words fail to express what my parents, my wife, my kids, my brothers, and my sisters have sacrificed for me. They helped me overcome all the obstacles I faced during my studies in the UK. I especially thank my wife, Aziza, and my kids, Lojain, Maysun, and Basil. You were there for me all the time.

To my friends in the UK, thank you so much for your help and support during the difficult times.

Finally, to my friends in Saudi Arabia, you believed in me from day one and gave me strength with your encouragement and continuous support, so I thank you from the bottom of my heart.

Table of Content

List of figures

List of tables

Chapter 1. Introduction

Phonetic characteristics of sounds are expected to be informative in terms of understanding the phonological aspects of languages. The gradual consensus among phoneticians and phonologists about the connection between phonetics and phonology has resulted in a number of empirical studies attempting to characterise the nature of this interaction. One of the most addressed topics in this field of study is the phonetic and phonological aspects of the voicing contrast in stops among languages. This topic has been used as an effective tool for constructing a theoretical justification for the parallelism between phonological entities and phonetic details. This dissertation attempts to analyse the phonetic and phonological aspects of voicing contrast in stops in Najdi Arabic which show the rare phenomenon of contrasting prevoiced and aspirated stops in word-initial position. Specifically, this study investigates the acoustic properties of Najdi Arabic Voiceless and Voiced stops (Voiceless and Voiced with capital letters refer to phonological voicing). The investigation is carried out with respect to different positions within the word and the utterance to determine how these acoustic details are implemented across phonetic contexts and to establish which of these correlates robustly differentiate Voiceless and Voiced stops. Moreover, additional processes that involve the voicing contrast and the implications thereof for the phonological representations are investigated, including passive and active voicing, final devoicing, speech-rate effects, and voicing assimilation.

Voicing contrast has been the focus of numerous studies in the field of phonetics and phonology. A plethora of experimental studies have recently provided precise and in-depth analysis of the phonetic realisation of voicing contrast and have used these details as tools to reveal how voicing is represented phonologically, and to emphasise the robustness of the ties between phonetics and phonology. The nature of the phonetic details and the phonological representations of this contrast have led to several models that interpret the interaction between the acoustic cues to voicing contrast and the distinctive features that specify this opposition in different languages (Kohler, 1984; Keating, 1984; Kingston and Diehl, 1995; Jessen, 1998, 2001; Harris, 1994; Iverson and Salmons, 1995; Jessen and Ringen, 2002; Honeybone, 2005; Beckman et al., 2013). Although these models might differ in terms of the nature and choice of the phonological features that specify the contrast, the primary arguments of these models are based on the proposed classification of languages into two categories: *aspirating* languages and *voicing* languages (see section 2.1 for more details). Based on the notion of VOT proposed by Lisker and Abramson (1964), these models demonstrate that, in initial position, aspirating languages such as English and German contrast aspirated stops (positive VOT: long lag) with unaspirated stops (short lag), while voicing languages such as Russian and French, on the other hand, contrast prevoiced stops (negative VOT: voicing lead) with unaspirated stops (short lag).

Various studies gradually started to pick up the thread and looked at stops in voicing and aspirating languages to disambiguate the acoustic differences between the two laryngeal systems in different phonetic contexts. It turned out that the differentiation between voicing and aspirating languages is not exclusive to initial stops, but it appeared in intervocalic or intersonorant stops as well. It has been found that Voiced intervocalic or intersonorant stops in aspirating languages such as German are not always voiced, and in case the voicing is present it is caused by the surrounding context (passive voicing) which can be acoustically identified as a weak broken voice bar during the closure course (Jessen, 2001; Jansen, 2004, 2007; Beckman et al., 2013; Al-Tamimi and Khattab, 2018). Voiced intervocalic or intersonorant stops in voicing languages such as Russian, on the other hand, are produced with active voicing caused by an intentional articulatory movement which can be acoustically identified as a strong visible voice bar covering the whole closure period (active voicing) (Ringen and Kulikov, 2012). Stops in utterance-final position, however, posit a challenge for the voicing/aspirating typology in that stops in such a context are expected to be neutralised in both voicing and aspirating languages (Jessen, 1998; Jansen, 2004). This neutralisation has been found to be incomplete suggesting that there are detectable acoustic traces in the neutralised stops in various languages that partially preserve the voicing distinction (Charles-Luce, 1985; Greisbach, 2001; Piroth and Janker, 2004). Therefore, it could be assumed that the neutralised stops can be distinguished in terms of the voicing/aspirating languages classification based on the possibility of these preserved acoustic traces to be reflective of their phonological representation (Iverson and Salmons, 2011).

The crucial discussions arising from the aforementioned acoustic characterisation of voicing contrast across positions are related to two broad issues: 1) how such an opposition is specified in the representational systems of voicing and aspirating languages, and 2) how the mapping between the phonological and phonetic representation is characterised. In relation to the first issue, the *binarity* vs *privativity* distinction forms the basis of the discussion within phonological theory with regard to voicing contrast representation. That is, are the phonological features specifying the voicing contrast defined as binary with positive and negative values ([+/- voice] in case of voicing languages, [+/- spread glottis] in case of aspirating languages) or defined by the privative presence and absence of terms ([voice] [Ø] for voicing languages, [spread glottis] [Ø] for aspirating languages)? What are the phonological ramifications for each view? This leads us to the second issue in relation to the connection between phonetics and phonology which rests on a lengthy discussion throughout the history of phonological theory. The modelling of the phonetics-phonology interaction ranged from proposing separation between the two domains with no interaction (Chomsky and Halle, 1968), separation with an interface that transforms the categorical entities to their gradient continuous characteristics (Keating, 1984; Pierrehumbert and Beckman, 1988; Cohn, 1990), and integration into one domain (Kohler, 1984; Ohala, 1990). Jessen (1998) postulated that the difference between the interface models and the integration models could be reduced by taking a middle approach that emphasises the two requirements crucial to both views: 1) the importance of the lexical representation to be made up of categorical forms (features), 2) the importance of the influence of phonetic details on lexical representation.

One of the approaches involving the strong interactions between the acoustic details and the phonological features is known as laryngeal realism (this term was first introduced by Honeybone, 2002; see details in section 2.2) (Iverson and Salmons, 1995; Jessen, 1998; Honeybone, 2002, 2005; Harris, 1994; Iverson and Salmons, 1995; Jessen and Ringen, 2002; Beckman et al., 2013), which the present dissertation adopts here. According to laryngeal realism, phonological specification can be evaluated through 1) examining the phonetic realisation of the obstruent across contexts, and 2) investigating the obstruent's behaviour in phonological processes (Iverson and Salmons, 2003; Jansen, 2004; Schwartz et al., 2019). Therefore, voicing languages like Russian contrast specified prevoiced stops [voice] with unspecified unaspirated stops [Ø] whereby aspirating languages like German contrast specified aspirated stops [spread glottis] with unspecified unaspirated stops [Ø] (Iverson and Salmons, 1995; Honeybone, 2005; Iverson and Salmons, 2011). Different types of evidence have been employed in laryngeal realism to investigate these assumptions. It has been proposed that varying the speech rate affects specified stops (slow speech rate increases the duration of aspiration/prevoicing) but not unspecified ones (Miller et al., 1986; Volaitis and Miller, 1992; Kessinger and Blumstein, 1997; Nagao and de Jong, 2007; Beckman et al., 2011; Magloire and Green, 1999; Solé and Estebas, 2000). Another common type of evidence proposed in laryngeal realism is related to regressive voicing assimilation. That is, specified stops trigger voicing/devoicing to the preceding sound while unspecified stops do not (Burton and Robblee, 1997; Barry and Teifour, 1999; Jansen, 2004; Kulikov 2012). (see section 2.2.3)

Some phonetic manifestations of voicing contrast among languages posit a challenge for laryngeal realism. Arabic, the target language of the present study, has been generally described in the literature as a voicing language (Yeni- Komshian et al., 1977; Khattab, 2002). However, instrumental studies of the voicing contrast in modern Arabic dialects have revealed considerable variation in the acoustic properties of Voiceless and Voiced stops

(Lebanese Arabic: Yeni- Komshian et al., 1977; Al-Tamimi and Khattab, 2018; Ghamidi dialect: Alghamdi, 1990; Kuwaiti Arabic: Mabrouk, 1981; Najdi Arabic: Flege and Port, 1981, Al-Gamdi et al., 2019; Jordanian Arabic: Mitleb, 2001; Cairene Arabic: Kabrah, 2008; Qatari Arabic: Kulikov, 2019, 2020). It has been proposed that Najdi and Qatari dialects, unlike others, contrast aspirated and prevoiced stops (Najdi Arabic: Flege and Port, 1981; Al-Gamdi et al., 2019; Qatari Arabic: Kulikov, 2019, 2020). Such variations raise crucial questions about the phonetic and phonological aspects specifying this contrast in Najdi and Qatari Arabic.

The present study aims to contribute to the field by investigating the phonetic realisation of voicing in stops in Najdi Arabic and its implications for the laryngeal features, using the diagnostics proposed in the literature of *laryngeal realism*. The outcomes of the present study will provide insight into the voicing contrast in Najdi Arabic and how interactions between phonetic realisation and laryngeal features enhance our understanding of the link between phonetics and phonology.

The next part of this chapter highlights fundamental topics that form the basic foundation for the current analysis. This includes the distinctive features that have been proposed to specify the voicing contrast, the aspects of phonological specification, and the phonetics-phonology interactions.

1.1 Distinctive features

The concept of distinctive features was first proposed in the early work of Jakobson (1939). Building on the notion of the *phoneme*, Jakobson emphasised the role of sound properties (sub-phoneme), rather than the *phoneme* to be employed as the fundamental element that signals the distinction between segments (distinctive function); this eventually led to the introduction of distinctive features (Jessen, 1998). That is, the distinction between /t/ and /d/ is not based on the concept of the phoneme as a whole; it is based on the voicing property which is part of each phoneme (sub-phoneme). Accordingly, the distinctive feature [voice] serves as holder for the distinctive function between the two English words *pad* and *pat*. Jakobson, Fant and Halle (1952) and Jakobson and Waugh (1987) continued the effort of conceptualising the distinctive features by identifying the articulatory, auditory, and acoustic aspects of distinctive features focusing on their role in the categorisation of sounds and their structure in the languages of the world. Jakobson and Waugh's work was considered the final version of the Jakobsonian approach in terms of determining the nature of the distinctive features as well as modelling the interaction between the distinctive features and phonetic reality (Jessen, 1998). Jessen (1998) summarised Jakobson and Waugh's view of the distinctive features as follows:

> "The distinctive feature concept as proposed by Jakobson and Waugh (1987) is characterized by a number of properties…: (1) distinctive features are binary; (2) they are defined in phonetically concrete terms; (3) they have universal validity; and (4) they provide a common, or nearly common, set across consonants and vowels" (Jessen, 1998, p. 10).

Several proposals aimed to identify a set of various distinctive features with a different way of characterisation. The *Sound Pattern of English* (SPE) by Chomsky and Halle (1968) was a landmark in the history of the field. In SPE, a set of distinctive features were proposed with a phonetic explanation identified only in articulatory terms. The Jakobsonian approach (the multiple studies conducted by Jakobson, Fant and Halle (1952) and Jakobson and Waugh (1979, 1987) will be referred to as Jakobson and colleagues) and SPE were probably the most crucial and influential studies demonstrating the role of distinctive features in phonetics and phonology (Jessen, 1998; Hall, 2001). The two approaches share the view that distinctive features are defined in, at least, some phonetic aspects. They also share the view that each feature is associated with a single phonetic cue (Ladefoged, 2006). The binarity of the distinctive features is another parallel between the two approaches. However, they differ in two major issues. First, SPE defined the distinctive features only in articulatory terms whereas the Jakobsonian approach defined them in articulatory, acoustic, and perceptual terms. The emphasis on the importance of acoustics in the Jakobsonian approach stems from the view that articulatory gestures enhance the saliency (perceptibility) of the sound which consequently highlight the essentiality of the products of the articulatory gestures; namely the acoustic signal. Second, SPE assumed a disassociation between the distinctive features and the phonetic reality whereas the Jakobsonian approach proposed a robust connection between the two, considering phonetics and phonology to be inseparable. The arguments posited in SPE and the Jakobsonian approach sparked a very long debate regarding the accepted set of universal features, the degree of abstractness in the distinctive features, and the robustness of the ties between the distinctive features and the continuous aspects of the acoustic signal. These topics will be discussed in the coming paragraphs in terms of voicing contrast which is the scope of investigation in the current study.

The two sides of the debate about proposals for feature theory have evidently been manifested in studies dealing with voicing contrast in stops. The enormous range of phonological and phonetic accounts of the voicing contrast in stops across languages have been employed as a window affording insights into the nature of the distinctive features and consequently the nature of interactions between distinctive features and acoustic details. Theoretical models of voicing contrast differ in relation to the distinctive features that specify the opposition. As pointed out in the introduction, some models assume the feature [voice] to account for voicing contrast in both aspirating and voicing languages (Keating, 1984; Kingston and Diehl, 1995; Wetzels and Mascaro, 2001). Other models argue that [voice] represents the voicing contrast in voicing languages while [spread glottis] accounts for this opposition in aspirating languages (Iverson and Salmons, 1995, 2003; Beckman et al., 2011; Beckman et al., 2013). Some researchers argue for the feature [tense] in aspirating languages following the Jakobsonian approach such as the work of Jessen (1998, 2001). The feature [fortis] was introduced by Kohler (1984) and implemented in various studies such as the study of voicing and gemination by Al-Tamimi and Khattab (2018). (see section 2.1 for detailed description of these models).

Among the features that represent the voicing contrast in recent work in laryngeal phonology, the features [voice] and [spread glottis] (see Jessen 1998 for more details with regard to the differences between [tense] and [spread glottis]), received the most attention but have been subject to controversy regarding their phonological specification and phonetic content among languages (Lombardi, 1991; Iverson and Salmon, 1995; Avery, 1997; Honeybone, 2005; Beckman et al., 2011). In this study, I argue for the effectiveness of [voice] and [spread glottis] to specify the voicing contrast in stops in Najdi Arabic based on the theoretical assumptions presented in the laryngeal realism approach. In that regard, I argue for a robust connection between the distinctive features and their acoustic correlates in the phonetic realisation of Voiceless and Voiced stops across various contexts and their behaviour in phonological processes. The current study adopts the monovalent structure (privative) as the representational system for the distinctive features for voicing contrast in Najdi Arabic. Proposing the presence of both [voice] and [spread glottis] in the system of Najdi Arabic indicates an *overspecification*. That is, both Voiceless and Voiced stops are associated with active phonological features. The issue of phonetic and phonological specification is addressed in the coming section.

1.2 Phonetic and phonological specification

Specification, in basic terms, refers to whether or not a segment is associated with a distinctive feature in the phonology. This leaves us with two possibilities: the segment is specified or unspecified (it is also called *underspecified* in some studies based on the traditional theory of underspecification proposed in generative phonology). As explained in the previous section, a distinctive feature is associated with the sound property (or properties) that hold the identity of the segment. Accordingly, if the segment is unspecified, it means that it lacks its distinctiveness. That is, an unspecified segment will be prone to coarticulation mechanisms which can be acoustically detectable in processes such as assimilation. The behaviour of unspecified segments has been the target of investigation for many phonologists and phoneticians because of the theoretical assumptions that can be proposed based on these diagnostics. Keating (1988b) depicted the aspects of specification (underspecification) and previewed various accounts for such a theory within different domains including generative phonology, articulation, and acoustics. Keating built her argument on two concepts: 1) variability of the phonetic details of unspecified segments (due to the surrounding context) and 2) phonetic transparency which refers to the degree to which the behaviour of an unspecified segment can reveal its phonological representation. She concluded that surface unspecification (phonetic level) can mirror the phonological specification. Cohn (1993) investigated specification by looking at the differences between English and French in terms of vowel nasalisation in nasal contexts. The motivation of his study was the fact that French, unlike English, has a lexical contrast between nasal and oral vowels. The results showed robust oral manifestation in French and nasalisation in English in the phonetic realisation. Cohen then concluded that the results indicated the presence of [-nasal] in French and the absence of this feature in English. That is, nasalised vowels in English is a product of coarticulation (unspecification).

Turning to the robust connection between distinctive features and phonetic details adopted in the present study, the notion of specification/unspecification in the privative system proposed in laryngeal realism has triggered a plethora of studies focusing on the laryngeal contrast in voicing and aspirating languages and considering the patterning of prevoicing and aspiration under the effect of positional and prosodic aspects within the word and the utterance. Insightful investigations have been carried out to explain the implications of the phonetic properties of laryngeal contrasts for phonological representations and how phonological specification is manifested in different phonetic contexts. In addition to the distinction in word-initial stops, various studies have shown that stops in non-initial position behave differently in aspirating and voicing languages, leading to a differentiation between active voicing that implies the presence of [voice] in the phonology, and passive voicing, that results from phonetic co-articulation (unspecification). The following paragraph deals with the articulatory and acoustic aspects of passive and active (de)voicing and combine them with the notion of specification.

One of the main assumptions about the distinction between (de)voicing that occurs in voicing languages compared to which occurs in aspirating languages is whether or not it is actively produced by articulatory gestures. The quality of the acoustic signal serves as a reflex of the activeness of the articulatory manoeuvres: robust voicing that overlaps with the hold phase indicates an *active voicing* whereas weak voicing that only partially overlaps with the hold phase indicates a *passive voicing* (Beckman et al., 2013; Jessen, 2001; Kohler, 1984). That is, active voicing results from active adjustments of the articulators (the lowering of the larynx, the expansion of the vocal tract, etc.) that aim to prevent a drop in the air pressure difference required to maintain vocal fold vibration (Ohala, 1997). It is worth noting that active voicing requires extra articulatory effort to counteract the spontaneous decrease of the air pressure difference. This spontaneous decrease of air pressure difference is known as *passive devoicing* (Jansen, 2004). Passive voicing, however, results from the surrounding phonetic environment which occurs as voicing leakage from an adjacent segment. *Active devoicing* is a consequence of the inhibition of passive voicing through articulation gestures such as actively raising the larynx and decreasing the size of oral cavity (Jansen, 2004).

Applying this mechanism to Voiceless and Voiced stops in aspirating and voicing languages, the activeness of (de)voicing in the two categories is as shown in (1):

(1a) Voiceless stops are actively devoiced in aspirating languages (specification).

(1b) Voiced stops are actively voiced in voicing languages (specification).

(1c) Voiced stops are passively devoiced in aspirating languages (unspecification).

(1d) Voiceless stops are passively voiced in voicing languages (unspecification).

These predictions highlight the importance of the two issues that Keating (1988b) proposed in her study: variability and transparency. First, the description of Voiceless and Voiced stops in (1a) and (1b) shows high transparency and low variability. To illustrate, the phonetic details carrying the distinctive function are expected to be reflective of [spread glottis] (or [tense]) in (1a) and of [voice] in (1b). This transparency can be detected by examining consistency in the phonetic details as well as resistance to change caused by the context or coarticulation. Therefore, for instance, Voiced stops will be produced with robust prevoicing in utterance-initial position in voicing languages despite it being aerodynamically challenging to do so. Second, the description of Voiced and Voiceless stops in (1c) and (1d) shows low transparency and high variability. That is, the unspecified stops in the two categories of voicing will show inconsistent behaviour that is shaped by the surrounding context or affected by spontaneous articulatory gestures. They are less transparent, however, when compared to actively (de)voiced stops due to the possibility of differences between languages with regard to some phonological processes such as voicing assimilation and final neutralisation (language-specific) (Keating, 1988).

It is worth mentioning that the difference between the mechanism of specification in aspirating and voicing languages is not clear-cut. Jansen (2004) proposed that Voiced stops in voicing languages in utterance-final position are passively devoiced. That is, Voiced stops might be specified by [voice] in the phonology but still affected by passive devoicing in such a position. Jansen (2004) also pointed out that Voiceless stops in voicing languages might act like an actively devoiced stop. Another example of cases that might posit a challenge to the specification mechanism is languages that contrast prevoiced and aspirated stops. In such a pattern, variability and resistance to change might be problematic in processes like regressive voicing assimilation in stop-stop cluster across word boundaries. As mentioned earlier, it is expected that actively specified stops trigger some (de)voicing in the preceding member of the cluster. In case both members are specified, the second member of the cluster (C2) is expected to regressively spread the (de)voicing, while the first member (C1) is expected to resist the change. In the current study, such issues will be tackled in Najdi Arabic. I argue that Voiceless and Voiced stops in Najdi Arabic are specified for [spread glottis] and [voice], respectively. This claim is examined by investigating the acoustic correlates of the stops in various contexts to account for how much this claim holds with respect to specification.

One of the main contributions this study aims to pursue in terms of specification is to investigate the compatibility between transparency and variability in the three diagnostics proposed in laryngeal realism: the activeness of (de)voicing, regressive voicing assimilation, and speech rate-effect on prevoicing and aspiration. As pointed out briefly in the opening introduction, it has been proposed that slow speech increases the duration of prevoicing and aspiration in the specified stops but not in unspecified ones (see section 2.2.1 for more details). Accordingly, the behaviour of stops in Najdi Arabic will be an interesting case to test the degree of symmetry between the three types of evidence postulated in laryngeal realism. The present work also aims to show a link between specification and acoustic correlates other than prevoicing and aspiration, including temporal and spectral correlates (see section 2.4 for details).

The issue of *overspecification* hinted at in the previous section refers to the voicing system that shows a voicing contrast specified by two features. Aspirating and voicing languages typically show a two-way voicing contrast in which a single feature is sufficient to mark the distinction, following the "economical representation" principle proposed by Chomsky and Halle (1968). This concept emphasises minimality in the number of features used to specify contrast in languages. The emergence of languages that contrast prevoiced with aspirated stops poses a challenge to the economical representation principle and opens possibilities for overspecification.

1.3. Phonetics and Phonology

The previous two sections shed light on the two basic foundations that the argument of the current study is built on: the nature of distinctive features and the mechanism of specification. The former embodies the phonological side of the debate whereas the latter epitomizes how a representational entity is implemented in the real world through articulation and acoustics, as well as the degree to which specified and unspecified segments mirror the representational entities. This leads us to try to draw the whole picture with regard to the difference between phonetics and phonology and what has been proposed in the literature in terms of their mutual interactions.

One of the most debated topics in the field of linguistics is defining *phonetics* and *phonology*. As shown so far, it is not a clear-cut distinction, as both disciplines focus on speech sounds in human languages. A common distinction between phonetics and phonology is to say that the former describes the physical aspects of the speech sound, which are gradient and continuous in nature, while the latter describes the abstract representations of speech sounds in the mind, showing how they are formed and categorised. The early work of Chomsky and Halle (1968) (SPE) postulated that phonetics is not a part of the *grammar* and should be discussed outside the field of linguistics, assuming disassociation between the phonetic component and the phonological representation. In SPE, the phonetic aspects are controlled by universal principles including the articulators' physiology and aerodynamics.

However, such an “extreme view” (Keating, 1984; Jessen, 1998) has been challenged by many empirical studies that revealed an evident and robust parallel between the two disciplines in speech production and perception (Keating, 1984; Jakobson and Waugh, 1987; Jessen, 1998). The gradient and continuous aspects of language have been highlighted as important, effective components of the study of the linguistic system. Moreover, it has been found that some phonetic aspects that have been traditionally proposed to be beyond the speaker’s control are in fact linked to phonological contrasts or governed by language-specific rules. For instance, the phonological contrast between Voiced/Voiceless stops shows notable variation in their phonetic properties because each language employs different values of VOT to mark the distinction (Cho and Ladefoged, 1999). Many subsequent experimental studies have led to a gradual consensus that there is a robust connection between the categorical and gradient aspects of speech sounds in human languages which consequently gave rise to research for invariant acoustic characteristics of the abstract phonological entities. Jessen (1998) summarised the theoretical models that assume an interaction between phonetics and phonology and divided them into two categories: *interface models* and *integration models*.

Jessen (1998) demonstrated that *Interface models* assume two separate components with an interface between them: the phonetic component and the phonological component. The interface is responsible for transforming the categorical phonological component into the continuous gradient component. The phase of the interface is derivational and unidirectional: from phonology to phonetics. *Integration models* propose that phonetics and phonology are deeply connected and inseparable. In integration models, phonetics and phonology “can be compared to a sheet of paper, one side devoted to phonology, the other to phonetics” (Jessen, 1998, p. 29). Such a view assumes a two-way interaction between phonetics and phonology. That is, to search for the phonetic basis of a phonological representation is as crucial as to “explore the consequences of phonetic reality for lexical representation” (Jessen, 1998, p. 31). Following the work of Jakobson and Waugh (1987), Jessen (1998) postulated that the difference between interface and integration models can be reduced by taking into account the crucial requirement each approach emphasized. This results in a model that insists on the importance of categorical representation on the phonological side and the importance of the phonetic reality and its impact on the representational system. Such a model was adopted by Jessen (1998) and Jakobson and Waugh (1987) and is adopted also in the current study.

In the current work, I argue for a perspective that emphasises the robust connection between phonetics and phonology as well as the two-way interactions between the two domains. This is demonstrated by investigating a range of voicing phenomena in stops, including the feature representation of the voicing contrast in the laryngeal system and the processes that involve voicing such as speech rate effect, voicing assimilation and final devoicing conditioned by the phonetic context.

This dissertation aims to answer the following questions:

1. What are the acoustic correlates of the stop-voicing contrast in Najdi Arabic and how are they implemented across the following phonetic contexts: utterance-initial, utterance-medial intervocalic, utterance-final, and across-word-boundary clusters?
2. Employing the laryngeal realism approach, how does the voicing system of Najdi Arabic behave in terms of the following processes: speech-rate effect on the acoustic correlates of stops across the examined phonetic contexts, the acoustic activeness of voicing/devoicing of stops across the examined phonetic contexts, and regressive voicing assimilation in across-word-boundary clusters.
3. In light of the results derived from the preceding inquiry, is Najdi Arabic a voicing or an aspirating language? What does that mean in terms of the phonological representation/specification?

1.4 Synopsis

Chapter 1 introduces a description of the purpose of the study and addresses the main theoretical basis for the work based on the phonetic and phonological aspects of voicing contrast in stops. It starts with a succinct explanation of the importance of investigating voicing contrast phenomenon by showing various phonetic patterns found in different studies, and how they contribute to the field by forming the experimental foundations for the theoretical models of voicing. The aspirating/voicing classification is discussed taking into account the phonetic manifestations of the stops' acoustic properties based on the position within the word and the utterance. Additionally, this chapter provides a brief sketch of the main aspects of laryngeal realism and the types of evidence that will be implemented in the current study. The opening section of the chapter ends the with justification for choosing Najdi Arabic as a target dialect and the research questions.

The remaining part of chapter 1 focuses on three main topics: distinctive features for voicing contrast, phonetic and phonological specification, and phonetics and phonology. These three topics are intended to provide the broad foundation and framework that the argument of the present study is established on. The first topic previews the distinctive features that have been proposed in the literature to specify voicing contrast in stops among languages. It shows various views regarding the suitable distinctive features that phonologically represent voicing contrast and how different features could be interpreted in terms of their articulatory and acoustic content. In addition, some discussion is provided on the conceptual status of distinctive features (monovalent or binary) and how that impacts the characterisation of voicing contrast. The second topic presents the phonetic and phonological specification in aspirating and voicing languages. It previews the acoustic characteristics of the specified stops compared to the unspecified ones across the phonetic contexts. The notion of passive and active (de)voicing is identified with respect to the aspirating/voicing classification. The third topic shifts the focus to the literature on the nature of the interactions between phonetics and phonology, showing the different views regarding the connection between the two disciplines which range from assuming the complete independence to proposing complete integration.

Chapter 2 shows the main phonological and phonetic aspects of voicing contrast in stops in the literature and their phonological implications. It begins with a description of the main arguments of the theoretical models of voicing that looked at the phonological and phonetic aspects of voicing contrast in stops among languages. A special section deals with the laryngeal realism approach and presents its main assumptions and the types of evidence employed in its literature, including speech rate effects on the duration of prevoicing and aspiration, passive and active (de)voicing, and voicing assimilation across word boundaries. The articulatory and aerodynamic aspects of voicing contrast are introduced with a description of voicing initiation, retention, and cessation. The chapter sheds light on the patterns of voicing and aspiration as well as temporal and spectral acoustic correlates that differentiate Voiced and Voiceless stops with a specific focus on prevoicing and aspiration (VOT). The distinction between *voicing* and *aspirating* languages is included in terms of the phonation and spectral patterns each category shows in different phonetic contexts including utterance-initial, utterance-medial intervocalic, and utterance final.

Chapter 3 previews the studies that focus on voicing contrast in stops in modern Arabic dialects. The chapter identifies the voicing contrast patterns which demonstrate that Arabic, in general, is a voicing language. The chapter presents two types of studies: 1) studies that focus only on the phonetic aspects of voicing contrast, and 2) studies that employ laryngeal realism in the investigation of voicing contrast. A growing number of instrumental studies, although relatively few, revealed interesting patterning of voicing among the dialects of Arabic. Drawing on the proposals in the studies that show Najdi Arabic has voicing contrast with both aspiration and prevoicing, crucial questions are raised with regard to the phonological features and the acoustic properties that signal the distinction in Najdi Arabic. A special section discusses the phonological specification of voicing contrast in Najdi Arabic with a focus on whether the inclusion of emphatics is required or not. The chapter ends with a section that demonstrates the motivation and justification of the current study.

Chapter 4 shows the methodology used in the current study in terms of participants, stimuli, data collection, acoustic analysis, and statistical analysis. The experiments designed to examine the acoustic features of Voiced and Voiceless stops in Najdi Arabic in different phonetic contexts including utterance-initial, utterance-medial intervocalic, utterance-final, and stop-stop clusters across word boundaries. Furthermore, for each phonetic context, a different phonological process is investigated. For utterance-initial position, the duration of prevoicing and aspiration is examined in normal and fast speech rate conditions. The purpose of this analysis is founded on the assumption that variability in the duration of prevoicing and aspiration as a result of speech rate effect implies the presence of an active phonological specification. For utterance-medial intervocalic stops, the strength and duration of voicing in the constriction phase are investigated to identify whether the voicing is a consequence of an active phonological specification (active voicing) or a consequence of a coarticulation process (passive voicing). For utterance-final position, the analysis focuses on the neutralisation phenomenon (final devoicing), considering the notion of incomplete neutralisation and its impact on the distinction between voicing and aspirating languages. For stop-stop clusters across word boundaries, the analysis covers the acoustic features of the first and second segments and how they interact with the phonological specification for each member of the cluster. This analysis is founded on the assumption that actively (de)voiced stops trigger some (de)voicing in the preceding member of the cluster.

Chapter 5, 6, 7, and 8 present the results of the acoustic analysis of Voiceless and Voiceless stops in the aforementioned phonetic contexts. The results for the patterns of voicing and aspiration and the temporal and spectral correlates of stops are presented through figures and tables in terms of voicing status, context, gender, speech rate, place of articulation, and vowel quality. The results indicate that the voicing contrast in Najdi Arabic shows the features of both aspirating and voicing languages. This claim is supported by the following: 1) Voiceless stops are aspirated in the majority of the tokens across the phonetic contexts, 2) Voiced stops are robustly prevoiced in utterance-initial and utterance-medial contexts, 3) aspiration and prevoicing are lengthened in response to slow speech rate compared to fast rate, 4) both Voiceless and Voiced stops trigger some regressive (de)voicing in stop-stop clusters. The results also show that Voiced stops are devoiced in the majority of tokens utterance-finally.

Chapter 9 discusses the results of the acoustic analysis with a focus on the parallels between Najdi Arabic and the acoustic features of stops in aspirating and voicing languages reported in the literature. The phonological implications of the acoustic analysis in each examined context is discussed. These arguments provide the answers for the first and second research questions of the current study. The third question is addressed by the focus on the position of Najdi Arabic with regard to the aspirating/voicing classification. The argument concludes that Najdi Arabic takes a middle position and the contrast is overspecified by two distinctive features [spread glottis] and [voice]. Phonological accounts for the proposed overspecification are discussed.

Chapter 2. The phonological and phonetic aspects of voicing contrast

In chapter 1, I identified the basic foundations and framework of voicing contrast in stops in terms of the nature of the representational system and the variety of proposals that attempt to capture specification, and interactions between the phonetic and phonological representations. From the perspective of this study, based on the previous literature, investigation of both phonological features and phonetic details is crucial for characterising the laryngeal system of a language.

This chapter summarises the phonetic and phonological aspects of Voiced and Voiceless stops on the basis of previous literature. Different implications of the phonetic details for the phonological representations are discussed when necessary. Given that the amount of work on the phonetic characteristics of voicing contrast is huge, I will focus on the main aspects in the field, taking into consideration their relevance to the present study, with special focus on the distinction between voicing and aspirating languages.

Section 2.1 presents the theoretical models of voicing and previews each model's perspective with regard to the phonetic and phonological aspects of the voicing contrast in stops among languages. Section 2.2 describes the basic assumptions of the laryngeal realism approach with a focus on three topics: speech rate effect (section 2.2.1), final devoicing (section 2.2.2), and regressive voicing assimilation in stop-stop clusters across word boundaries (section 2.2.3). Section 2.3 introduces the articulatory aspects of voicing and aspiration which are essential in analysis of the acoustic correlates of the distinction between Voiced and Voiceless stops. It also forms the foundation for the difference between passive and active voicing. Section 2.4 identifies the temporal and spectral acoustic correlates of the opposition between Voiced and Voiceless stops with a special focus on VOT and its interaction with linguistic and non-linguistic factors. The associations between the acoustic correlates and the phonological features are discussed in the light of the distinction between voicing and aspirating languages.

2.1 The theoretical models of voicing

This section aims to briefly preview the main aspects of the theoretical models of voicing contrast within feature theory in terms of the nature of the phonological representation, the interactions between the phonological and phonetic components, and the manifestation of voicing and aspiration with respect to the aspirating/voicing distinction. This preview primarily focuses on stops in two-way contrast languages and pays more attention to the acoustics but refers sometimes to articulation and auditory aspects proposed in some of the models.

As pointed out earlier, both sides of the debate, embodied in the Jakobsonian approach from one side and SPE from the other, shaped the foundations for the competing views proposed in theoretical models of voicing. That is, the origins of the differentiation in characterisation of theoretical models of voicing lies in the conceptual development of *distinctive features* throughout the history of sound structure studies, which ranged from being purely formal functioning abstract entities responsible for sound categorisation, with no interaction with the phonetic reality (abstractness), to being phonetically grounded and deeply connected to phonetic events (naturalness) (Rooy and Wissing, 1998). Some models took a middle position by drawing a clear separation between the categorical and continuous components but assuming a phonetics-phonology interface. By using the arguments of SPE and the Jakobsonian approach as a starting point to direct the discussion in this section, the following paragraphs start by giving a succinct description of SPE followed by a detailed description of two models that are closely connected to SPE, including the model of Halle and Steven (1971) and the model of Keating (1984) (Keating's model is intended to be an improvement on SPE, as explicitly stated in her study). The second part will focus on the models that assume a strong connection between phonetics and phonology (integration models). The models that adopt this view include Kohler's model (1984), Kingston and Diehl's model (1995), and Jessen's model (1998, 2001). The third part of this section previews the predictions of the aforementioned models in terms of aspirating/voicing classification and its relevance to the current study.

2.1.1 The SPE model by Chomsky and Halle (1968)

In SPE, lexical items are represented phonologically as matrices of features with binary values "+/-" in which each row presents a feature whereas each column presents a segment. In the phonetic level, the same features are used but with scalar values rather than binary values. The phonological rules transform the phonological components into phonetic structure while the phonetic rules convert the binary values into quantitative values that are shaped by universal phonetic principles. The output of the universal phonetic principles is proposed to be not specified by any rules which are language-specific. In SPE, the features are defined in articulatory terms describing the active articulators' movement. The distinctive features proposed in SPE to describe voicing contrast in stops are: [voice], which is associated with a vocal cord adjustments that enable voicing to occur (not the actual vocal cords' vibration); [tense], which indicates articulatory gestures (the tension of walls of the vocal tract) that inhibit voicing; [heightened subglottal pressure], which is associated with "extra energy for aspiration"(Keating, 1988a, p. 17); and [glottal constriction], which is associated with the gestures that enhance voicing. These features were centred around describing the articulatory adjustments required at the moment of the release to initiate voicing and aspiration, which make them complicated and not straightforward (Keating, 1988a). Keating (1988a) states that "The various features interacted in somewhat complicated ways in determining vibration, aspiration, etc., and were therefore perhaps too hard to learn to use, rather than theoretically unacceptable; there was also little evidence presented in their support". (Keating, 1988a, p. 17).

2.1.2 Halle and Stevens' model (1971)

Another model that used articulatory parameters was proposed by Halle and Steven (1971). Halle and Stevens (1971) proposed a set of binary features that characterise the articulatory settings of the vocal cords at the moment of the release. The features were postulated to specify all possible voicing system types in all languages. Beside voicing and aspiration, they intended to account for aerodynamics (airflow mechanism), phonation type, and fundamental frequency. The features were as follows: [±spread glottis], [±constricted glottis], [±stiff vocal cords], [±slack vocal cords]. The features [±spread glottis] and [±constricted glottis] were associated with glottal opening or vocal cords abduction, while [±stiff vocal cords] and [±slack vocal cords] were associated with vocal cords stiffness. To account for the phonetic aspects of segments found in languages, Halle and Stevens espoused the notion that "these four feature are not completely independent" (Halle and Stevens, 1971, p. 50). That is, a combination of these features should depict the articulatory configuration needed for the classification of segments in terms of aspiration, voicing, and glottalization. Focusing on the phonetic categories expected to occur in aspirating and voicing languages, prevoiced stops were described as [-spread glottis, -constricted glottis, -stiff vocal cords, +slack vocal cords], unaspirated stops were described as [-spread glottis, -constricted glottis, +stiff vocal cords, -slack vocal cords], and aspirated stops were described as [+spread glottis, -constricted glottis, +stiff vocal cords, -slack vocal cords]. It can be noticed that the opening and stiffening of vocal cords enhance aspiration and suppress voicing whereas the closing and slacking of vocal cords enhance voicing and suppress aspiration. With regard to F0, the model posited that the effect on F0 at the onset of the following vowel is an automatic by-product ramification of vocal cords' stiffening in case of F0 raising, and vocal cords' slacking in case of F0 lowering.

It is evident that the model of Halle and Stevens presented a meticulous characterisation of the articulatory manoeuvres of the vocal cords that take place during the production of Voiceless and Voiced stops at the moment of release. However, it can be noticed that both SPE and Halle and Stevens' models discussed the articulatory setting required for voicing initiation but not the actual vocal cords vibration (Keating, 1988a). However, [±spread glottis] and [±constricted glottis] were widely employed by many experimental models because of their association with aspiration and glottalization which can be accurately detected in the acoustic signal. Additionally, among experimental approaches, the notion of VOT classification proposed by Lisker and Abramson (1964) was more useful and straightforward when compared to Halle and Stevens' model in terms of characterising the phonetic manifestation in languages (Keating, 1988a). Some of the articulatory adjustments such as the stiffening of vocal cords in the production of Voiceless stops were not confirmed by articulatory studies (Jessen, 1998; Keating, 1988a).

2.1.3 Keating's model (1984)

Keating (1984) proposed a model for voicing in stops in languages with two-way contrast specified phonologically with a single feature [± voice]. This model was an attempt to deal with some problems in SPE and Halle and Stevens' models in terms of the features at the phonetic and phonological levels. Keating emphasized that the problem in these models emerged because the features were proposed to be the same in both the phonetic and phonological levels in order to account for all the phonetic differences found in languages. To solve these issues, she emphasized the importance of the acoustic features of voicing and aspiration as a source of determining the voicing typology among languages. She also posited that the function of features in the phonological level is to categorize natural classes, and they do not contain any phonetic details. Keating proposed a new level consisting of what she called "the major phonetic categories" that have three phonetic features based on the VOT patterns proposed in the work of Lisker and Abramson (1964); voice lead, short lag, and long lag. The phonetic features in this new level were {voice}, {vl.unasp}, and {vl.asp}. This level is separate from the phonological level. It is noteworthy that these proposed features are earlier in the derivation than the output level that contains the continuous aspects of the acoustic signal. The features in the posited new level were meant to be categorical and discrete, and abstract in nature, to allow for free acoustic manifestations among languages. Based on this model, aspirating and voicing languages (although the aspirating/voicing classification was not used in Keating's model) employ [± voice] in the phonological level but differ in the major phonetic categories. Aspirating languages usually employ {vl.unasp} and {vl.asp} to signal the distinction while voicing languages use {voice} and {vl.unasp}. With regard to the possibility of a two-way voicing system that employs {voice} and {vl.asp}, Keating proposed that this is highly unlikely to be the case.

One of the challenges for Keating's model is the mapping between the major phonetic categories and the phonetic (acoustic and articulatory) output which is continuous in nature. Keating dealt with this issue by proposing that the mapping from the major phonetic categories into the continuous output is affected by universal phonetics, but the process of choosing between the three features is language-specific and affected by the context. To account for positional variations in voicing and aspiration, Keating posited that {voice} stops show prevoicing in word-initial and voicing during closure when preceded by a sonorant. This voicing realisation might differ in its timing and strength. {voice} stops in final position might not be released but might show some voicing in the closure. Regardless of the amount of voicing, "in all cases, … the stop closure crucially contains some low-frequency vibration" (Keating, 1984, p. 295-296). In terms of {vl.asp}, {vl.asp} stops show aspiration after the release in both initial and medial positions (relatively shorter than word-initially). The closure phase of {vl.asp} might show a very short voicing tail from the preceding voiced segment. Finally, {vl.unasp} stops differ from {voice} stops in relation to the relative absence of voicing in the closure, and from {vl.asp}by showing shorter aspiration.

One of the main assumptions of Keating's model is the flexibility of the phonetic manifestations assumed in the model. Although all languages in Keating's model are limited to choose from the three major phonetic categories, their implementation in the output level is controlled by language-specific rules for each context. This modelling approach allows for variation within aspirating and voicing languages. In that regard, Voiced stops ([+voice]) in English could be realised as {voice} or {vl.unasp} whereas Voiceless stops ([-voice]) could be realised as {vl.unasp} or {vl.asp}. In languages like French, Voiced stops are always realised as {voice} while Voiceless stops are realised as {vl.unasp}. The phonetic and phonological identity of {vl.unasp} stops is an interesting case in terms of the aspirating/voicing classification. Keating postulated that the differences between {vl.unasp} stops that implement [-voice] (in voicing languages) and that implement [+voice] (in aspirating languages) are not contrastively employed. However, the differences between them depend on the phonetic characteristics of the counterpart category in their system to achieve the maximal distinction. This principle was called *polarisation* in Keating's model. Based on this principle, the main purpose of {vl.unasp} is to "act as a swing category" (Keating, 1984, p. 309) that maximizes the distinction from their counterpart segments ({vl.asp} in aspirating languages and {voice} in voicing languages).

To sum up the differences between the basic assumptions of the aforementioned models, it can be noticed that Keating's model is different in various ways. First, Keating's model introduced a new level with abstract discrete features that aim to describe the phonetic manifestation typology (articulatory and acoustic) of voicing and aspiration among languages, and, more importantly, these features are separable and independent from the phonological feature [±voice]. Second, the phonological feature [± voice] in Keating's model does not have phonetic details. Third, the phonetic details in the output level are language-specific. One of the essential aspects of Keating's model is the minimality of the features in the phonological level; a single feature. This is not the case in the SPE and Halle and Stevens models in which more than one feature is required to describe voicing and aspiration (redundancy) (Jessen, 1998). The impact of context on the phonetic manifestation is accounted for in Keating's model which makes it more useful for experimental studies.

In terms of similarities, all three models adopted binarity in their representational system. However, (+) and (–) in SPE and Halle and Stevens models are equivalent to the terms "with" and "without", respectively, whereas in Keating's model they are equivalent to "more" and "less" to express relational terms. SPE and Keating's models also showed symmetry in terms of the abstractness of the distinctive features. With regard to the phonetics and phonology interaction, Keating's model adopted the view that the categorical components are transformed into continuous structures through an interface, called the phonology-phonetics interface. Accordingly, as in SPE, it is unidirectional process from the phonology to the phonetics. To illustrate, an investigation would concentrate on looking for the phonological consequences for phonetic details but not the opposite.

Unlike these interface models, integration models assume two broad issues: 1) a robust connection between the phonetic and the phonological components, 2) two-way interactions between phonetics and phonology in which integration models "explore consequences of phonetic reality for lexical representation" (Jessen, 1998, p. 31) and vice versa. The integration models gained more attention with the advancement of experimental phonetic tools that showed compelling evidence to support the robust connection between the phonetic and phonological components. In the following paragraphs, I will present models that adopt this view including Kohler's model (1984), Kingston and Diehl's model (1995), and Jessen's model (1998, 2001).

2.1.4 Jakobson and colleagues' model (1952, 1979, 1987)

Before delving into the details of more recent models, this is a brief description of the main aspects of the Jakobsonian approach for voicing contrast which proposed an early integration framework between distinctive features and the phonetic component. It is important to mention that the concept of distinctive features in the work of Jakobson and colleagues concentrated on the distinctive function associated with the sound property that signals oppositions across all sources of variation. That is, the distinctive features were defined in articulatory, acoustic, and auditory terms based on the "common phonetic denominator" (the relational invariance across all factors including language, context, speaker etc). This generalisation of the sound property across sources of variation is called by Jessen (1998, 2001) *relational invariance*. The characteristics of the distinctive features that are associated with a specific language or context are called the *correlates*. The distinctive features associated with the voicing contrast in the work of Jakobson and colleagues were the binary features [±voice] and [±tense]. The feature [±voice] is defined by the presence or absence of vocal cord vibration which can be acoustically identified as the low frequency voicing bar in the holding phase of the stop. [tense], on the other hand, is associated with the duration of aspiration, closure, and the preceding vowel. According to Jakobson and colleagues, in languages with two-way contrast, some languages employ [voice] such as French while others use [tense] such as English.

Jakobson and colleagues' approach to examining distinctive features inspired a plethora of experimental studies that aimed to reduce the distinction between the phonological component and phonetic reality. At the heart of the experimental framework lies the notion of *distinctness* in the conceptualization of the distinctive features. The distinctive feature that specifies voicing opposition by actual articulatory and acoustic voicing should be [voice]. On the other hand, if the opposition is signalled by aspiration, the feature should be [spread glottis] (or [tense]). This type of theoretical assumption forms the basis of integration models that postulate a two-way interaction in which the phonetic reality is as important as the phonological components in characterising sound systems among languages.

2.1.5 Kohler model (1984)

Another model that postulated an integration perspective between the phonological component and the phonetic reality to characterise the voicing contrast has been proposed by Kohler (1984). Kohler's model of voicing distinction in obstruents emphasizes the role of articulatory, acoustic and perceptual details in the definition of the distinctive features, with no interface level. Unlike the interface models, timing was accounted for in the interaction between the phonological features and the phonetic level. Kohler's model is based on the binary feature [±fortis]. This feature is related to articulatory power which is defined as "power in the supraglottal movements and in the air stream, and with tension, especially in the larynx' (Kohler, 1984, p. 168). Voiceless obstruents are fortis while Voiced ones are lenis (according to Jessen 1998, fortis vs lenis or tense vs lax share the same meaning in the majority of the literature; in the present work I use Voiceless/Voiced terminology). The binary feature [±fortis] in this model consists of two components: 1) the supraglottal gestures and 2) the properties of the glottis which include voicing and aspiration. According to Kohler, the first component is potentially universal while the second is language specific. For the glottis action, a voicing distinction between stops can be observed by investigating the presence or absence of pre-voicing in the stop closure or by measuring the duration of aspiration in the case of aspirated and unaspirated stops. These acoustic features differ based on the language and the position within the word and utterance. For aspirating languages, the distinction is manifested by aspiration in all contexts or in non-final contexts only (depending on the language). Voicing languages, on the other hand, contrast Voiceless and Voiced stops through the absence of voicing in the former and the presence of it in the latter in non-final stops. In intervocalic position, Kohler posited that the articulatory power of Voiceless stops is weakened, which results in a shorter closure course, shorter preceding vowel, shorter aspiration, and the possibility of passive voicing in the closure. As for intervocalic Voiced stops, they show the features of approximants. In final position, the glottis actions (voicing and aspiration) are difficult to maintain which leads to a role for the timing of the articulators' movement (closure duration and preceding vowel duration) to signal the distinction.

Kohler's model is maximally different form SPE and Keating's models in terms of the distinctive features, the phonology-phonetics connection, and the abstractness of the phonological components. The distinctive feature [±fortis] was proposed to account for all the phonetic manifestations of voicing contrast in stops within a single concept: the articulatory power. With regard to phonetics-phonology interaction, Kohler's model assumes complete integration between the two domains in which the phonetic reality is directly associated with the distinctive features with no interface level. To account for all the phonetic variability within and across languages, a coordinative structure between the oral, velopharyngeal (associated with soft palate and larynx), and glottal valves has been proposed. The coordinative mechanism aims to account for contextual variation by assigning an importance degree to each valve. For instance, the importance of glottis action is high in utterance-initial position which leads to longer aspiration (aspirating languages) or longer voicing (voicing languages). The consequences of articulatory power for f0 in the following and preceding vowel was also discussed in Kohler's model which considers laryngeal tension to be accompanied by high F0 values. As for the possibility of a language with two-way contrast employing both aspiration and voicing in Voiceless and Voiced stops, respectively, Kohler postulated that "if the lenis feature is manifested in active closure voicing, the fortis feature does not include aspiration; if the fortis feature is accentuated by aspiration, the lenis feature does not require active voicing" (Kohler, 1984, p. 154). This is in agreement with what was proposed in Keating's model (1984) in relation to the possibility of a voicing system that contrasts prevoiced with aspirated stops.

With regard to the phonetics-phonology connection, Kohler's model proposed a two-way interaction in which the phonetic reality is crucial for characterising phonological features. As for the abstractness of the distinctive features, it can be noticed that there are similarities between the work of Jakobson and colleagues and Kohler's model in terms of dispensing with the abstractness of the distinctive features proposed in the SPE tradition. However, the difference between the two models is related to the number of features employed to capture the distinction: [tense] and [voice] in Jakobson and colleagues and [fortis] in Kohler's model. One of the consequences of assuming a single feature in Kohler's model is the need for a coordinative structure in which a combination of articulatory manoeuvres mutually participates in the output of the acoustic signal (Docherty, 1992).

2.1.6 Kingston and Diehl's model (1995)

Another model to explain voicing distinction has been proposed by Kingston and Diehl (1995), which argues for an auditory explanation of the voicing contrast and forms the basis for the auditory enhancement hypothesis. Kingston and Diehl's model is based on the binary feature [±voice], but the authors proposed that there is an intermediate level between the distinctive feature and the acoustic details, namely, intermediate perceptual properties (IPP). They state that "speakers covary articulation precisely because their acoustic consequences are auditorily similar enough to be integrated into more comprehensive perceptual properties, intermediate between the acoustic properties and distinctive feature values" (Kingston and Diehl, 1995, p. 7). This intermediate level combines further acoustic cues to describe the distinctive features, and for the feature [voice], three properties have been proposed in the (IPP): aspiration, low frequency property, and consonant/vowel duration ratio. Low frequency property has three sub-properties: F1 onset, F0 onset, and closure voicing. C/V duration ratio has three sub-properties: preceding vowel duration, closure duration, and closure duration/vowel duration interaction. The sub-properties of each IPP are expected to be perceived by the listener as implementations of the same property. Accordingly, F0/F1 and closure voicing perceptually share the same manifestation, namely, low-frequency property.

The model was based on a set of accurately built perception experiments using synthesized non-speech tokens that aimed to investigate the effectiveness of the proposed IPP properties in identifying the voicing distinction. Specifically, they investigated to what extent the sub-properties mutually enhance the saliency (perceptibility) of Voiced and Voiceless stops and which of them integrate to achieve the same perceptual target. As seen in figure 2.1 below, the results of the experiments reveal that closure duration is integrated with preceding vowel duration, closure duration is integrated with closure voicing, and F1 onset/offset is integrated with closure voicing. The results also showed that F1 integrates with closure voicing, F0 integrates with closure voicing, F1 and F0 do not integrate with each other (Kingston et al., 2008). The results for the synthesized tokens showed that when closure voicing in Voiced stops is accompanied with low F1/F0 in the adjacent vowels, this led to the best identification of this category. Similarly, when voiceless closure in Voiceless stops is accompanied by F1/F0 lowering, this resulted in more accurate identification of Voiceless stops. These results suggest that they mutually play a perceptual role in the phonological opposition of voicing in stops. The model did not elaborate on the sub-properties of aspiration and the status of aspiration remained unclear.

The main assumption of Kingston and Diehl's model is that the phonological oppositions, is better explained by perception than articulation which means the speaker fits the articulation process to meet the requirements of the perceptual mechanism implemented by the listener. Kingston and Diehl state that "This perspective implies that speakers have knowledge of the mechanisms that listeners apply to the task of recognizing speech sound, and that this knowledge prescribes reorganizations of articulatory behaviours to take advantage of these mechanisms. It should be clear that such recognitions require the phonetic component to be controlled" (Kingston and Diehl, 1994, p. 446). The postulation that these acoustic properties are controlled argues against automatic effect accounts which assume that variations in F1/F0 onset values are by-products of the laryngeal articulatory adjustments. According to Kingston and Diehl's model, F0/F1 raising or lowering signal the phonological opposition regardless of the phonetic voicing manifestation. Therefore, [+voice] stops are combined with low F1/F0 whether or not there is a voicing in the closure.

The contribution of Kingston and Diehl's model is valuable in terms of providing a perceptual basis for the phonological feature [± voice] and it was supported by well-constructed experiments. The phonological representation in the model seems to assume that [± voice] specify the voicing contrast in both voicing and aspirating languages. In this view, aspiration was considered as an IPP property for [± voice] but with no sub-properties in the acoustic signal. The model can be useful in terms of the voicing/aspiration classification when discussing the closure voicing. Considering the acoustic manifestation of voicing in aspirating languages, it has been pointed out in the literature that Voiceless stops are prone to passive voicing in intervocalic position. In this model, the presence of passive voicing is not expected to signify a voicing feature (F1/F0 lowering) since F0/F1's role is to indicate phonological voicelessness. This might be problematic when accounting for the role of aspiration in F0/F1 raising that is expected to occur in Voiceless stops in aspirating languages. Such issues, with regard to the interchangeability of impact between aspiration as well as the closure voicing on F1/F2, led Jansen (2004) to propose that speakers of aspirating languages rely more on aspiration whereas speakers of voicing languages rely more on low-frequency properties. More studies are needed to confirm this assumption or reject it.

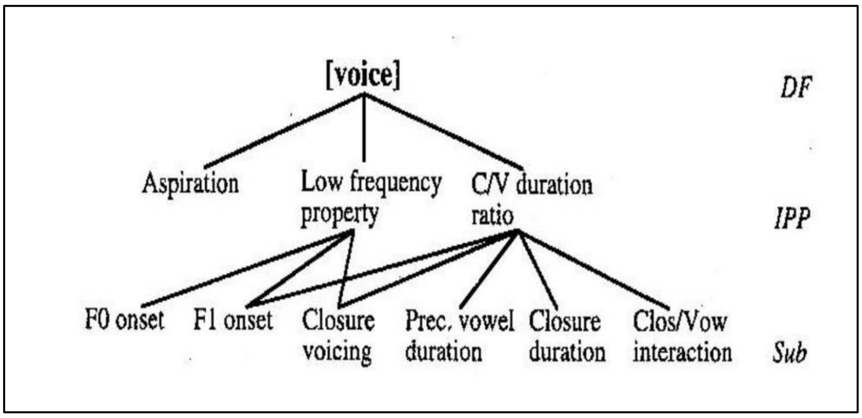


Figure 2.1 The interaction between IPP and the acoustic cues Kingston and Diehl's model. (Adopted from Jessen, 1998, p. 266)

Kingston and Diehl's model inspired many phoneticians and phonologists to work on the notion of enhancement and on the compensation mechanism between acoustic correlates in voicing and aspirating languages in various contexts within the utterance and the word. The model of voicing contrast proposed by Jessen (1998, 2001) provided a detailed account that looked at various acoustic correlates and in various contexts.

2.1.7 Jessen's model (1998, 2001).

Jessen (1998, 2001) considered two features in his model of voicing distinction: [voice] and [tense]. His model is based on explaining distinctive features through *acoustic invariance* which refers to the acoustic property that is consistent and generalisable across all sources of variation (following Jakobson and colleagues' assumption of a common phonetic denominator). Jessen (2001) highlighted the importance of the acoustic correlates of the distinctive features for two reasons: 1) acoustic analysis is a prerequisite for perceptual analysis, and 2) acoustic analysis can be reflective of articulatory gestures. Jessen classified the acoustic features of [voice] and [tense] into two types: 1) basic correlates which occur in all (or most) of the contexts and 2) non-basic correlates which occur under limited conditions. A basic correlate, unlike a non-basic correlate, shows "contextual stability" and "perceptual saliency". The former means that the correlate signals the distinction in the majority of contexts whereas the latter means "manipulation of that correlate alone in a speech perception experiment leads to a categorical perception" (Jessen, 2001, p. 243). The role of non-basic correlates is to enhance the basic correlates or to replace them if the basic correlates are weak or not present. It is evident that Jessen interrelates the features [voice] and [tense] with acoustic properties which helps in explaining the laryngeal systems of both voicing and aspirating languages. The hierarchical classification of acoustic correlates into basic/non basic is another aspect that sheds light on the notion of a compensation mechanism that might occur in certain phonetic contexts. For instance, the preceding vowel duration signals the distinction in final stops in English instead of aspiration, due to neutralisation.

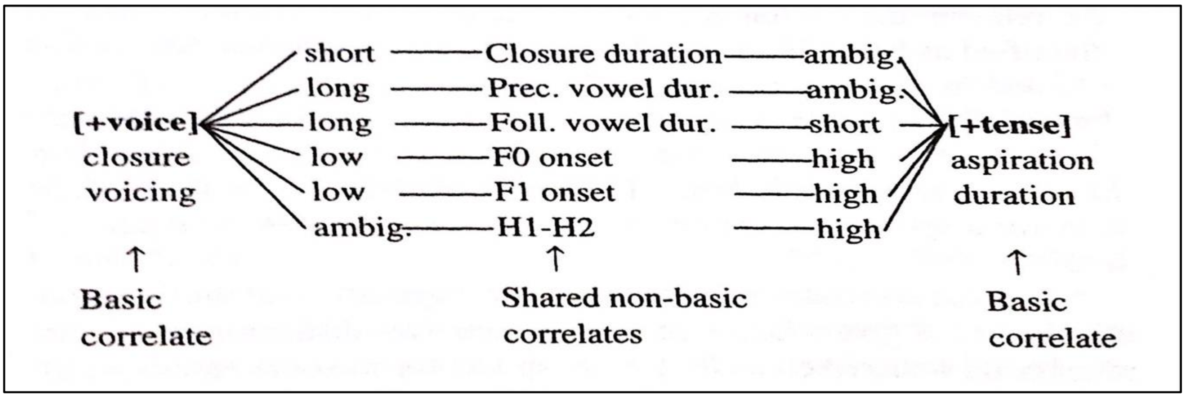


Figure 2.2 Basic and non-basic correlates of [voice] and [tense]. (Adopted from Jessen, 2001, p. 224)

Jessen's model follows the Jakobsonian approach in terms of defining the distinctive features in concrete phonetic terms with respect to articulation, acoustics, and perception (The acoustic correlates will be discussed in detail in section 2.4). He reintroduced the feature [tense] and proposed it as a phonological feature that specifies the voicing contrast in German. [tense] in Jessen's model is mainly defined as increased duration in "events in the consonant and its surrounding" (Jessen, 2001, p. 242), which include aspiration, closure, preceding and following vowel. The set of correlates proposed for [tense] and [voice] are supported by many studies and can be implemented to test the phonetic manifestation of voicing contrast in both voicing and aspirating languages. Another advantage of Jessen's model is the clarity of contextual variations and the enhancement mechanism between the basic and non-basic correlates.

Jessen presented a comparison between [tense] and [spread glottis] and concluded that [tense] is better for specifying the voicing contrast in German. The justification for his view was based on two observations: 1) aspiration does not occur in the majority of tokens, 2) [spread glottis] is associated with the moment of release only and cannot account for the durational aspects in the surrounding context such as preceding vowel and closure duration. Jessen pointed out the distinction between [tense] in his study and [fortis] in Kohler's model. [fortis] in Kohler's model accounts for voicing contrast in both voicing and aspirating languages by incorporating timing and articulatory power in the articulatory and acoustic manifestation of the contrast. By incorporating timing, it can be noticed that the conceptualisation of distinctive features in Kohler's model is scalar in nature unlike that proposed in Jessen's model which showed dichotomy approach (Jessen, 1998). This difference led Jessen to limit [tense] to aspirating languages and adopt [voice] for voicing languages.

2.1.8 Summary and discussion of theoretical models of voicing

The review presented above shows the previous models that attempt to identify the phonetic and phonological aspects of voicing contrast in stops in a wide range of languages. As noted, each model has its own merits and contributes to the field by proposing descriptions that aim to accurately address the phonetic and phonological aspects of voicing contrast. In the following paragraphs, I sum up the similarities and differences between these models in terms of the interaction between the phonological component and the phonetic reality, the nature of the distinctive features chosen to specify the opposition, the phonetic content of the distinctive features, and the phonetic manifestation in the model with relation to aspirating/voicing classification.

In terms of the features chosen to specify the voicing contrast in stops and their nature, we have seen that some of the models posited a single feature: the binary feature [± voice] is used in Keating's model (1984) and Kingston and Diehl (1995), or the binary feature [±fortis] in Kohler's model (1984). There are also models that employed two features, [±voice] and [±tense], to capture the difference between voicing and aspirating languages such as the model of Jakobson and colleagues (1952, 1979, 1987) and the model of Jessen (1998). There are also models that proposed four features. The SPE model (1986) proposed four features [±voice], [±tense], [±heightened subglottal pressure], and [±glottal constriction]. Halle and Stevens' model (1971) proposed four features [±spread glottis], [±constricted glottis], [±stiff vocal cords], [±slack vocal cords]. There are several reasons that led to variation with respect to the number and nature of the features. Some of the models define the features in articulatory terms such as the model of SPE, Halle and Stevens, and Kohler. It can be noticed that the articulatory terms were reflected in the names of some of the proposed features such as [±spread glottis], [±constricted glottis], [±stiff vocal cords], [±slack vocal cords]. Other models consider acoustics to be primarily the basis of the phonetic content in defining the features such as the models by Keating and Jessen.

Another reason for the variation among the proposed features stems from the models' perspectives on the connection between distinctive features and phonetic details. The models that emphasized a robust connection between the phonetic and phonological component used phonetic details (articulatory, acoustic, perceptual) as a mirror that reflects what the feature is in the phonology. These models adopted the a two-way interaction between phonetics and phonology such as the model of Jakobson and colleagues, Kohler, Kingston and Diehl, and Jessen. Another source of differentiation between the models is the scope of focus within the sequence of events during the production of Voiceless and Voiced stops among languages. To illustrate, it can be noted that SPE and the model of Halle and Stevens concentrated on the articulatory adjustments that initiate or suppress voicing and aspiration (Keating, 1984, 1988a). Kohler's model solved this problem by accounting for all the articulatory events and representing variation through a coordinative mechanism that considers timing and phonetic power. Other models concentrated on the actual articulatory gestures during the production of voicing and aspiration such as the models of Jakobson and colleagues, Kingston and Diehl, and Jessen.

The status of prevoicing and aspiration (VOT) was crucial in the models of Jakobson and colleagues, Keating, Kohler, and Jessen. The importance of prevoicing and aspiration in these models took various degrees. In the model of Keating, the major phonetic categories were based only on prevoicing and aspiration patterns among languages. Other acoustic correlates were discussed in Keating's model as a consequence of prevoicing in voicing languages and aspiration in aspirating languages. The model of Jessen emphasized the role of prevoicing and aspiration by considering them as basic correlates of [voice] and [tense], respectively. The acoustic correlates, besides prevoicing and aspiration, were described and accounted for as non-basic correlates that might be as important as prevoicing and aspiration in some contexts (phonologization). Kingston and Diehl described the sub-properties as controlled factors that are adjusted by the speaker to fit the listener's need to perceive the distinction.

The models were focused on a relatively small number of languages. The Germanic languages were the primary scope of investigation for aspirating languages whereas Slavic languages were the focus of examination in case of voicing languages. In contrast, the present study aims to build a detailed description of voicing contrast in stops in Najdi Arabic which appears to show an uncommon distinction in its laryngeal system by contrasting prevoiced and aspirated stops (Flege and Port, 1981; Al-Gamdi et al., 2019). Languages that show features of both aspirating and voicing patterns have not received sufficient attention in the literature. Keating (1984) and Kohler (1984) pointed out that it is unlikely for languages to employ both prevoicing and aspiration to signal the voicing contrast. Keating, however, mentioned that, based on the principle of polarisation, some voicing languages might increase the aspiration of their unspecified stops (unaspirated stops) to maximize the voicing distinction. It is worth noting that Keating mentioned that the principle of polarisation needs more results from diverse languages to be confirmed. The current study aims to examine the phonetic and phonological aspects of voicing contrast in Najdi Arabic, a variety that employs aspiration and prevoicing in its voicing system.

To identify the phonetic and phonological aspects of this contrast, I adopt the model of Jessen (2001) that emphasized the effectiveness of various acoustic correlates to account for the voicing contrast in voicing and aspirating languages. Some modifications on Jessen's model are applied to fit the purposes of the present study. First, I investigate the presence of both [spread glottis] and [voice] in the voicing system of Najdi Arabic in which Voiceless stops are specified by the privative feature [spread glottis] (rather than [tense]) whereas Voiced stops are specified with the privative feature [voice]. Second, the acoustic properties of the release burst (duration and intensity) are added to the model. Third, the acoustic correlates F0/F1/H1-H2 offset are added to the analysis to account for the utterance-medial intervocalic and utterance-final stops. To test the notion of contextual stability, sources of variability including place of articulation, vowel quality, gender, stress (trochaic/iambic for utterance medial stops) are considered in the analysis for each context.

It will be clear from the proposed analysis in the previous paragraph that this study adopts the first premise of the phonetics-phonology two-way interaction framework, which is to investigate the repercussions of the phonetic reality for the phonological representation. To investigate the implementation of phonological features in the acoustic signal, the present study adopts the types of evidence proposed in the realm of *Laryngeal Realism* including passive/active voicing, speech rate effect, final devoicing, and regressive voicing assimilation in stop-stop clusters across word boundaries, which are employed as tools for characterising the specification mechanism that accounts for the implementation of the distinctive features in the phonetic details. Laryngeal realism and the model of Jessen (1998, 2001) emphasize the basic foundations for the analysis, namely the robust parallelism between the phonetic details and the phonological features proposed in the two-way interaction framework. The following section gives a succinct description of laryngeal realism and the types of diagnostics in its perspective.

2.2 Laryngeal realism

Early generative phonologists posited that distinctive features were binary in nature and loosely associated with their acoustic or articulatory consequences. The theory of laryngeal realism (Avery and Idsardi, 2001; Jansen, 2004; Honeybone, 2002, 2005; Iverson and Salmons, 1995 2003; Brown, 2016; Harris, 1994; Jessen and Ringen, 2002; Vaux and Samuels, 2005; Beckman et. al, 2011,2013; Schwarz et al., 2019) departs from the abstractness of distinctive features that proposed in SPE and emphasizes the connection between the phonological features and their phonetic realization. As noted earlier, this approach was proposed in the early work of Jakobson and colleagues and embraced by the integration models. It pays attention to cross-linguistic variation in articulatory events and their acoustic consequences that speakers use to produce the distinction between Voiced and Voiceless stops.

The term laryngeal realism was first introduced by Honeybone (2001). The laryngeal realism approach is in agreement with the Jakobsonian approach regarding the robust connection between phonetics and phonology which is manifested through the phonetic grounding of distinctive features. The evidence used in laryngeal realism was based on phonetic, phonological, and diachronic aspects in the analysis of languages (Iverson and Salmons, 1995; Honeybone, 2002, 2005). However, the laryngeal realism approach proposes privativity for the distinctive features that specify voicing contrast in stops. Unlike the binary system, laryngeal realism supports the privative or monovalent system in which negative values of distinctive features are meaningless, and a segment that lacks a specific feature should be left unspecified (representational absence). The reason for this assumption is that negative values are conceptually problematic in that they do not require an articulatory movement to be achieved at the phonetic level (Honeybone, 2002). Moreover, if privativity is confirmed to be appropriate for features like [nasal], it is highly useful to consider it also in the case of [voice] and [spread glottis] (or [tense]) which evidently leads to more simplicity in the representational system (Iverson and Salmons, 2003). By acknowledging that phonological features are privative and phonetically grounded, languages within the laryngeal realism approach are classified in terms of laryngeal contrast in stops. On the one hand, voicing languages like Dutch have Voiced stops in word-initial position with long lead VOT and Voiceless stops with short-lag VOT. Aspirating languages like German, on the other hand, have unaspirated stops in word-initial position with short-lag VOT and aspirated stops with long-lag VOT. Laryngeal realism argues that the phonological feature that precisely marks the distinction between homorganic stops in voicing languages is the privative feature [voice] in which Voiced stops are associated with [voice], and Voiceless ones are unspecified. In aspirating languages, however, the contrast is specified with the privative feature [spread glottis] that specifies the aspirated stops, while unaspirated stops are unspecified.

The manifestation of [voice] and [spread glottis] in voicing and aspirating languages has been characterised in laryngeal realism by a set of phonetic details and phonological processes. Honeybone (2005) presented a set of aspects that describe the voicing contrast in aspirating and voicing languages (he used the term "type A" for aspirating languages and "type B" for voicing languages). According to Honeybone (2005), in aspirating languages (type A), 1) Voiceless stops are aspirated in most contexts, 2) Voiced stops might show passive voicing, and 3) in clusters, it is common to find assimilation to voicelessness not to voicedness. In voicing languages (type B), 1) Voiced stops show robust prevoicing, 2) Voiceless stops are unaspirated, and 3) in clusters, it is common to find assimilation to voicedness. As pointed out in several sections of this study, the types of evidence proposed in laryngeal realism to form the basis for classifying languages on aspirating or voicing include the acoustic manifestation of the acoustic correlates across contexts, passive and active voicing, speech rate effect on aspiration and prevoicing, final devoicing, and regressive voicing assimilation in stop-stop clusters across word boundaries. The first two of these types have been discussed in the previous sections on several occasions. The remaining three will be discussed in the following sections.

2.2.1 Speech rate effect on voicing contrast

Several studies have shown that speaking rate affects the temporal acoustic correlates of Voiced and Voiceless stops (Miller et al., 1986; Volaitis and Miller, 1992; Kessinger and Blumstein, 1997; Nagao and de Jong, 2007; Beckman et. al, 2011; Magloire and Green, 1999; Solé and Estebas, 2000; Kulikov, 2019, 2020) (see table 2.1). The results of these studies have revealed that the amount of prevoicing in Voiced stops in voicing languages (e.g., Spanish and Russian) increases as speaking rate declines, and the amount of aspiration in Voiceless stops in aspirating languages follows the same pattern, as well. In both aspirating and voicing languages, however, stops with short lag VOT are not affected by the speaking rate. The range for short aspiration to differ in response to speech rate is narrower when compared to prevoicing and long aspiration. Upon a closer look, it could be suggested that the behaviour of unaspirated stops is dependent on the counterpart category (prevoiced stops in voicing languages and aspirated stops in aspirating languages). This reminds us of Keating's description (1984) of short lag stops in which they act as a swing category to maximize the phonological opposition. If we apply this principle to speech rate effects, short lag aspiration in Voiceless stops in voicing languages would be lengthened in normal or slow speech similar to the lengthening of the voice lead in Voiced stops to increase the distinction. On the other side, short lag aspiration in Voiceless stops in aspirating languages would be shortened (or disappear) to increase the distinction as well. These two scenarios usually do not occur, however, which supports the view that unaspirated stops in aspirating and voicing languages are unspecified.

Based on the privative feature models, stops with prevoicing and aspiration are specified for [voice] and [spread glottis], respectively, while stops with short lag are unspecified, a disparity which suggests a robust parallel between the phonetic aspects of voicing contrast and the active phonological features in the laryngeal system. That is, speaking rate affects specified stops, not unspecified stops, making correct predictions possible in languages with laryngeal overspecification. Beckman et. al (2011) demonstrate that stops in Swedish contrast prevoiced stops with aspirated stops in utterance-initial position. By using speaking rate effect as a tool to address laryngeal specification, Beckman et al. (2011) found that the amount of prevoicing and aspiration demonstrates an inverse relation to the speaking rate. Thus, it is reasonable to assume that speaking rate effect can be used to identify the active laryngeal features in the phonology.

Using the speaking rate paradigm as a source for revealing the phonological specifications specifying voicing contrast has attracted the attention of phoneticians and phonologists because it is simple and straightforward. The consistency of the speaking rate effect on the VOT of specified stops, but not the unspecified ones, in aspirating and voicing language, strengthens the usefulness of such a hypothesis. Nevertheless, such a claim requires an explanation. A number of theoretical justifications have been postulated to account for the speaking rate effect on VOT. One explanation is that the speakers' aim in lengthening/shortening the aspiration or prevoicing, but not the short lag (unspecified stops), is to maintain the contrast to be detectable by the listener (Kessinger and Blumstein, 1997). Yet another explanation, an articulatory justification, is that to lengthen the aspiration (short-lag unspecified stops), an active gesture is required, in this case, glottal opening (Kessinger and Blumstein, 1997). A more phonologically oriented account for speaking rate effect on VOT is proposed by Beckman et al. (2011) who stated:

> "One of the reasons that speakers slow down, of course, is to make their speech easier to understand. Thus, when the rate is slower, then speakers

are able to produce more of whatever acoustic property they are trying to produce. So, if the features specified in the phonology reflect speakers' goals, then we might expect a speaking rate to affect the phonetic cues associated with the features [voice] and [sg], but to not affect categories defined by their absence" (Beckman et al., 2011, p. 6).

The implementation of speech rate effects on prevoicing and aspiration in languages that contrast prevoiced and aspirated stops has not gained enough attention. Previous studies that employed speech rate effect to examine such a pattern in Swedish and Qatari Arabic are shown in table 2.1. The results revealed an intriguing pattern in which both prevoicing and aspiration are lengthened in response to slowing the speech rate. Operating under the phonological specification mentioned earlier, these results suggest that Voiceless stops are specified with [spread glottis] and Voiced stops specified with [voice].

Language type	language	stop	slow	fast	Reference
Aspirating	English	Aspirated	78	49	Allen and Miller (1999)
		Unaspirated	13	14	
	English	Aspirated	107	79	Kessinger and Blumstein (1997)
		Unaspirated	19	18	
Voicing	Russian	Prevoiced	125	78	Kulikov (2012)
		Unaspirated	14	13	
	Spanish	Prevoiced	69	46	Magloire and Green (1996)
		Unaspirated	12	19	
Asp-voice	Swedish	Prevoiced	107.9	78.5	Beckman et al. (2011)
		Aspirated	74.5	55.5	
	Qatari	Prevoiced	76	57	Kulikov (2020)
		Aspirated	55	43	

Table 2.1 Mean aspiration/prevoicing as a function of speech rate for various languages.

The present analysis adopts the speaking rate effect on VOT as a tool to reveal the phonological specifications underlying the voicing contrast. The present study will go further and investigate speech rate effect across contexts. Furthermore, the compatibility between the speaking rate hypothesis and other proposals, that have been presented in the literature as a method for examining the parallelism between the phonetic cues and the phonological features, is investigated in the present study. That is, the contribution of the [voice] stops and [spread glottis] stops in the phonological processes of final devoicing and voice assimilation is examined (specified stops trigger voice assimilation to the neighbouring sound). The results of final devoicing and regressive voicing assimilation processes are compared to the speaking rate effect findings to examine the consistency between the hypothesises made for the same goal, namely, using the phonetic cues to identify the active phonological features.

2.2.2 Final devoicing

Final devoicing refers to phonological process that results in a loss of contrast between Voiced and Voiceless stops in word-final environments and is also identified as *final laryngeal neutralization*, a phenomenon found in many languages. In the literature, there is a distinction between two views regarding laryngeal neutralization: *complete neutralization* which means the contrast is completely absent, and *incomplete neutralization*, in which the contrast is partially preserved. The former was proposed within formal theoretical frameworks based on categorical data whereby the latter was postulated within experimental frameworks relying on the continuous data (gradient) (Iverson and Salmons, 2007; Kirby, 2010).

The foundation for the debate regarding final devoicing is predicated on identifying the acoustic features of the stops in a context that enables the listener to detect the contrast and hence the lexical meaning. The results provided by the experimental accounts of final laryngeal neutralizations provide more insightful views regarding the nature of the process and its implication for the phonological specifications. Several attempts to describe final devoicing in both aspirating and voicing languages have been proposed. Based on the tightness between the phonetic cues and the phonological features proposed in the realm of laryngeal realism, it is reasonable to expect a difference in the behaviour of final devoicing in aspirating and voicing languages in which the former is associated with [spread glottis] while the latter is associated with [voice].

Final devoicing has often been regarded as *weakening or strengthening*. Iverson and Salmons (1999) proposed that the change of Voiced stop to voiceless in aspirating languages implicates a *strengthening* process (adding [spread glottis]) whereas the change of Voiced stops to voiceless in voicing languages implicates a *weakening* process (loss of [voice]). Acoustically, this means Voiced stops will be aspirated in final position in aspirating languages while they will be devoiced in voicing languages. Another classification in the literature of final devoicing suggests that final devoicing or laryngeal neutralization has three types: lenition which implies the removal of [voice] as in Dutch and Polish (Iverson and Salmons, 2006), fortition which implies adding [spread glottis] as in German (Iverson and Salmon, 2007) or adding [constricted glottis] as in Thai (Henderson, 1965).

To explain the reasons behind final devoicing, several proposals have been posited in the literature. A source of difficulty in maintaining voicing in final environments is that some languages tend to signal the edges of the phonological phrase by spreading the glottis in case of adding [spread glottis] or constricting the glottis in case of glottalization (adding [constricted glottis] (Blevins, 2004; Ohala, 1999 1997; Iverson and Salmons, 2007). In both cases, it is challenging for voicing to be initiated due to the absence of the required gesture for the vibration of the vocal cords which is glottal tension (Halle and Stevens, 1971; Ohala, 1983, Iverson and Salmons, 2007). A cue-based analysis of final devoicing raises crucial questions with regard to the differences between aspirating and voicing languages. As proposed in laryngeal realism, the distinction in final position is expected to be preserved at least some of the acoustic cues. It has been found that native English speakers employ the preceding vowel duration to signal the distinction whereas French speakers employs the release burst (Flege and Hillenbrand, 1987). Jessen (2001) proposed that non-basic correlates cover for the absence of the prevoicing to signal the contrast.

The present study adopts the feature addition or deletion processes proposed in Iverson and Salmons' study (1999) and tests this principle by examining the basic and non-basic acoustic correlates proposed in Jessen's study. Operating under these assumptions, in languages that contrast prevoiced stops with aspirated stops word-initially, the manifestation of final devoicing entails important theoretical and experimental implications for how the contrast is signalled. If Najdi Arabic shows glottis spreading to signal the final edge of the utterance, it is expected that Voiceless and Voiced stops will be aspirated. The manifestation of aspiration is crucial in this case in which the degree of aspiration in Voiced stops might be reflective of the addition/deletion process. The percentage of devoicing in the closure is another aspect that can be indicative of the voicing system of Najdi Arabic. The manifestation of other acoustic correlates beside voicing and aspiration might form a foundation for deciding which feature is specifying the Voiced and Voiceless stops in utterance-final position.

2.2.3 Voice assimilation across word boundaries

Voicing assimilation in the Generative Phonology approach was assumed to result in complete neutralisation, since the phonological rule precedes the phonetic rule; consequently it is a low-level process and should be treated outside the phonological component of grammar (Chomsky and Halle, 1968). However, several studies have noted a connection between the occurrence of regressive voicing assimilation (RVA) at word boundaries and the negative VOT which is a basic correlate of the feature [voice] (Kohler, 1984; Wissing and

Roux, 1995; Iverson and Salmons, 1995); specifically, word-initial stops trigger RVA to preceding word-final stops in different languages (Wissing, 1991; Katz, 1987; Wells, 1982). Because of the assumption that the occurrence of RVA might be phonetically conditioned (resulting from coarticulation and, consequently, C1 and C2 should not be phonetically identical), several studies have investigated RVA using direct quantitative evidence of acoustic data (Burton and Robblee, 1997; Barry and Teifour, 1999; Jansen, 2004; Kulikov, 2012). By investigating the acoustic characteristics of C1 and C2 at word boundaries, these studies have demonstrated that the assimilation of C1 to C2 is incomplete (gradient). Jansen (2004, 2007) postulated that stops trigger RVA only if they are actively (de)voiced which means that they are driven by an active articulatory event. Accordingly, RVA can be used to identify the active laryngeal features in the phonology that specify the voicing contrast such that specified stops trigger RVA, while unspecified ones do not.

Based on this assumption, a distinction is expected between aspirating and voicing languages in terms of RVA. For aspirating languages, it is expected that actively devoiced stops (aspirated stops specified by [spread glottis]) trigger a degree of devoicing for the preceding Voiced stop. In terms of voicing languages, on the other hand, actively voiced stops (prevoiced stops specified by [voice]) will trigger a degree of voicing in the preceding Voiceless stop. This is also in agreement with what has been proposed by Honeybone (2005) with regard to the differentiation of aspirating and voicing languages in their behaviour in clusters, in which the former usually show assimilation to voicelessness while the latter show assimilation to voicedness. It is noteworthy that the types of acoustic correlates expected to spread from the trigger to the target stop in RVA are the ones driven by an active articulatory gesture (Jansen, 2004). For instance, F0 lowering is proposed to be a correlate that accompanies Voiced stops. If it is driven by an active gesture (larynx lowering, expanding pharyngeal cavity, slacking the vocal cords), it is expected to spread from C2 to C1. This assumption might strengthen the argument for the specification of the stops and the activeness of the articulatory gesture they are associated with.

By considering the case of languages that contrast aspirated and prevoiced stops, investigating the occurrence of RVA can provide evidence for the reliability of the phonological overspecification in which both Voiced and Voiceless stops are expected to spread their (de)voicing characteristics to a preceding stop. Most of the studies that looked at RVA in languages with both prevoicing and aspiration did not experimentally examine this process. Instead, the analysis of categorical data was performed using the Optimality Theory framework (Ringen and Helgason, 2004; Petrova et al., 2006). It is also worth mentioning that it is expected for specified stops to resist changes caused by the phonetic context. If the two members of the cluster are presumed to be specified, the manifestation of the voicing assimilation is crucial to examine.

By adopting the coarticulation accounts proposed in the work of Jansen (2004, 2007) in the investigation of voicing assimilation in stop-stop clusters across word boundaries, the present study aims to reveal results with respect to three main issues: 1) complete/incomplete neutralisation, 2) passive/active (de)voicing, and 3) aspiration/voicing categorisation. For the first issue, incomplete neutralisation is indicative of interaction between the acoustic signal and the phonological representation, as proposed in integration models which emphasize the importance of phonetic concreteness in understanding the phonology. For the second issue, passive and active (de)voicing in stop-stop clusters are related to their ability to participate in assimilation in which specified stops are supposed to show active (de)voicing. In that regard, Jansen (2004) differentiated between passive and active (de)voicing and connected them with *regressive* and *progressive* voicing assimilation. That is, regressive voicing assimilation (from C2 to C1) indicates active (de)voicing whereas progressive voicing assimilation (from C1 to C2) indicates passive (de)voicing. The reason for this is that the former is an anticipatory effect of an active articulatory gestures while the latter is a carryover spontaneous spread form the trigger (C1) to the target (C2). For the third issue, as noted earlier, Voiceless stops are expected to trigger voicelessness in aspirating languages whereas Voiced stops are expected to trigger voicedness in voicing languages.

The basic assumption of laryngeal realism with respect to specification is explained through two-way interactions between the phonetic and phonological components. Therefore, employing phonetic concreteness in the study of voicing assimilation is a crucial tool to characterise the active features in the laryngeal systems in aspirating and voicing languages (Iverson and Salmons, 1995; Honeybone, 2002, 2005). Taking into account the different degree of assimilation reported in many studies, it is hard to draw any conclusion about extent to which the voicing targets of the assimilated stops will be affected. What might help in this case is considering numeric values of distinctive features proposed in the work of Beckman et al (2013). By combining the manifestations of Voiceless and Voiced stops across contexts with their behaviour in the voicing assimilation process, these findings might provide evidence for the numeric value that should be chosen for the distinctive features specifying the contrast.

Section 2.1 and section 2.2 above present the phonetic and phonological accounts of voicing contrast in stops in the theoretical models of voicing. The former reviewed the theoretical assumption of the traditional theories with regard to distinctive features, the levels of representation, the connection between phonetics and phonology, and the treatment of the aspirating/voicing distinction. The latter identified the basic assumptions proposed in the realm of laryngeal realism including the main assumption of this approach and the types of evidence employed to examine the interaction between the phonological features and the phonetic reality. Table 2.2 below sums up the key aspects of the aforementioned models in the previous two sections. The remaining part of the chapter focuses on two main issues: 1) the articulatory gestures in the production of voicing and aspiration and 2) the acoustic correlates of voicing contrast considering the aspirating/voicing distinction as well as the factors that are expected to affect the correlates.

Model	Main propositions	Featural system	Laryngeal features	Phonetic aspects
Jackobson et al. (1952, 1979, 1987)	Articulatory, auditory, and acoustic aspects of distinctive features.	binary	[voice], [tense]	[voice]: vocal folds vibration, [tense]: timing and energy of the acoustic components.
Chomsky and Halle (1968)	Phonetics and phonology disassociation.	binary	[+ voice], [-voice]	[voice]: vocal cord adjustments that enable voicing to occur. [tense]: gestures that inhibit voicing [heightened subglottal pressure]: extra energy for aspiration [glottal constriction] gestures that enhance voicing
Halle and Stevens (1971)	Accounting for laryngeal phonology through glottal events.	binary	[±spread glottis], [±constricted glottis], [±stiff vocal cords], [±slack vocal cords].	[+spread glottis, -constricted glottis, +stiff vocal cords, -slack vocal cords]: aspirated [-spread glottis, -constricted glottis, +stiff vocal cords, -slack vocal cords]: unaspirated [-spread glottis, -constricted glottis, -stiff vocal cords, +slack vocal cords]: prevoiced
Keating (1984)	A new level with major phonetic categories	binary	[+ voice], [-voice]	{vl.asp}: long lag aspiration {vl.unasp}: short lag aspiration {voice}: voicing lead
Kohler (1984)	The role of timing in laryngeal specification	binary	[+fortis], [-fortis]	[-fortis]: prevoicing in the closure [+fortis]: aspiration duration
Kingston and Diehl Model (1995)	Auditory explanation that assumes acoustic cues combining to mutually enhance the distinction	binary	[+voice], [-voice]	[voice] is associated with aspiration, low frequency properties, and C/V duration ratio
Jessen (1998, 2001)	Acoustic invariance and the hierarchal classification of correlates (basic/non-basic)	binary	[+voice], [+tense]	[+tense]: duration of stops properties and surrounding vowels [+voice]: closure voicing.
Laryngeal realism	The privativity of laryngeal features and the robust connection between features and acoustic signal	privative	[voice], [spread glottis]	[voice]: prevoicing [spread glottis]: aspiration

Table 2.2 Summary of theoretical models of voicing

2.3. The production of voicing

The production of a stop can be described as a sequence of articulatory movements that lead to several acoustic signals. This includes making a constriction in the oral cavity, maintaining the constriction for a period of time, and releasing the constriction, respectively (Stevens, 2000). Accordingly, the main component of this process is the airflow that escapes out of the lungs and passes through the larynx to the oral cavity. The place of the constriction distinguishes the categories of the stops and generates the acoustic difference between them in terms of temporal and spectral correlates. The acoustic and aerodynamic differences between stops with different places of articulation are caused by the changes in the vocal tract configuration.

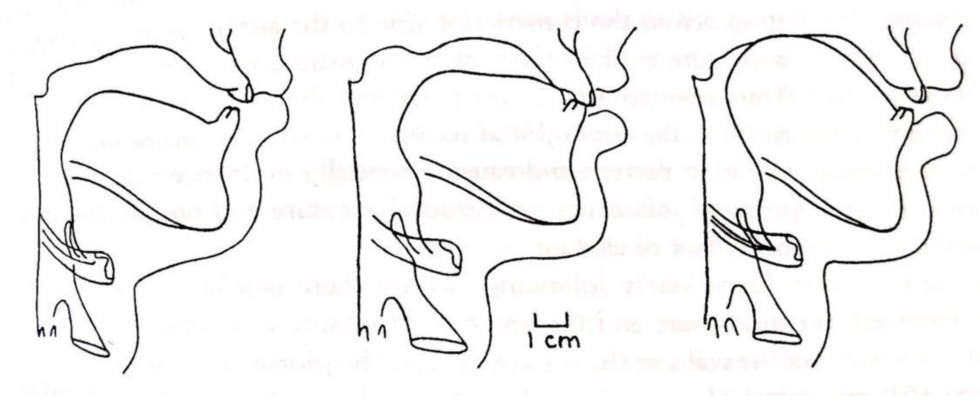


Figure 2.3 The production of stops based on the place of constriction (Adopted from Stevens, 2000, p. 325)

As shown in figure 2.3, the constriction is made by the lips in bilabial stops, by the tip of the tongue in alveolar stops, and by the back of the tongue in velar stops. Such articulatory movements might be simple in their structure, but they are complex in terms of their acoustic consequences. The closure period, if the vocal folds do not vibrate, is a complete silence in terms of acoustics, and it is a time to build up the pressure behind the constriction in the oral cavity in terms of aerodynamics. The soft palate rises during this period to prevent the air from escaping through the nasal cavity (Hayward, 2014). Building up the pressure leads to the release phase whereby the hold phase ends and is followed by a burst of noise that differs in its duration and intensity based on the place of the constriction and the timing of the voicing onset of the following vowel. Understanding the physical mechanism of voicing and aspiration production is key for the investigation of the acoustic properties of voicing. The importance of the articulatory and aerodynamic conditions for voicing initiation and sustainability stems from the fact that they provide the basic explanation behind the variations in the acoustic quality and quantity of voicing (Stevens, 2000; Jansen, 2004). Vocal fold vibration is the main articulatory event in the production of voicing. The manipulation of this event has important consequences in the production of Voiced and Voiceless stops in terms of aerodynamics, articulation, and hence acoustics. The vocal folds are contained within the larynx, which is basically formed of two cartilage structures: the cricoid and the thyroid (figure 2.4). Many studies have offered varied explanations for when, and to what extent, the vocal folds are likely to vibrate and thus satisfy the required aerodynamic and articulatory conditions for initiating voicing during the hold phase of a stop (van den Berg, 1958; Westbury, 1983; Ohala, 1997; Westbury and Keating, 1986). For voicing to be initiated, there must be a degree of tension in the vocal folds and a difference, about twice as large, between subglottal and supraglottal air pressure (Baer, 1975). Such a difference allows the air coming from the lungs to flow through the vocal folds, causing the vibration to occur.

The challenging part, however, is maintaining the vibration in the production of Voiced stops since the air pressure difference will be reduced due to the closure of the oral cavity and the increased air pressure above the glottis. Cavity enlargement is the main solution to maintain the transglottal air pressure difference, and thus facilitates voicing. Various articulatory mechanisms for cavity enlargement have been described in the literature including tongue root advancement, larynx lowering, soft palate raising, and pharyngeal expansion (Perkell, 1969; Westbury, 1983; Keating, 1984; Ohala, 2011; Sole, 2011). Differences in the articulatory nature of voicing initiation and maintenance mechanisms result in variations in the acoustic signals, and hence the voicing contrast patterns among languages occur.

In terms of aspiration, Lisker and Abramson (1964) defined aspiration as a delay in the onset of voicing for a following vowel. Kim (1970) proposed a different definition of aspiration and mentioned that it is associated with a spread glottis configuration in the larynx, resulting from a glottal opening. Simply put, after the stop's release, during the movement of vocal folds beginning to come together for the production of voicing of the following vowel, the air that is passing through the vocal folds prior to the glottal tension is perceived by the listener as aspiration (Iverson and Salmons, 1995). Kim (1970) states that "it seems to be safe to assume that aspiration is nothing but a function of the glottal opening at the time of release. This is to say that if a stop is *n* degree aspirated, it must have an *n* degree glottal opening at the time of release of the oral closure" (Kim 1970, p. 111). Later, various studies showed that aspiration is not only a consequence of a delay of voice onset (Cho and Ladefoged, 1999), but that it is a correlate contributing to the categorisation of voicing contrast among languages (Ladd and Schmid, 2018). Regarding another related issue, Kingston and Diehl (1995) emphasized the effect of word position and stress on the glottal opening in English. They state that "Glottal opening is simply smaller inter-vocalically than initially and before unstressed than before stressed vowels, and this smaller opening leads to shorter voicing lags (VOTs) and thus less aspiration" (Kingston and Diehl, 1995, p. 431). Thus, it could be assumed that [spread glottis] is active in the phonological system of English but only fully implemented in foot-initial position (Iverson and Salmons, 1995).

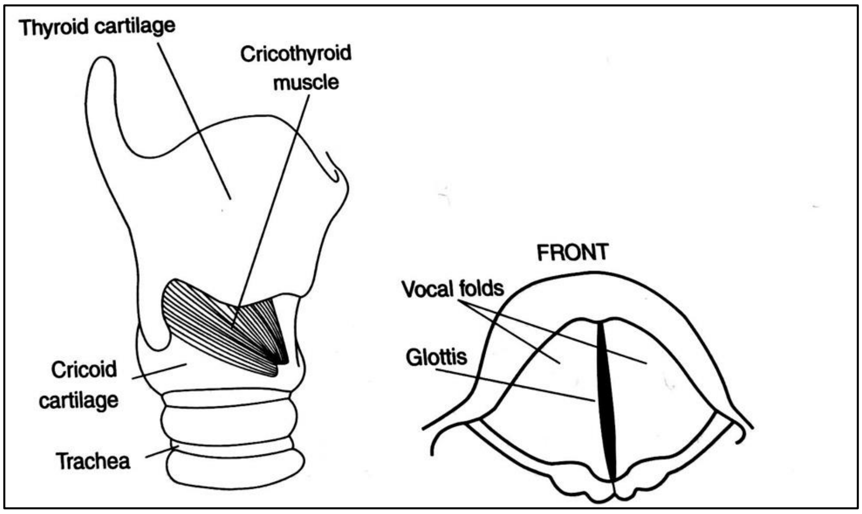


Figure 2.4. The structure of the larynx and the vocal folds (from Ashby and Maidment, 2005, p. 23)

An argument for the differentiation of languages with voice lead/short lag distinction and languages with short/long lag distinction is that voicing and aspiration have different articulatory mechanisms (Keating, 1990). This supports [spread glottis] as a feature specifying the laryngeal contrast in aspirating languages. It is worth mentioning that voicing is an extra event that in addition to the stop's closure which requires more articulatory adjustments. The differences between VOT patterning across phonetic contexts are based on the articulation and aerodynamic processes and how the stop maintains its characteristics. In initial and final positions, the required air pressure difference for vocal fold vibration diminishes early before the release in case of initial stops, and quickly after the onset of the closure in case of final stops (Westbury and Keating, 1986). The extra effort to maintain voicing in initial and final positions gives a robust indication of the behaviour of the language in terms of its classification as voicing or aspirating language.

2.4. Acoustic correlates of the voicing contrast

Previous studies focusing on the empirical basis of phonological features have documented that each phonemic contrast has several acoustic correlates. Perceptual studies, using synthetic stimuli, have shown that listeners' judgments vary as a result of manipulating these correlates (Abramson and Lisker, 1970; Williams, 1977; Flege and Hillenbrand, 1987; Chen, 1970). These studies also demonstrate the effectiveness of gradient data in enhancing our understanding of the voicing contrast (Kohler, 1984; Keating, 1984; Kingston and Diehl, 1995; Iverson and Salmons, 1995; Jessen, 2001). This section revisits the acoustic correlates of voicing contrast in stops across aspirating and voicing languages and discusses the association between these correlates and their phonological representations. Upon inspection of the theoretical models of voicing, Jessen's model (2001) provides a detailed acoustic description of the correlates that are expected to signal the distinction between Voiceless and Voiced stops in aspirating and voicing languages. This study adopts Jessen's model but also expands the scope to additional correlates proposed in the literature, to account for the contrast in various positions within the word and the utterance in Najdi Arabic. The very few previous accounts of Najdi Arabic discussed the presence of both prevoicing and aspiration in initial position but did not examine other contexts and processes. The participation of the privative features [voice] and [spread glottis] the phonological processes, along with their acoustic signal associations across various contexts, form the basic experimental analysis upon which the present study is built.

In the acoustic literature, the opposition between Voiced and Voiceless stops is cued by multiple temporal and spectral correlates. Some of the acoustic properties are found in the stop itself and some in the preceding or following vowels. Temporal correlates include Voice Onset Time (VOT), closure duration, voicing in the closure, preceding vowel duration, following vowel duration, and release burst duration. Spectral correlates include fundamental frequency (f0) and the frequency of the first formant (F1) at the onset of the following vowel and at the offset of the preceding vowel, the amplitude of the first and second harmonics (H1-H2) of the preceding vowel, and release burst intensity.

2.4.1. VOT and laryngeal features with respect to positional variations

VOT, defined as the time between the release of a stop and the onset of the laryngeal vibration, has been found to be a crucial correlate to distinguish Voiced and Voiceless stops in voicing and aspirating languages (Lisker and Abramson, 1964). Lisker and Abramson investigated stops in prevocalic **utterance-initial position** in eleven languages and concluded that VOT could be grouped into three categories: prevoiced stops produced with voicing lead (voicing begins before the release, negative VOT: -100 ms), unaspirated voiceless stops produced with short lag (voicing starts immediately after the release: 0-25 ms), and aspirated stops produced with long lag (voicing begins after the release, positive VOT: above 60). In their study, they found that languages could be grouped into three categories: 1) languages with a two-way contrast have prevoiced stops and unaspirated voiceless stops, as exemplified in *voicing* languages, such as Tamil and Spanish; 2) languages with a two-way contrast have unaspirated voiceless stops and aspirated stops, as noted in *aspirating* languages, such as American English and Cantonese; and 3) languages with a three-way contrast have the three VOT categories, such as in Thai and Eastern Armenian. Numerous studies have followed Lisker and Abramson's proposal investigating the VOT dimension and its usefulness in differentiating stop categories. The reason behind such focus on VOT in the previous research is to precisely address the phonetic characteristics of voicing contrast across languages and how they are represented in the laryngeal system. Lisker and Abramson's VOT categories were effective in addressing the contrast in many languages. For example, most Germanic languages contrast Voiceless unaspirated stops with aspirated stops (English: Flege, 1982; Smith 1978; German: Jessen, 1998), while many Romance and Slavic languages contrast prevoiced stops with unaspirated voiceless stops (Spanish: William, 1977; Portuguese: Jesus and Hall, 2010).

However, a growing number of studies have shown that VOT-patterns in some languages flout the traditional VOT typology. For example, it has been found that word-initial stops in sentence-medial position are realised with prevoicing in some aspirating languages (English: Docherty, 1992; Flege, 1982; Davidson, 2016) and Voiceless stops are produced in initial position with aspiration in some voicing languages (Najdi Arabic: Flege and Port, 1981; Al-Gamdi et al., 2019; Qatari Arabic: Kulikov, 2020; Turkish: Feizollahi, 2010; Norwegian: Ringen and Dommelen, 2013). Other studies have shown that Voiced stops in aspirating languages in medial position were realised with voicing in the hold phase (Beckman et al., 2013; Jessen 2001). Table 2.3 below presents mean VOT values for utterance-initial stops in aspirating languages, voicing languages, and asp-voice languages (show prevoicing and long lag aspiration).

Language type	language	Mean of aspiration	Mean of prevoicing	Reference
Aspirating	German	75	21	Jessen (1998)
	English	107	14	Kessinger and Blumstein (1997)
Voicing	Russian	20	-74	Ringen and Kulikov (2012)
	Dutch	19	-77	Van Alphen and Smits (2004)
	Spanish	14	-94	Dmitrieva et al. (2015)
Asp-voice	Swedish	75	-108	Beckman et al. (2011)
	Qatari	55	-76	Kulikov (2020)
	Turkish	41	-77	Ünal-Logacev et al. (2018)

Table 2.3. Mean VOT values for utterance-initial stops in some aspirating, voicing, and asp-voice languages.

To address such complexity, it is crucial to discuss the positional variation of VOT considering phonologically motivated voicing (active) and the phonetically motivated voicing (passive) and how these are implemented with respect to laryngeal feature specification. As noted earlier in chapter 1, two different views have been postulated about the laryngeal features that describe voicing contrast across languages. Some studies have argued that the voicing contrast is specified by the binary feature [±voice] in both aspirating and true voice languages (Keating, 1984, Kingston and Diehl, 1995; Wetzels and Mascaro, 2001). However, in privative models (Iverson and Salmons, 1995; Honeybone, 2005; Beckman et al., 2013; Jessen and Ringen, 2002), the binary feature [±voice] is argued to not sufficiently describe the voicing contrast in aspirating languages. The motivation behind this view is the assumption that the presence of active voicing (voicing lead) indicates the presence of the feature [voice], while the presence of active devoicing (long lag), indicates the presence of the feature [spread glottis]. Accordingly, in aspirating languages such as English and German, Voiceless stops are specified with the feature [spread glottis], while Voiced stops are unspecified [Ø]. On the other hand, in voicing languages such as Russian and Spanish, Voiced stops are specified with the feature [voice], while Voiceless stops are unspecified [Ø]. This pattern of laryngeal feature specification seems straightforward in **utterance-initial position**.

Voicing in **intervocalic position** is assumed to be complicated but still informative for determining what features are active in the laryngeal system. A number of studies have examined the acoustic features of VOT in intervocalic stops in aspirating and voicing languages. It has been found that Voiced stops (short lag) in aspirating languages were realised with passive voicing during the closure in intervocalic position (Keating, 1984;

Lisker, 1986; Beckman et al., 2013; Ringen and Kulikov, 2012; Jansen, 2004). However, the quality of the passive voicing is not acoustically equal to that of the active voicing which occurs in intervocalic Voiced stops in voicing languages in which the former, unlike the latter, shows an amplitude drop during the hold phase (Beckman et al., 2013; Ringen and Kulikov, 2012). These findings support the usefulness of privative models of features in describing the voicing contrast in aspirating and voicing languages. The primary observation, that in aspirating languages, intervocalic Voiced stops are passively voiced, has led to the conclusion that stops with no laryngeal specification are prone to phonetically motivated processes. Accordingly, Voiceless stops in voicing languages should analogously exhibit passive voicing in intervocalic position. Yet interestingly, it has been found that Voiceless stops (short lag) in some voicing languages, which are meant to be unspecified, do not always have passive voicing in this position such as Russian (Ringen and Kulikov, 2012; Beckman et al., 2013). It has also been found that unspecified stops in some aspirating languages like Mandarin and Danish (Jessen, 2001) do not undergo passive voicing inter-vocalically (Deterding and Nolan, 2007).

To counter these arguments, Beckman et al. (2013) proposed a numerical specification in which laryngeal features are assigned different numerical values (1 means inactive, 9 means highly active); these values are language-specific (see table 2.4). In their analysis, Beckman et al. note that, in voicing languages, the phonologically specified stops (Voiced) receive a high numerical value [9voice] whereby the phonologically unspecified stops (Voiceless) become [1voice]. Similarly, for aspirating languages, the Voiceless stops (long lag) will be specified with the feature [9spread glottis], whereby the phonologically unspecified Voiced stops (short lag) become [1spread glottis]. In this analysis, passive voicing is a phonetic process, and "such phonetic processes cannot change a numerically specified phonological feature" (Beckman et al., 2013, p. 280). Therefore, for passive voicing to occur in a stop, 1) the stop should lack a [voice] value and, 2) it should not receive a high numerical value of [spread glottis]. Accordingly, intervocalic Voiceless stops in Russian do not undergo passive voicing because of the [1voice] value. Similarly, they do not exhibit passive voicing in Icelandic due to the relatively high numerical value of the [5spread glottis], resulting from a language-specific feature (great glottal width).

Aspirating languages		
German type		
Fortis	[9sg] [∅ voice]	Passive voice cannot apply, glottal width too great
Lenis	[1sg] [∅ voice]	Passive voice applies, small glottal width, no numerical specification for [voice]
Icelandic type		
Fortis	[9sg] [∅ voice]	Passive voice cannot apply, glottal width too great
Lenis	[5sg] [∅ voice]	Passive voice cannot apply, glottal width too great

True voice languages		
Russian		
Fortis	[1voice] [∅ sg]	Passive voice cannot apply; phonetic rules do not change numerical specifications
Lenis	[9voice] [∅ sg]	Active voice

Table 2.4. Numerical values of laryngeal features (adopted from Beckman et al., 2013, p. 281)

The previous paragraphs provide a brief review of the laryngeal contrast in voicing and aspirating languages in utterance-initial and intervocalic stops, and it is seen that the distinction between voicing and aspirating languages is maintained in these two positions. The challenge is to establish how this distinction is applied in **utterance-final stops**, a phenomenon which receives little attention in the preceding literature (Jansen, 2004; Iverson and Salmons, 2011). Operating under the assumption there should be some degree of consistency across phonetic contexts in terms of voicing contrast in a language, that is, if Voiced stops in a voicing language are prevoiced in initial position and actively Voiced in intervocalic position, it could be further assumed that they would be actively Voiced in final position, as well. However, due to the possibility of laryngeal neutralisation or final devoicing processes, it has been found that stops in final position behave differently in voicing and aspirating languages. Furthermore, it has been shown that voicing and aspirating languages use other acoustic correlates instead of, or in addition to, VOT (e.g. French: release burst properties, English: preceding vowel duration) to mark the laryngeal contrast in final position (Flege and Hillenbrand, 1987; Mack, 1982).

Views on the behaviour of voicing and aspirating languages in marking the laryngeal contrast in final position raise important points. It seems reasonable and possible to use the numerical specification approach proposed by Beckman et al. (2013) to address the acoustic variation in VOT in stops across the phonetic context. That is, stops in initial position in stressed syllables will receive a higher numerical value than stops in unstressed syllables or final position. Hence, VOT in stops with low numerical value could be affected by the devoicing process. In such cases, other correlates would be used to signal the voicing distinction besides, or as alternatives to, VOT.

2.4.2 Languages with prevoicing and aspiration and their phonological specification

As highlighted in the previous section, some languages contrast prevoiced with aspirated stops in initial position. This unusual pattern has been found in a number of languages that are generally proposed to be either voicing or aspirating languages (Najdi Arabic: Flege and Port, 1981; Al-Gamdi et al., 2019; Swedish: Beckman et al., 2011; Qatari Arabic: Kulikov, 2020; Turkish: Ünal-Logacev et al., 2018; Norwegian: Ringen and Dommelen, 2013). The VOT pattern in these languages posits a challenge to the VOT categorisation of voiceless stops to short and long lag (Lisker and Abramson, 1964; Keating, 1984). This led some researchers to propose the possibility of intermediate VOT categories that take a middle position between short and long lag aspiration (Cho and Ladefoged, 1999). Cho and Ladefoged (1999) presented VOT values for a number of languages which show intermediate level and argued against the discreteness of VOT categories proposed in Keating' model (1984). Moreover, researchers differ in terms of determining the values of the intermediate VOT category. Some researchers followed the original categorisation of Lisker and Abramson (1964) and set the values of intermediate VOT between 25-60 ms such as Riney et al. (2007). Rafael et al. (1995) considered 30-50 ms to be the range of intermediate VOT. Keating (1984) proposed the {vl.unasp} stops to be 0-35 ms.

Most of the previous studies that looked at languages with two-way contrast which show prevoicing and aspiration focused only on initial stops and did not go further by looking at the manifestation of VOT in various contexts nor account for the behaviour of stops in phonological processes. In the following paragraphs, I present three studies that did expand the scope of the investigation of languages with such a pattern by considering various contexts or examining a phonological process.

Ringen and Dommelen (2013) investigated prevoicing and aspiration in Voiced and Voiceless stops in Norwegian in three contexts: utterance-initial, intervocalic, and utterance-final considering different places of articulation and different vowel contexts. The results for utterance initial stops showed that 63% of Voiceless stops were produced with long lag aspiration around 52 ms whereas 37% of Voiced stops were produced with prevoicing around -75 and 63% were produced with short lag aspiration 17 ms. The results for intervocalic stops showed that the majority of Voiced stops were produced with 93% voicing in the closure while the voiceless stops were produced with short lag aspiration 17 ms. For utterance-final, the results showed that the majority of Voiced stops were produced with 86% voicing in the closure and with release duration around 106 ms. For Voiceless stops, they were produced with aspiration around 173 ms. Based on the acoustic results, Ringen and Dommelen proposed that the voicing system of Norwegian showed features of both aspirating and voicing languages. The occurrence of long lag aspiration in initial position indicates the presence of [spread glottis] whereas the presence of robust voicing in intervocalic and final stops indicate the presence of [voice]. They also proposed that there is maybe an ongoing change in Norwegian's voicing system moving towards the aspirating languages because of the high percentage of devoicing in the Voiced stops in initial position.

Beckman et al. (2011) investigated the manifestation of prevoicing and aspiration in Swedish and draw some phonological conclusions based on speech rate effects. The target words included stops in initial position with different places of articulation and different vocalic contexts. They were produced at different rates (isolation, slow, fast). Voiced stops were produced with prevoicing: around -79 ms in fast rate and -107 in slow rate. The labial stops showed higher values of prevoicing than coronals and velars. As for Voiceless stops, they were realised with aspiration: around 56 ms in fast speech and 74.5 ms in slow speech. Stops in all places of articulation were produced with aspiration while the velars showed the highest values. A similar study that looked at prevoicing and aspiration in word-initial position in Qatari Arabic was carried out by Kulikov (2020). The results revealed that Voiced stops were realised with prevoicing and Voiceless stops with long lag aspiration; both were lengthened in response to slowing the speech rate (Kulikov's study is reviewed in Chapter 3 when discussing the modern dialects of Arabic). The study of Swedish (Beckman et al., 2011) and Qatari Arabic (Kulikov, 2020) are important for the present study because they looked at the speech rate effect as a tool to draw phonological implications. The VOT means for Swedish and Qatari stops are presented below.

	Voiceless		Voiced	
	slow	fast	slow	fast
Qatari	55	66	-76	-63
Swedish	74.5	55.8	-107.9	-78.5

Table 2.5 Mean VOT for Voiceless and Voiced stops in Qatari and Swedish in slow/fast speech rates.

The results of the aforementioned analysis of Norwegian, Swedish, and Qatari are very important for the present study. The study of Ringen and Dommelen showed VOT manifestation in various contexts which allow for proposing more accurate phonological analysis. The studies of Beckman et al. (2011) and Kulikov (2020) employed speech rate effect as a tool to test the phonological specification in prevoiced and aspirated stops.

The acoustic manifestation of VOT in languages that contrast prevoiced and aspirated stops is at the core of the phonetic and phonological analysis of the current study. The examination of this pattern might provide crucial findings for the interactions between distinctive features and phonetic reality. Most of the accounts for this pattern in the previous literature did not go beyond the acoustic measurement of prevoicing and aspiration in certain contexts. Little attention was paid to applying more in-depth analysis by looking at the behaviour of stops in various contexts in different phonological processes and interpreting the results in the light of the phonological specification. To examine specification in a laryngeal system, Jessen (1998, 2001) proposed the concept of contextual stability or relational invariance which emphasizes the consistency of robust voicing or long lag aspiration in the majority of contexts to be indicative of the presence of [voice] and [tense] (or [spread glottis]), respectively. Keating (1984) and Kohler (1984) proposed the rarity of such a pattern in languages with two-way contrast. Halle and Stevens (1971) discussed moderate aspiration only in languages with three-way contrast such as Korean. In this study, I aim to analyse prevoicing and aspiration in Najdi Arabic with a detailed acoustic and statistical investigation across different sources of variability taking into account various phonological processes.

2.4.3 VOT interactions with linguistic factors

It is well established that VOT is sensitive to various linguistic factors which are implemented because of language-specific characteristics (Docherty, 1992; Cho and Ladefoged, 1999). **The place of articulation** is one of the factors that evidently affects VOT cross-linguistically (Cho and Ladefoged, 1999).

In Voiceless stops, it has been found that there is a tendency for velar stops to have longer VOT (aspiration). Generally, the VOT increases as the place goes back in the mouth. These results have been found with some variability in aspirated stops in aspirating languages such as English (Caramazza et al., 1973; Klatt, 1975; Suomi, 1980; Docherty, 1992; Nearey and Rochet, 1994; Yao, 2009), and German (Jessen, 1998). It has also been found in voicing languages such as French (Abdelli-Beruh, 2009), Spanish (Rosner et al., 2000), Portuguese (Lousada et al., 2010), and Arabic (Yeni-Komsh et al., 1977). This pattern is also found in Swedish which is proposed to have prevoicing and aspiration (Helgason and Ringen, 2008). For Voiced stops, if they are realised with aspiration in aspirating languages, they follow the same pattern found in aspirated stops in which the duration of VOT increases as the place goes back in the mouth. Such a result has been found in English (Klatt, 1975; Zue, 1976; Docherty, 1992), and German (Jessen, 1998). In terms of prevoicing in voicing languages, the duration of pre-voicing, unlike aspiration, decreases as the place goes back in the mouth. Such a result is found in French (Jacques, 1984), Polish (Rojczyk, 2009), and Spanish (Lisker and Abramson, 1964).

Several proposals have been put forward in the literature to explain the reasons behind the impact of place of articulation on the duration of pre-voicing and aspiration. For aspiration, the size of the front cavity in case of /k/ leads to more air pressure behind the constriction which consequently means more time for the decreasing of the trans-glottal pressure necessary to initiate voicing for the following vowel (Cho and Ladefoged, 1999). For prevoicing, however, Ohala (1983) claimed that the expansion of the vocal tract needed to maintain the pressure difference between subglottal and supraglottal air is the main reason for longer prevoicing in bilabial stops than in alveolar and velar stops. That is, /b/ has more forward place of articulation hence a bigger oral cavity which leads to a greater chance of maintaining voicing through expansion of vocal tract. Such a claim has been found to be the case in both initial and intervocalic stops (Keating, 1984).

One of the contextual factors that affect VOT or the duration of pre-voicing and aspiration is **the quality of the following vowel**. Some early studies that focused on VOT concluded that there was no interaction between VOT and the quality of the following vowel (Lisker and Abramson, 1967). However, later, many studies showed a robust interaction between VOT and the quality of the following vowel (Klatt, 1975; Smith, 1978; Port and Rotunno, 1979; Weismer, 1979; Flege et al., 1998) in which that VOT is longer before high and tense vowels than before low vowels; this is found in Voiced and Voiceless stops in aspirating languages such as English (Docherty, 1992; Smith, 1987; Klatt, 1975) and German (Jessen, 1998). A similar vowel impact on VOT was found in some voicing languages including Arabic (Yeni-Komshian et al., 1977) and Italian (Esposito, 2002). This interaction is explained in the literature in terms of aerodynamics and physiology. It has been proposed that high vowels are articulatorily associated with a larger oral cavity, which is not the case for low vowels. Accordingly, for the voicing to be initiated, more time is needed for the supra-glottal air pressure to drop in the case of high vowels, which leads to more time for VOT. (Smith, 1978; Ohala 1983).

There are several other linguistic factors that were found to have an impact on VOT. **The syllable structure** is found to play a role in VOT variations such that VOT is longer in a monosyllabic word than in bi-syllabic words (Lisker and Abramson, 1967; Klatt, 1975; Flege et al., 1998). The speaking style in which the target words are produced has an impact on VOT values. **The speaking style** includes words in isolation, in a sentence, and in spontaneous speech. It has been found that the VOT of stops produced in isolated words is longer than those produced in a sentence context (Lisker and Abramson, 1967). Moreover, stops produced in words in spontaneous speech have shorter VOT than the ones produced in isolation (Theodore et al., 2009; Gosy, 2001). One of the factors that might affect VOT values is **stress**. It has been claimed that stops produced in a stressed syllable have longer VOT than those produced in an unstressed syllable for both Voiced and Voiceless stops (Klatt, 1975; Jacques, 1987; Kahn, 1976; Lavoie, 2001; Al-Tamimi and Khattab, 2018).

2.4.4. Other acoustic correlates of voicing contrast

Various studies focusing on the voicing contrast in stops across and within languages have shown that multiple acoustic correlates are employed to mark this contrast in aspirating and voicing languages. (Slis and Cohen, 1969; Ohde, 1984; Kingston and Diehl, 1995; Jessen, 2011). It has also been found that there is a trade-off relationship between some of these correlates whereby the weakness of a correlate is a by-product of the strength of the other and vice versa (Haggard et al., 1970; Coetzee et al., 2018). It is worth mentioning, too, that these correlates differ across phonetic contexts (Jessen, 2001; Lousada et al., 2010). Considering the aforementioned points, the following paragraphs offer a brief discussion of other acoustic correlates associated with the voicing contrast in the previous literature and how they mutually enhance the laryngeal contrast in aspirating and voicing languages.

Voiced stops are realised with weaker **release burst** than Voiceless stops in both aspirating and voicing languages (Halle et al., 1957; Zue, 1976). It has also been found that release burst duration is affected by phonetic context in that both Voiced and Voiceless stops have longer burst duration word-initially than in medial and final positions (Lousada et al., 2010; Lavoie, 2001). Among perceptual studies, it has been proposed that the role of the release burst is crucial in signalling the distinction between Voiced and Voiceless stops in final position, and some studies have shown a distinction between aspirating and voicing languages in that regard. For example, French listeners (a voicing language) rely on the release burst in their judgement of final stop voicing, whereas English listeners (an aspirating language) use the duration of the preceding vowel to signal the contrast. Different interpretations have been proposed to clarify the relationship between the voicing contrast and the release burst. It has been demonstrated that the strong and long release burst in Voiceless stops is a by-product of the relatively high pressure in articulating the constrictions (van Alphen and Smits, 2004).

The **preceding vowel duration** has been proposed in the literature as a correlate for voicing contrast in postvocalic and intervocalic stops whereby vowels before Voiced stops tend to be longer than vowels before Voiceless stops (Chen, 1970; Alghamdi, 1990; Kluender et al., 1988; Luce and Charles Luce, 1985). It has been found that this difference is not universal and some languages such as Arabic showed no significant effect of voicing on the preceding vowel duration (Mitleb, 1984; Munro, 1993; De Jong and Zawaydeh, 2002; Al-

Gamdi, 2013). De Jong and Zawaydeh (2002) posited that vowels in Arabic is different from English in that the vowel contrast in the former robustly depends on duration. Accordingly, they concluded that voicing effect is inhibited to retain the quantity contrast.

Some studies have shown that preceding vowel duration is opaque for several reasons. The correlation between **closure duration** and the duration of the preceding vowel shows an inverse relation in that if the former is long the latter is short, and vice versa (Kohler, 1984; Kluender et al., 1988). Thus, it could be suggested that the preceding vowel duration is dependent on the closure duration and, consequently, is not an important correlate for marking the voicing contrast if operating under the assumption that the duration of the closure is language-specific (Lehiste, 1970). However, the study of French and English final stops by Flege and Hillenbrand (1987), as cited above, found that English speakers, unlike French speakers, employ preceding vowel duration as a cue in their judgments, which is an indication that speakers of aspirating languages use temporal cues more than speakers of voicing languages (Jansen, 2004). On the other hand, it could be proposed that speakers of voicing languages employ spectral cues relatively more than speakers of aspirating languages, and there is reason to assume this is the case, specifically where other spectral correlates, such as **voicing in the closure**, are typically used by speakers of voicing languages but not by speakers of aspirating languages.

Closure duration has been found to mark the distinction between Voiceless and Voiced stops in aspirating and Voicing languages in which closure duration for Voiceless is longer than for Voiced stops. (Lisker, 1957; Kohler, 1984; Jacques, 1980). The results of various studies showed variation depending on the speaking style. The difference is relatively larger in isolated words than in words produced in sentences (Chen, 1970). It has also been proposed that the difference in closure duration is salient and can be perceived by listeners in perception studies (Slis and Cohen, 1969). Jessen (1998), however, found the opposite pattern for German in word-medial stops in which closure duration for Voiced was significantly longer than for Voiceless. This led Jessen to postulate that closure duration is an ambiguous correlate in aspirating languages due to the possibility of interaction between closure duration and aspiration duration. That is, closure duration might be reduced as an enhancement for the perception of aspiration duration (Jessen, 2001). This pattern has been found in Danish in which closure duration for Voiced is longer than for Voiceless stops (Hutters, 1985). Based on these results, it can be postulated that closure duration difference in voicing languages is more stable than in aspirating languages because of the lack of aspiration in the former and its presence in the latter (Jessen, 2001).

The **following vowel duration** has been considered in the literature as a correlate for the distinction between Voiced and Voiceless stops in prevocalic and intervocalic positions. It has been found that vowels after Voiced stops are longer than vowels after Voiceless stops in both aspirating and voicing languages. This distinction has been found in English (Allen and Miller, 1999), in French (Fischer-Jorgensen, 1968), in Ghamidi Arabic (Alghamdi, 1990), and in Lebanese Arabic (Al-Tamimi and Khattab, 2018). The explanation of this correlate is straightforward in terms of the trade-off relationships between correlates, as noted earlier. That is, it is expected that aspiration, in aspirated stops, causes reduction of the following vowel duration, which is not the case in Voiced stops where the release phase is shorter and weaker. Such a difference could be interpreted with respect to saliency and perceptibility in that the shortness of the following vowel duration makes aspiration more perceptible by the listener (Jessen, 2001). By proposing an interaction between aspiration and the following vowel duration, it is problematic to assume the same pattern to occur in voicing languages in which Voiceless stops are unaspirated. It could be assumed that in the case of voicing languages, the effect on the duration of the following vowel is not as evident as in aspirating languages. Such an assumption is supported by some studies such as the work of Iwata and Hirose (1976) on Mandarin.

It has been proposed for decades that voicing contrast in many languages is cued by a difference in the **fundamental frequency (F0)** of a following vowel (House and Fairbanks, 1953; Lehiste and Peterson, 1961). The distinction can generally be identified as the fundamental frequency being higher after Voiceless stops than after Voiced stops. This distinction has also been reported in the vowel preceding the stop but to a lesser extent (Jansen, 2004; Kingston and Diehl, 1995). The difference in F0 in response to voicing in most studies does not exceed 30 Hz for female speakers (Jansen, 2004). The connection between F0 and aspiration is proposed based on an aerodynamic explanation. It has been assumed that aspiration triggers the rise of F0 in the onset of the following vowel in which the high airflow associated with aspiration and glottal opening induces increased F0 values (Ohala, 1983; Stevens, 2000). This distinction is found in aspirating languages such as German (Jessen, 1998), Cantonese (Zee, 1980), English (Kingston and Diehl, 1995). It has been found also in languages with a three-way stop system such as Korean (Bang et al., 2018) and Madurese (Misnadin et al., 2015). Moreover, some studies show a connection between closure voicing and the lowering of F0 in the onset of the following vowel, as explained through the progressive impact of vocal folds' tension (during closure voicing production) on the F0 (Halle and Stevens, 1971; Hombert et al., 1979). Some other recent studies proposed that the lowering or raising of F0 is an active gesture and not a product of closure voicing or aspiration, respectively (Jessen and Roux, 2002; Chen, 2011). That is, the lowering/raising of F0 might be phonologically motivated. Kirby and Ladd (2016) investigated CF0 (consonant-induced F0) in two voicing languages: French and Italian. The results revealed that F0 was lowered following Voiced stops but was raised after Voiceless stops. This finding contradicts the prediction of laryngeal realism in which unspecified stops should not show any active gestures including F0 raising. Kirby and Ladd (2016) further argue that F0 lowering/raising should not be linked to the presence or absence of phonological specification and should be considered as a phonetic enhancement and a language-specific approach.

However, operating under the assumption of laryngeal realism, it could be inferred that the raising of F0 after aspirated stops is a by-product of aspiration (aspirating languages) whereas the lowering of F0 after prevoiced stops is a by-product of closure voicing. In terms of languages that are proposed to contrast prevoiced and aspirated stops such as Swedish (Beckman et al., 2011), Najdi Arabic (Flege and Port, 1981; Al-Gamdi et al., 2019), and Qatari Arabic (Kulikov, 2020), it could be suggested that the difference in F0 value might be greater (due to having both aspiration and closure voicing which are assumed to be the reasons for F0 raising and F0 lowering based on the articulatory justification, respectively) in comparison to aspirating languages (unaspirated/aspirated distinction) and to voicing languages (prevoiced/unaspirated distinction).

F1 frequency at the onset of the following vowel and the offset of the preceding vowel is another spectral parameter closely similar to F0, as a cue for the distinction between Voiced and Voiceless stops. Roughly speaking, F1 is also higher after aspirated stops than after unaspirated or Voiced stops (House and Fairbanks, 1953; Ohde, 1984; Kingston and Diehl, 1995; Jessen, 1998). The associations between **F0/F1 frequencies** and **closure voicing** are that they share the same function as low-frequency events (Kingston and Diehl, 1995; Jansen, 2004). In fact, it has been proposed that low frequency events (F0/F1 frequencies and closure voicing) robustly signal the voicing contrast in word-initial position more than in medial position (Kingston et al., 2008). Despite the fact the F0/F1 differentiation signalling voicing distinction is relatively small across many studies, they contribute to the identification of stops as Voiced or Voiceless in perception studies (Haggard et al., 1970; Kingston and Diehl, 1995). It is also noteworthy that the impact has been found to be much more noticeable in the onset of the following vowel than in the offset of the preceding vowel (Jansen, 2004). As for F1 lowering, this has been considered as a correlate for active [voice] in that the articulatory adjustments to maintain voicing, such as larynx lowering and vocal tract expansion, result in F1 onset lowering (Jessen, 2001).

One of the correlates that captures the voicing contrast in stops is a difference between the amplitude of the first and second harmonic (**H1-H2**). It is found that H1-H2 is higher after Voiceless stops than after Voiced stops (Jessen, 2001). The high value of H1-H2 is a consequence of breathy voice (Stevens and Hanson, 1995; Klatt and Klatt, 1990). It also has been proposed that there is a robust connection between aspiration and the rise of H1-H2 value as both are associated with the same gesture which is glottal opening (Jessen, 2001). This correlate has been found in aspirating languages such as English (Chapin Ringo, 1988) and German (Jessen, 1998). Surprisingly, some studies showed that H1-H2 is higher after Voiceless stops in voicing languages as well. This distinction has been found in Italian and French (Ni Chasaide and Gobl, 1993). However, the difference was much smaller than the difference found in German and English, suggesting that low H1-H2 is associated with [voice] in case of voicing languages and high H1-H2 is associated with [spread glottis] in case of aspirating languages.

This section provides a brief review of the temporal and spectral acoustic correlates that signal the distinction between Voiced and Voiceless stops with special focus on the difference between aspirating and voicing languages. The nature of the variation in acoustic correlates (continuous data) and their interactions with the phonetic contexts (positional and prosodic variation) might cause ambiguity in terms of characterising the laryngeal systems among languages. Accordingly, many attempts have been made to describe the acoustic correlates in voicing and aspirating languages by taking into consideration the phonological features that specify voicing contrast and how the interactions between the phonological representations and the acoustic signal shape the laryngeal systems among languages. One of the most important findings that has enhanced the usefulness of the acoustic correlates in identifying the phonetic and phonological aspects of voicing contrast is the mutuality of the acoustic correlates in marking the distinction (Kingston and Diehl, 1995; Jessen, 2001). The trade-off relationships between the correlates means that a specific correlate is stronger in a phonetic context than in any other correlate. The strength of the correlate is determined by whether it is a consequence of an active articulatory event which implies the presence of an active phonological feature (Jansen, 2004). Mutual enhancement means that non-basic correlates simultaneously strengthen the distinction in case the basic correlate is weak due to co-articulation or aerodynamic or articulatory factors (Kingston and Diehl, 1995; Jessen, 2001).

The importance of the non-basic correlates in terms of their role in signalling the phonological opposition between Voiced and Voiceless stops is that they enhance the distinction which accordingly means they are parts of the phonetic manifestation of the distinctive features specifying the contrast (Jessen, 2001; Kingston and Diehl, 1995). Therefore, acoustic correlates have been linked to the phonological features because they provide the information needed to confirm whether the stop is phonologically specified or not. Perception experiments have confirmed this claim by emphasizing the impact of the acoustic correlates on listeners' judgments.

2.5 Conclusion

In this chapter, I reviewed the literature on the phonetic and phonological aspects of voicing contrast. The theoretical models of voicing were presented, and their theoretical assumptions were discussed in the light of the key aspects of the current study. Phonological processes that involve voicing contrast were discussed, as well, including speech rate effect, final devoicing, and regressive voicing assimilation. The articulation process and the acoustic correlates that signal the distinction between Voiced and Voiceless stops were presented with a special focus on the phonetic manifestations and the phonological specifications of such a contrast among languages.

The modern dialects of Arabic show variation in their phonetic and phonological aspects of voicing contrast. This raises questions about the similarities and the differences between these dialects and what this variability means with regard to their laryngeal systems. The next chapter sheds light on the voicing contrast in stops in the modern dialects of Arabic.

Chapter 3. Voicing contrast among modern Arabic dialects

This chapter provides a concise description of voicing contrast in stops in modern dialects of Arabic, which share the same origin of Najdi Arabic, the target dialect of the present analysis. The importance of showing some aspects of voicing contrast in stops in modern Arabic dialects is associated with presenting how voicing contrast in Najdi is similar to or different from other dialects in terms of phonetic cues and phonological features. Despite expected variation in the findings of the wide range of studies that investigate voicing in Arabic with different methodologies, it is still possible to find relatively similar results. There are relatively few studies that focused on voicing contrast in the modern dialects of Arabic in comparison to other languages. Generally, most of these studies that looked at voicing contrast focused on the phonetic aspects only without discussing the implications for the phonology. Until recently, it had been proposed that Arabic shows a voicing language pattern by contrasting prevoiced and unaspirated voiceless stops (Yeni-Komshian et al., 1977; Khattab, 2002). Yet, a number of studies that focused on some modern Arabic dialects revealed a different pattern in which Voiced stops were realised with prevoicing while Voiceless stops were realised with aspiration that falls within the long lag range (Al-Gamdi et al., 2019; Alanazi, 2018; Kulikov, 2020). These findings raise questions about the phonetic and phonological aspects of voicing contrast in the modern dialects of Arabic. The structure of this chapter is as follows:

Section 3.1 reviews studies that investigated some of the phonetic aspects of voicing contrast in Arabic. Section 3.2 describes studies that adopted the laryngeal realism approach and explored the acoustic properties of voicing contrast and discusses their implications for the phonological representation. Section 3.3 reviews the status of emphatics in the discussions of specification and the type of voicing contrast (two-way vs three-way contrast). Section 3.4 presents Najdi Arabic with a special focus on its phonetic and phonological aspects as shown in the previous studies. Section 3.4, finally, sorts out the motivation and rationale for experimentally investigating voicing contrast in Najdi Arabic and how such an acoustic analysis reflects on theory.

3.1 Phonetic aspects of voicing contrast in modern Arabic dialects

Yeni-Komshian et al. (1977) investigated VOT in word-initial Voiced and Voiceless stops in modern Standard Arabic produced by native Lebanese Arabic speakers. The stops investigated include: /b/, /d/, /t/, /k/ (/g/ is not found in standard Arabic), followed by the three vowels that exist in modern standard Arabic (a, i, u). The results revealed that the stops showed a two-way voicing system with an opposition between prevoiced and unaspirated stops in the examined places of articulation. The mean VOT values for Voiced stops were -65 for /b/ and -56 of /d/ whereas in Voiceless stops they were 25 ms for /t/ and 28 ms for /k/. Regarding the effect of the following vowel on VOT, the results showed that prevoicing is shorter before /i/ than before /a/ and /u/. Aspiration (short-lag), on the other hand, is slightly longer before /i/ than before /a/ and /u/. Similar findings, with subtle differences, have been reported for Syrian Arabic which showed a word-initial position contrast between prevoiced and unaspirated voiceless stops (Radwan, 1996; Jesry, 1996). The work of Khattab (2002), later, unlike the aforementioned studies, investigated colloquial Lebanese Arabic but again found that, in the word-initial context, there was a contrast between prevoiced and unaspirated stops which is the pattern found in voicing languages. Mitleb (2001), however, found that in Jordanian Arabic the Voiced stop /d/ was realised as unaspirated voiceless stops with mean VOT around 10 ms whereas /t/ was realised with a longer aspiration but within the range of short lag as well /t/ = 37 ms. The results showed that VOT was affected by the quality of the following vowel (VOT is shorter before short vowels than before long vowels).

There are noticeable variations between the results of these studies and there are some issues, as well. One of the obvious issues is that the participants in some of these studies (Yeni-Komshian et al., 1977; Radwan, 1996; Jesry, 1996) produced the target words or sentences using Standard Arabic not their colloquial dialect although the results were generalised as phonetic aspects of the colloquial dialect. This is problematic considering that there might be differences between standard Arabic, which is only used in formal contexts, and colloquial varieties spoken in daily life. There are other issues related to the participants chosen for the investigation. For instance, the participants in Mitleb's study (2001) were university students in an English department which increased the possibility of second language effects. Another issue regarding the previous studies is that the number of participants was relatively not sufficient.

Giving that the studies of voicing contrast in modern dialects of Arabic are even fewer than other languages, studies that discuss VOT in Saudi Arabic are fewer and rare. Alghamdi (1999) investigated the durational correlates that signal the distinction between Voiced and

Voiceless stops in Ghamidi dialect, a dialect spoken in the southwestern region of Saudi Arabia. The stops were investigated in initial, intervocalic, and final positions within the word. The test words were real words, embedded in a phrase, and were produced by the participants in their colloquial dialect. The results showed that Voiceless stops in word-initial position were produced slightly aspirated, but longer than the short-lag range found in voicing languages (/t/: 38 ms, /k/: 50 ms), and they were unaspirated in intervocalic (/t/: 21 ms, /k/: 25 ms) and word-final (/t/: 26 ms, /k/: 27 ms) positions. For Voiced stops, the results showed that they were produced with full voicing during the closure in all contexts (98% of initial stops, 100% in intervocalic stops, 92 % of final stops). The results of Alghamdi's study, unlike other studies, found aspiration in case of word-initial stops. The inclusion of intervocalic and final positions is another strength in Alghamdi's work so that the results might give general indications about the laryngeal system in Ghamidi dialect. There are some issues in the study however. The notion of passive and active voicing was not considered in the analysis despite the possibility of them occurring due to the phonetic context. That is, the initial stops were preceded by a vowel /a/ and the final stops were followed by a glide /w/ in the carrier phrase (ʔamla ...... wasakat) 'he dictated' which make them prone to passive voicing. Also, the study focused only on temporal correlates.

Another Saudi dialect investigated by Flege and Port (1981) is the Najdi dialect. They investigated Voiced and Voiceless stops in initial and final positions in the context of /a/ vowel, and the test words were embedded in a carrier phrase. The results showed that Voiceless stops were slightly aspirated in word-initial position (/t/: 37 ms, /k/: 52 ms), but no results were presented regarding stops in final contexts. Voiced stops were produced with prevoicing (fully Voiced) in word-initial position (/b/: 85 ms, /d/: 82 ms, /g/: 75 ms) and with voicing in the hold phase that covered half of the closure duration. Flege and Port's study was the first that focused on voicing contrast in Najdi Arabic. The occurrence of aspiration in the production of the participants' speech could be a result of the effect of their English background (they were university students in USA at the time of participation).

More recently, Alanazi (2018) investigated the voicing contrast in North Saudi Arabic in word initial stops. The initial stops were investigated at three places of articulation in the context of /a/, /i/, and /u/, and the test words were embedded in a carrier phrase (ʔna ʔgu:l ….. wa ʔru:h elbe:t) 'I say ….. and go home'. The results for the monolingual speakers of North Saudi Arabic showed that they contrasted prevoiced stops with aspirated stops (/t/: 58 ms, /k/: 72 ms, /b/: -77 ms, /d/: -81 ms, /g/: -78 ms). It could be noticed that the presence of aspiration and prevoicing has been found in several Saudi dialects including Najdi Arabic (Flege and

Port, 1981; AL-Gamdi et al., 2019), Ghamidi dialect (Alghamdi, 1999), and North Saudi dialect (Alanazi, 2018).

The observed variation among Arabic dialects in terms of voicing contrast are worthy of further investigation to consider all interactions at the phonetic level as well as their implications for phonological representations. One of the obvious issues in the studies that looked at voicing contrast in Arabic dialects is that more attention has been paid to VOT, rather than other spectral and temporal acoustic correlates which may enhance the distinction between Voiced and Voiceless stops. The next section discusses other acoustic correlates besides VOT found and discussed in the previous studies of Arabic dialects.

The majority of the previous studies of voicing contrast in stops in Arabic dialects focused on the durational acoustic correlates including VOT, preceding vowel duration, following vowel duration, and closure duration. Flege and Port (1981) in their work on Najdi dialect found that **preceding vowel duration** did not significantly mark the distinction between Voiced and Voiceless stops. **The closure duration**, however, did slightly differ with 10 ms more for Voiceless stops. Alghamdi (1990) found that **the preceding vowel duration** is significantly longer before Voiced stops than before Voiceless stops in intervocalic and final positions. He also found that **the following vowel** duration was significantly longer after Voiced stops than after Voiceless stops in initial and final positions. Mitleb (1984), however, found no significant effect of voicing on **the preceding vowel duration** in Jordanian Arabic. De Jong and Zawaydeh (2002), with more in-depth analysis, investigated the intrinsic and extrinsic factors that affect the durations of vowels in Arabic. They concluded that the duration of vowels in Arabic before Voiced and Voiceless stops, unlike English, is associated to the phonemic length and not affected by voicing.

3.2 The Laryngeal Realism approach in modern Arabic dialects

The Laryngeal Realism approach highlights the importance of the phonetic reality in exploring the phonological representation. A growing number of studies started to examine the theoretical proposals of laryngeal realism in various languages with different phonetic patterns. Arabic dialects received little attention so far in the literature investigating the manifestation of voicing contrast with respect to the diagnostics proposed in laryngeal realism. In the following paragraphs, I present three studies that investigated voicing contrast in stops in three different Arabic dialects by adopting the laryngeal realism approach; they are: Lebanese Arabic (Al-Tamimi and Khattab, 2018), Najdi Arabic (Al-Gamdi et al., 2019), and Qatari Arabic (Kulikov, 2020).

A study that focused on both the temporal and spectral acoustic correlates and their phonological implications was conducted very recently by Al-Tamimi and Khattab (2018). They examined the interaction between two types of contrast in Lebanese Arabic: singleton-geminate contrast and voicing contrast in order to test voicing patterns crossed with phonological length in word-medial intervocalic position. The temporal correlates included voicing in the closure, closure duration, the preceding and following vowel duration, **the release burst duration**, and aspiration duration. The spectral acoustic correlates included **F0 and F1** at the onset of the following vowel and the offset of the preceding vowel, and the difference between the first and second harmonics **(H1-H2)** in the offset of the preceding vowel and the onset of the following vowel. The results revealed that VOT in Lebanese Arabic falls within the voicing languages category by contrasting prevoiced stops and unaspirated voiceless stops in both singleton and geminate stops. The results surprisingly showed that closure duration is the most important correlate that marks the distinction in the four-way contrast in both voicing and gemination. The correlates found to be significant in the distinction between Voiced and Voiceless singleton stops included voicing in the hold phase and the preceding and following vowel duration. The results also revealed a tendency to decrease for F0/F1 at the onset of the following vowel and at the offset of the preceding vowel, and also the difference in the amplitude between the first and second harmonics (H1-H2), in the context of Voiced stops.

As for the phonological implications, the authors adopted the numerical values of phonetic distinctive features and the privativity of the representational system proposed in the work of Beckman et al. (2013). The patterns of voicing in the closure of singleton and geminate stops showed variation in which Voiced stops showed passive devoicing in the two categories while Voiceless stops showed a moderate degree of passive voicing. For the release phase the geminate stops in both voicing categories showed a minor feature of spread glottis compared to the singleton stops. Based on these findings, the authors concluded that Voiceless singletons are associated with [3 voice], [Ø] [2tense], Voiced singletons are associated with [8 voice], [Ø] [0tense], Voiceless geminates are associated with [1 voice], [Ø spread glottis] [4tense], and Voiced geminates are associated with [6 voice], [Ø spread glottis] [3tense].

Al-Tamimi and Khattab's work is a pioneer attempt in applying the laryngeal realism approach to modern Arabic dialects. One of the main contributions of this study is to investigate passive and active voicing in intervocalic singleton and geminate stops and discuss their acoustic details in the light of phonological specification. It can be noticed that in Lebanese Arabic Voiceless stops acted as unspecified segments by showing passive voicing in their closure and by being unaspirated, in contrast to Voiced stops which showed robust voicing but with various degrees of passive devoicing during their closure. Although the numeric values of the features seem somehow impressionistic, they allow for hierarchical differentiations within voicing languages in terms of the robustness of voicing in the closure. Some voicing languages such as Russian showed robust voicing in the closure in almost all tokens as reported in various studies (Ringen and Kulikov, 2012; Kulikov, 2012; Beckman et al., 2013).

Another study that adopted the laryngeal realism approach was conducted by Al-Gamdi et al., (2019). They investigated the acoustic properties of voicing contrast in Najdi Arabic word-initial stops. Temporal and spectral correlates have been investigated in stops in different places of articulation /b, d, g, t, k/ followed by the eight vowels that exist in Najdi Arabic (/a:/, /a/, /i:/, /i/, /u:/, /u/, /e:/, /o:/). The test words were produced in a carrier phrase /ʔana ʔagu:l …../ 'I say…..'. The results showed that Voiceless stops were realised with heavy aspiration whereas Voiced stops were realised with prevoicing. Mean aspiration for Voiceless stops was 76.2 ms and -75.1 ms for prevoicing in Voiced stops. In terms of the effect of place of articulation, /b/ showed the highest value for prevoicing with -82.2 ms while /k/ showed the highest value for aspiration with 83.9 ms. For closure duration, the results showed that closure duration in Voiceless stops tended to be longer than for Voiced stops with overlap between the two categories. The results showed that F0 onset was a robust acoustic correlate with 25 Hz difference between the two categories: 140 Hz for Voiced and 165 Hz for Voiceless. Based on the acoustic results, the authors proposed that Najdi Arabic shows features of both voicing and aspirating languages which implies that the distinction is overspecified with two features [spread glottis] and [voice].

Al-Gamdi et al.'s work was the starting point for the current study. The findings revealed that both aspiration and prevoicing were employed by Najdi Arabic speakers to mark the phonological opposition between Voiced and Voiceless stops. The present study aims to identify the manifestations of prevoicing and aspiration across various sources of variability to test the contextual stability of these properties following the model of Jessen (1998, 2001). Furthermore, the current work goes further by examining the behaviour of stops in phonological processes which will form the basic foundation for determining the features that specify voicing contrast in Najdi Arabic.

One of the studies that considered the interaction and tightness between phonetic cues and phonological laryngeal features is Kulikov (2020), which focused on voicing contrast in Qatari Arabic in word-initial position. The test words were embedded in a carrier phrase and produced by the participants at slow and fast speech rates. The reason for testing the speech rate effect is to evaluate VOT behaviour in these two conditions and what it implies for phonological specifications. The results showed that Qatari Arabic contrast prevoiced and aspirated stops and that VOT in both categories was affected by speech rate. Accordingly, Kulikov concluded that there is overspecification with two phonological features [voice] and [spread glottis] that specify the distinction. Additionally, several spectral cues have been investigated, including spectral centre of gravity (SCG) of the burst, the fundamental frequency F0 at the onset of the following vowel, and the F1 frequency at the onset of the following vowel. The results showed significantly low values of all the examined spectral cues in the context of Voiced stops and high values in the context of Voiceless stops.

Kulikov's study revealed crucial findings that reflect on the theoretical debate in terms of the interactions between phonetics and phonology. It is also worth mentioning that the origins of the Qatari population go back to some tribes that arrived to Qatar from Saudi Arabia (Alsudairi and Abusharaf, 2015) which might indicate that Saudi Arabic and Gulf Arabic, in general, differ from the rest of the dialects in terms of the phonetic manifestations and the phonological features describing voicing contrast. This is also supported by the results of voicing contrast in the other Saudi dialects including Najdi Arabic (Al-Gamdi et al. 2019), North Saudi dialect (Alanazi, 2018), and Ghamidi dialect (Alghamdi, 1999). Although Kulikov's work adopted the laryngeal realism approach, the only type of evidence used in the study is speech rate effect and in word-initial stops only. Various types of evidence have been proposed in the literature of laryngeal realism that could be used to evaluate the laryngeal systems in order to address the phonological specifications and their correspondent phonetic cues. This gap will be filled in the present analysis by considering all types of evidence that have been employed in the literature of laryngeal realism.

The previous two sections (3.1 and 3.2) provide a description of the acoustic correlates reported in the literature regarding voicing contrast in modern Arabic dialects and the phonological aspects that based on the laryngeal realism approach. The temporal and spectral acoustic correlates have been discussed with more attention to temporal correlates, due to the rarity of studies that focus on spectral correlates. More research on the laryngeal system of Arabic dialect is crucial for various reasons. The small number of studies is an obvious reason. More importantly, variation in acoustic correlates across modern Arabic dialects entails questions about the phonological features active in the phonological systems of each dialect. In fact, to accurately investigate the laryngeal system of a language, it is important to consider both the phonetic cues and the phonological representations by looking at the parallelism between the two levels in various contexts with a focus on phonological processes of laryngeal features. This is the approach that the present analysis attempts to pursue.

One of the topics related to the phonological features that specify voicing contrast in the modern Arabic dialects is whether their voicing systems show two-way or three-way contrast. This issue is discussed in the next section.

3.3 Two-way vs three-way contrast

It is well-known that most, if not all, dialects of Arabic have an emphatic / plain consonant distinction. In terms of stops, some researchers discuss the distinction between emphatic and plain stops separately from the discussion of voicing contrast, considering that VOT is not the primary correlate for this contrast (Al-Masri and Jongman, 2004; Khattab et., al 2006; Zawaydeh and de Jong 2011, Heselwood and Maghrabi, 2015; Al-Tamimi, 2017). Other studies consider VOT as correlate that differentiate between Voiceless stops, Voiced stops, and emphatics leading to the conclusion that Arabic dialects can be divided into two types: 1) dialects with two-way contrast (unaspirated voiceless stops including both plain and emphatics with short lag vs prevoiced stops), 2) dialects with three-way contrast (unaspirated Voiceless stops which include emphatics only vs aspirated Voiceless stops vs prevoiced Voiced stops) (Bellem, 2014). When accounting for the specification and the features in the phonological representation, there are two options: the single feature [voice] specifies the contrast in case of the first type; and the two features [voice] and [spread glottis] specify the voicing contrast in case of the second type. Based on this classification, Saudi Arabic might fit the description of the first type in that it has three-way contrast between unaspirated Voiceless, aspirated Voiceless, and Voiced stops (Bellem, 2014).

Other studies that looked into the acoustic correlates of emphatics in Arabic proposed that VOT is not the primary acoustic correlate in the distinction between plain vs emphatic stops, and consider the lowering of F2 in the adjacent vowels to be the main correlate for this opposition (Al-Masri and Jongman, 2004; Khattab et., al 2006; Zawaydeh and de Jong 2011, Al-Tamimi 2017). Accordingly, Arabic is different from languages which have three-way contrast that is predicated only on the presence of prevoicing and aspiration such as Thai (Kessinger and Blumstein 1997) and Eastern American (Lisker and Abramson 1964). Based on this view, it is justifiable to propose that the voicing contrast in Arabic dialects that contrast aspirated and prevoiced stops (such as Najdi Arabic: Al-Gamdi et al., 2019; Northern Saudi: Alanazi, 2018) has a two-way contrast system. Moreover, the notable variation in the values of VOT reported in studies that discussed acoustic correlates of emphatic stops raises questions about the reliability of VOT as a correlate for emphatic/plain distinction.

3.4 Najdi Arabic

Najdi Arabic is spoken in the middle region of Saudi Arabia, traditionally called "Najd province" (Al-Sweel, 1987). Najdi Arabic is the best-known dialect in Saudi Arabia because it is the dialect spoken in the capital city and used by the royal family. In fact, Najd, in terms of geography, refers to a large region that extends from Yemen in the south to Jordan in the north, and from Hijaz in the west to Ahsa to the east (Figure 3.1) (Abboud, 1979; Al-Sweel, 1987; Ingham, 1994). Ingham (1994) states that there are different sub-dialects distributed in the Najd area. These sub-dialects, according to Ingham, differ in terms of morphology but are highly similar in terms of phonology. These dialects were categorised in Ingham's study as the follows:

"1- Central Najdi. The dialects of Central Najd as described above and the central Bedouin tribes also the 'Anizah of the Syrian desert.

2- Northern Najdi. The dialect of Jabal Shammar and of the Shammar tribes of Northern Najd and the Jazirah.

3- Mixed Northern-Central. The dialect of Qasim and of the Dhafir tribe.

4- Southern. The dialect of Najrān and the Ghatān tribe of the south and of the Ä1 Murrah and 'Ājmān tribes of the east" (Ingham, 1994, p. 5).

The main focus of this study is on the dialect spoken in Riyadh, which belongs to the sedentary population in Central Najdi based on Ingham's classification. It has been proposed that all different dialects of Najdi Arabic are phonologically similar, but they differ in terms of morphology (Ingham 1994).

Najdi Arabic has voicing contrast in stops and fricatives and it occurs in word-initial, word-medial and word-final contexts. The contrast occurs in alveolar and velar stops but not in bilabial stops because of the absence of /p/ in the Najdi Arabic inventory. Najdi Arabic has some features and phonemes that do not exist in Classical Arabic such as the voiced velar stop /g/, the voiceless affricate /ts/, the mid front vowel /eː/, and the mid back vowel /oː/ (Ingham, 1994). Initial consonant clusters are another feature found in Najdi that is not found in Classical Arabic (Alghmaiz, 2013). Tables 3.1 below shows the inventory of consonants of Najdi Arabic as presented in the work of Ingham (1994). Najdi Arabic has five vowels: /i/, /eː/, /a/, /u/, /oː/. The vowels /i/, /a/, and /u/ have a phonemic length contrast.

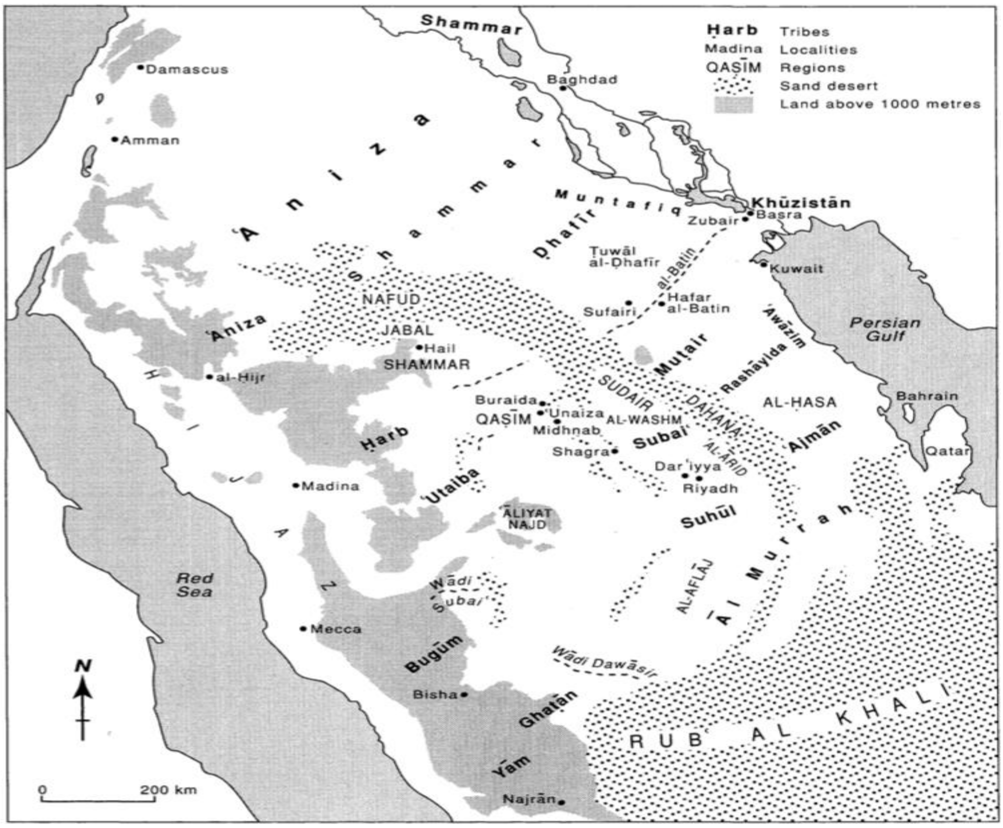


Figure 3.1. The map of Najd and surrounding regions (Adopted from Ingham, 1994, p. xvii).

	Bilabial	labiodental	Dental	Alveolar	postalveolar	palatal	Velar	Uvular	Pharyngeal	Glottal
stop	b			t d tˤ			k g			ʔ
Nasal	m			n						
Affricate					dʒ					
Fricative		f	θ ð ðˤ	s z sˤ	ʃ		x ɣ		ħ ʕ	h
Approximant				r		j				
Lateral approximant				l						
Glide	w									

Table 3.1. The phonemic inventory of consonants in Najdi Arabic.

Previous research on Najdi focused on various aspects including affrication and syllable structure (Johnstone, 1967; Ingham, 1994), vowel quality (Lehn, 1967; Asweel, 1990), the stress pattern in verbs (Abboud, 1979; Prochazka, 1988), syllable structure (Alezetes, 2007) and initial consonant clusters (Alghmaiz, 2013). The studies that focused on voicing contrast were carried out by Flege and Port (1981) in the field of second language acquisition and Al-Gamdi et al. (2019) which looked at temporal and spectral acoustic correlates in word-initial stops. As for other acoustic features, and the interaction between phonetics and phonology, it is evident that Najdi Arabic received little attention in the previous literature despite the dialect showing some features that are important to look at.

3.4 The rationale of the current study.

The present study, building on the work of Al-Gamdi et al. (2019), marks the first attempt to examine the phonetic and phonological aspects of voicing contrast in Najdi Arabic, so its importance is implicit. The present study investigates voicing contrast in stops in Najdi Arabic by taking into consideration various positions within the word and the sentence. Based on what has been proposed in the pre-existing literature, the positions examined in the present study include all the phonetic contexts expected to differentiate voicing and aspirating languages. Furthermore, processes that involve the properties of voicing contrast and their phonological representations are investigated in the present study, including speech rate effect, passive and active voicing, final devoicing or final laryngeal neutralization, and regressive voicing assimilation. The importance of these processes lies in the established correlation determined in the previous literature in the laryngeal realism approach between their phonetic aspects and the active phonological features in both aspirating and voicing languages. Since it is assumed that stops in Najdi Arabic contrast between prevoicing and aspiration and possess features of both voicing and aspirating languages, the outcomes of the present study add crucial insights to the field of voicing contrast in stops among languages and enrich the voicing models in the literature which have been predicated on a small number of languages.

Within the literature that posits a relationship between phonetics and phonology, the accuracy in characterising the nature of the interaction between phonetics, which is a physical science, and phonology, an abstract one, remains opaque. Two major approaches have been postulated in the literature. The first approach proposes the independence of the two domains and argues for an interface component which converts phonological entities into phonetic details (Keating, 1984). The second approach assumes integration between the two domains whereby they interact with each other consistently (Kohler, 1984; Jessen, 1998). The present study contributes to testing the effectiveness of acoustic correlates in explaining the empirical and theoretical aspects of voicing contrast. The major line of inquiry in this study can be described as follows: to what extent do the acoustic correlates of voicing contrast afford indications of the activeness or specificity of the phonological features, and how are the phonological features implemented in the acoustic signals? To pose a challenge for such approaches, the target dialect chosen in the present study shows features of both voicing and aspirating languages (Flege and Port, 1981; Al-Gamdi et al., 2019), and the designated phonetic contexts to be investigated are various and expected to be markedly affected by different factors. Therefore, delimiting the phonological representation and the phonetic realisation of voicing contrast are expected to be problematic in these cases. By pursuing the aims of the present study, the approaches that characterise the interactions between phonetics and phonology will be tested in a dialect that shows an uncommon phenomenon by contrasting prevoiced and aspirated stops. Many studies attempted to build models and systemize the interaction between phonological representation and the phonetic aspects of voicing contrast, but they might limit the investigations to laryngeal systems that show a typical pattern with regard to voicing/aspirating classification. While this is not a unworthy approach and crucial in terms of identifying the typical patterns among the two categories (voicing/aspirating), investigating languages that show features of both voicing and aspirating languages might enrich the theory and afford insightful contributions into the connection between phonetics and phonology. More focus is needed on different languages and dialects which is the main contribution of the present study.

Laryngeal realism theory employs three types of evidence to address laryngeal features and phonological specifications: 1) the phonetic cues of the segment, 2) speech rate effect on prevoicing and aspiration, and 3) stop behaviour in phonological processes including final devoicing and regressive voicing assimilation. Most of the previous studies discuss one or two of these types. Given that all three types were proposed to be effective in various studies, the present analysis uses the full range of available evidence and considers the linguistic factors that affect the acoustic correlates. Conducting this investigation will enable a clear description of the acoustic details of voicing contrast, including the temporal and spectral correlates, in addition to their implications for phonological representations. In this dissertation, I aim to answer the following questions:

1. What are the acoustic correlates of stop-voicing contrast in Najdi Arabic and how are they implemented across the following phonetic contexts: utterance-initial, utterance-medial intervocalic, utterance-final, and across-word-boundary clusters.
2. Employing the laryngeal realism approach, how does the voicing system of Najdi Arabic behave in terms of the following processes: speech-rate effect on the acoustic correlates of stops across the examined phonetic contexts, the acoustic activeness of voicing/devoicing of stops across the examined phonetic contexts, and regressive voicing assimilation in across-word-boundary clusters.
3. In light of the results derived from the preceding inquiry, is Najdi Arabic a voicing or an aspirating language? What does that mean in terms of the phonological representation/specification?

The present study starts with the following primary predictions of voicing contrast in Najdi Arabic:

a) In utterance-initial position, Voiced stops will be realised with prevoicing (voicing lead), while Voiceless stops will be realised with aspiration (long lag).
b) The distinction between stop categories will be extended to intervocalic stops. Voiced stops will be realised with strong active voicing throughout the hold phase. Voiceless stops, on the other hand, will not exhibit passive voicing.
c) In utterance-final position, the closure phase in Voiced stops will be devoiced due to the process of final devoicing.
d) Other acoustic correlates will be employed to signal the voicing contrast besides, or as alternatives to, VOT in utterance final position.
e) The speaking rate will affect the duration of prevoicing and aspiration in Voiced and Voiceless stops, respectively, in that they will be shortened under fast speech rate condition.
f) In terms of regressive voicing assimilation, the voicing of C1 will be affected by the voicing of C2 in both Voiced and Voiceless stops.

Chapter 4. Methods

4.1 Participants

The participants involved in the study were 40 monolingual native speakers of the central Najdi dialect (20 females, 20 males), aged between 19- 25. The participants were university students at King Saud University in Riyadh, who shared a similar socio-economic background, who lived in Riyadh, and who originated from an urban and sedentary population.

None of the participants reported any speech or hearing problems. The researcher ensured the students were born and raised in Riyadh and did not continue foreign language learning since they had graduated from high school. Each participant was asked to fill in a form to confirm the aforementioned information. (Demographic survey provided in Appendix A).

4.2 Stimuli

The stimuli of the present study consisted of words and phrases that include Voiced and Voiceless stops /t, k, b, d, g/ in the following contexts: utterance-initial, word-medial intervocalic, utterance-final, and stop-stop clusters at words boundaries. All the words and phrases were embedded in natural sentences: they were placed at the beginning of the sentence in case of utterance-initial, word-medial intervocalic, and stop-stop cluster across word boundaries, and they were placed at the end in case of utterance-final stops. The stimuli consisted of a hundred and five natural sentences that share similar length (3 to 4 words for each). For instance, the word b*uː**k*** 'your father' was embedded in the sentence *gid gaːbalt buː**k*** 'I've met your father' to examine ***k*** in utterance-final context; and the word ***t**iːn* 'figs' was embedded in the sentence ***t**iːn abha na ðˤiːf* 'Abha's figs are clean' to examine ***t*** in utterance-initial context. The sentences were revised by three native speakers of Najdi Arabic to ensure the target words were familiar and frequently used by Najdi speakers. The stimuli were divided into three parts: 1) utterance-initial stops, 2) utterance-medial stops (word-medial intervocalic), 3) utterance-final stops, and 4) stop-stop clusters at the word boundaries. The full list of the test words is included in the Appendix (Appendix B). The following sections present the structure of the test words in every context.

4.2.1Uterance-initial stops

The list included twenty-five words embedded in natural sentences with Voiced and Voiceless stops /t, k, b, d, g/ at three places of articulation: bilabial (Voiced only), alveolar, and velar.

The words were monosyllabic **C**V:C and they included long vowels that differed in height and backness [iː, eː, aː, oː, uː]: e.g. ***b**aːt* ‘slept’, ***t**iːn* ‘figs’, ***d**uːd* ‘worms’, ***k**oːm* ‘group of’.

4.2.2 Word-medial intervocalic stops

The list included forty-five words embedded in natural sentences with Voiced and Voiceless stops /t, k, b, d, g/ at three places of articulation: bilabial (Voiced only), alveolar, and velar. The test words were disyllabic with trochaic (word-initial stress) and iambic (word-medial stress) stress patterns. In iambic stress, (CV1ˈ**C**Vː2C), V1 was controlled (/a/) while Vː2 included long vowels differ in height and backness [iː, aː, uː]: e.g. *ʃaˈ**b**aːb* ‘youth’, *dʒaˈ**d**iːd* ‘new’, *kaˈ**t**uːm* ‘secretive’. In trochaic stress (ˈCVː1**C**V2C), V2 was controlled (/a/) while Vː1 included long vowels that differed in height and backness [iː, aː, uː]: e.g. *ˈbiː**g**at* ‘stolen’, *ˈduː**d**ah* ‘worm’.

4.2.3 Utterance-final postvocalic stops

The list included twenty-five words embedded in natural sentences with Voiced and Voiceless stops /t, k, b, d, g/ at three places of articulation: bilabial (Voiced only), alveolar, and velar. The words were monosyllabic CV:**C** and they included long vowels that differed in height and backness [iː, eː, aː, oː, uː]: e.g. *baː**b*** ‘door’, *fiː**k*** ‘in you’, *foː**g*** ‘up’. Short vowels were not included to avoid having final position geminate stops because of the bimoraicity of syllable structure in Arabic (Kiparsky, 2003).

4.2.4 Stop-stop clusters at the word boundaries

The list of phrases included five pairs of words (ten tokens) embedded in natural sentences with Voiced and Voiceless C1-C2 clusters at the word boundaries. C1 was a postvocalic stop in word-final position (CVː1**C**1), while C2 was a prevocalic stop in word-initial position (**C**2Vː2C). V1 and V2 were controlled (V1: eː, V2: aː). The clusters examined included **kb**, **tg**, **bk**, **gt**, **dg** in the following real words: *beː**t** **g**aːsim* ‘Gasim’s house’, *geː**d** **g**aːsim* ‘gasim’s handcuff’, *ʃeː**b** **k**aːmil* ‘full grey hair’, *ʃeː**k** **b**aːsim*, and sw*eː**g** **t**aːmir* ‘Tamir’s market’. The baseline context for C1 was the same stop in utterance-final position whereas the baseline context for C2 was the same stop in utterance-initial position (same vocalic environments as well: /e:/ preceding C1, /a:/ following C2). The words for the baseline context were selected from the stimuli for utterance-initial and utterance-final stops in the present study. Table 4.1 below presents the baseline contexts for C1 and C2 for each cluster.

Cluster	context	C1 base (utterance-final stop)	C2 base (utterance-initial stop)
bk	*ʃe:**b** **k**a:mil*	*ʃe:**b***	***k**a:l*
kb	*ʃe:**k** **b**a:sim*	*bre:**k***	***b**a:t*
dg	*ge:**d** **g**a:sim*	ke:***d***	***g***a:m
tg	*be:**t** **g**a:sim*	*be:**t***	***g***a:m
gt	swe*:**g** **t**a:mir*	bge:***g***	***t**a:b*

Table 4.1 The clusters and the baseline contexts for C1 and C2.

4.3 Procedures

Data were recorded in a soundproof recording studio at King Saud University in Riyadh. The participants were asked to read the sentences (presented in colloquial form) from a computer screen one by one. In the first experiment, they were asked to read the sentences naturally in their normal tone (normal speaking rate). In the second experiment, they were asked to read the sentence as fast as they could without sacrificing comprehensibility (fast speaking rate). The advantage of this simple way in testing the speech rate effect is to keep the participants' production as natural and real as possible. The stimuli for stop-stop clusters were not included in the second experiment. Each sentence was repeated three times in each experiment. The sentences and repetitions were randomised so that the same word did not appear consecutively. The total number of tokens was 24000 tokens (first experiment = 105 x 3 repetitions = 315/ second experiment = 95 x 3 repetitions = 285; 315 + 285 = 600/ 600 x 40 speakers = 24000 tokens).

A special instruction asked the participants to read the sentences in their native dialect and allowed them to correct themselves if they mispronounced the test words or pronounced the sentences in Standard Arabic. To ensure that the participants read the sentences in the colloquial form, the writing style of the sentences was not following the syntactic rules of Standard Arabic '*fusha*'. For example, the sentence ***قد قابلت بوك*** /*gid ga:balt bu:**k***/ 'I have met your father' is written in the colloquial style. In Standard Arabic, it should be written as follows: ***قَد قَابلتُ أباك***/ *qad qa:baltu aba:**k***/.

The production was recorded using a Zoom H6 Handy Recorder which was placed 15 cm away of the mouth. The recording was made at a sampling frequency of 44,100 Hz, 16-bit quantisation in mono-channel. The recording session lasted approximately 30 minutes for each speaker. The participants were given a break for 10 minutes before the fast speech experiment to achieve the experiment with the required accuracy and fluency.

4.4 Acoustic analysis

The software PRAAT (Boersma and Weenink, 2016) was used to perform the acoustic analysis. All tokens of the target words were manually transcribed and segmented in

Textgrids with reference to waveforms and spectrograms. In the present study, I follow what has been proposed in Turk et al. (2006) which emphasises the role of constriction in segmenting the sounds. However, I did not consider aspiration to be part of the following vowel. To obtain accurate acoustic measurements, automatic procedures (Praat scripts) adopted from Al-Tamimi and Khattab (2018) were used for two purposes: the adjustment of measurement points and quantifying the degree of voicing in stops across phonetic contexts. The first script employs F0 computation of the intensity within the glottal cycle to precisely adjust onset/mid/offset positions for each segment, a step that reduces errors resulting from the automatic extraction (Al-Tamimi and Khattab, 2018). The second script was the Praat voicing detection algorithm (VUV function) which creates a new tier with boundaries determining the voicing portions in each sound file. This measure helps in distinguishing true voice portions which are strong in their amplitude from passive voice which are not.

The manual and automatic segmentation, clarified above, yielded the following tiers and intervals:

a) A segment tier which presents the segments of the target words transcribed phonetically.

b) An acoustic properties tier that shows the acoustic properties for the target stops, including **CD**: for closure duration (intervocalic and final stops), it was determined from the end of the preceding vowel till the onset of the first visible burst. **B:** for release burst, it was determined as the beginning of the visible burst which appears in the spectrogram as vertical line. In case of multiple bursts, the ones separated by less than 5 ms were measured together as the burst (Al-Tamimi and Khattab, 2018). **ASP:** aspiration in Voiceless stops (t,k), and **A:** for aspiration in Voiced stops (b,d,g) were determined from the offset of the burst to the onset of the following vowel for initial and medial stops and to the disappearance of the aspiration in terms of utterance-final stops, respectively.

c) **VI** and **VF:** for vowel utterance-initial and utterance-final stops, respectively. **V1M**I and **V2MI:** for the preceding and following vowels in word-medial intervocalic stops (iambic). **V1MT** and **V2MT**: for the preceding and following vowels in word-medial intervocalic stops (trochaic). They were measured from the onset till the offset of the vowel formants including formant transitions (Turk et al., 2006).

d) A VUV tier which shows voiced **V** and unvoiced portions **U** in the acoustic properties. **V** intervals in case of Voiced stops were used to measure prevoicing in utterance-initial stops and voicing in the closure in word-medial intervocalic and utterance-final stops. Prevoicing was measured from the onset of the voicing cycle till the end of it at the onset of the release burst (Lisker and Abramson, 1964).

e) A word tier which simply presents the target word.

f) A position tier which shows the phonetic context for each stop, comprising **initial**: for utterance-initial stops, **final:** for utterance-final stops, **midT:** for intervocalic stops in trochaic foot, **midI:** for intervocalic stops in iambic foot, **finalC:** for the first member of the cluster, and **initialC:** for the second member of the cluster.

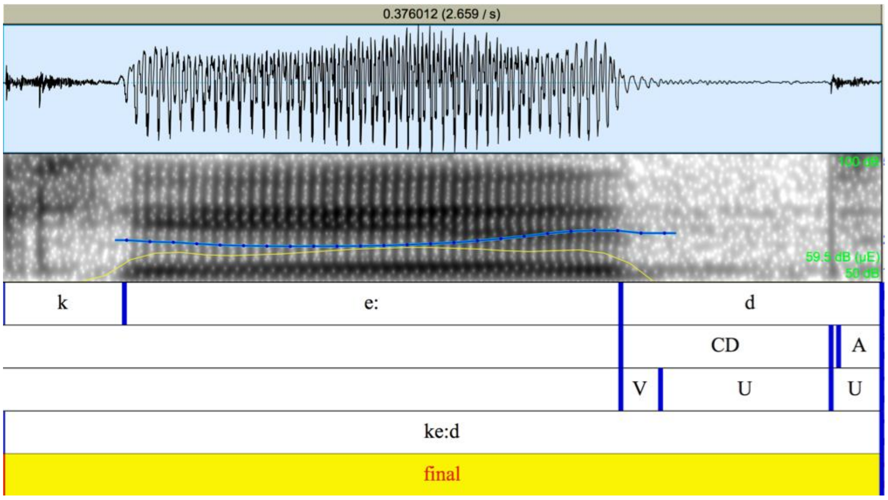


Figure 4.1. The segmentation of the word *ke:d* ‘conspiracy’ produced in utterance-final position produced by S13 in normal speech rate.

Acoustic measurements were extracted using a PRAAT script that automatically takes the measurements for the acoustic labels which had been manually checked. All the acoustic correlates were measured in the chosen phonetic contexts in both speech rates conditions (normal, fast). The target acoustic properties differed based on the phonetic context. For utterance-initial stops, prevoicing, aspiration, and release burst duration and intensity were measured. For the following vowels, the acoustic measurements included the difference between the amplitude of the first and second harmonics (H1*-H2* onset), F0 and F1 onsets, in addition to the absolute duration. For word-medial intervocalic stops, the following features were measured: prevoicing, aspiration, closure duration, voicing in the closure, release burst duration and intensity, preceding and following vowel duration, F1 and F0 at the offset of the preceding vowel and at the onset of the following vowel, and H1*-H2* at the offset of the preceding vowel and at the onset of the following vowel. For absolute final stops, voicing in the closure, closure duration, release burst duration and intensity, the preceding and following vowel duration were measured, F1 and F0 at the offset of the preceding vowel, and H1*-H2* at the offset of the preceding vowel. To test for voicing assimilation in stop-stop clusters at word boundaries, the acoustic characteristics of C1 and C2 were investigated. For both C1 and C2, voicing duration the closure, burst intensity for C1 and C2, F0/F1 onset (following C2), and F0/F1 offset (preceding C1). (See table 4.1 for a summary).

Utterance-initial stops	Intervocalic-word-medial stops	Utterance-final stops	Stop-stops clusters
VOT	Closure duration	Closure duration	Voicing duration in C1 and C2
Release burst duration and intensity	% voicing in the closure	% voicing in the closure	F0 offset (before C1)
F0 onset	Release burst duration and intensity	Release burst duration and intensity	F0 onset (after C2)
F1 onset	PV duration	PV duration	F1 offset (before C1)
H1-H2 onset	FV duration	H1-H2 offset (PV)	F1 onset (after C2)
FV duration	F0 onset (FV)		Burst intensity for C1 and C2
	F1 onset (FV) H1-H2 offset (PV) H1-H2 onset (FV)		

Table 4.2 The summary of the acoustic measurements for each phonetic context.

4.4.1 Characterisation of voicing, devoicing, and aspiration patterns.

This section presents the criteria for describing patterns as voicing, devoicing or aspiration in the examined tokens under normal and fast speech rate conditions. It is important to identify the terminology used in the results section for the acoustic description of prevoicing, devoicing, and aspiration. *Prevoicing* refers to the duration of the voice bar that precedes the onset of the release in utterance-initial Voiced stops (the traditional voicing lead VOT). The values for prevoicing are presented in positive not negative values following Jessen (1998) to avoid confusion when presenting the change in response to speech rate (for instance, -67 is mathematically higher in value than -80). It is also important to note that prevoicing results are represented only for Voiced stops (no token was found for Voiceless stops showing prevoicing). The results also do not account for closure duration in utterance-initial stops (it was not possible to detect the closure onset in utterance-initial position). *Voicing duration* refers to the absolute duration of voicing in the closure in utterance-medial intervocalic and utterance-final Voiced and Voiceless stops. The reason for measuring voicing duration in Voiceless stops is to account for any possible passive voicing caused by the surrounding context. The results for voicing duration are also presented without including the closure in order to examine the effect of speech rate on the absolute duration of voicing. *% voicing* refers to the percentage of voicing in the closure in utterance-medial intervocalic and utterance-final Voiced and Voiceless stops. *Aspiration* refers to the aperiodic noise that follows the release burst in Voiced and Voiceless stops across the positions. The burst, which is the observed transient, is not included in the aspiration following Klatt (1970) and Al-

Tamimi and Khattab (2018). The term *aspiration* was used instead of VOT for two reasons: 1) aspiration was measured for all tokens of Voiced stops regardless of their phonetic voicing which contradicts the definition of VOT, 2) aspiration was measured in utterance-final stops which is called by some researchers Voice Offset Time (Jansen 2004); so the term aspiration was used instead of Voice Offset time to avoid ambiguity the correlates. The term *short aspiration* was used to describe unaspirated stops (UASP: 0-34 ms), *moderate aspiration* for stops with (MASP: 35-60 ms), and *heavy aspiration* for stops with (HASP: above 60 ms).

A special coding system was created to characterise the phonetic realisation of each stop accurately in each context (see Figure 4.2). This coding aims to provide a more precise description for the voicing and aspiration status of the examined stops. It is more sophisticated for the description of the phonetic voicing than the common terms (fully Voiced, partially Voiced), at least for the purpose of this section. The present study departs from what has been proposed in previous studies regarding the standards for considering a stop a phonetically Voiced one, such as in Whalen and Abramson (2017) who consider 50% to be the threshold for voicing characterisation. In the present study, if a Voiced stop in word-medial intervocalic position gets the symbol **V**, it means the voicing percentage is 100% whereas if it is **V80**, it means the percentage is within the range of 80%–99%. With regard to aspiration, the proposed coding system enhances the accuracy of describing aspiration in Najdi Arabic. Additionally, the classification of aspiration can be used to check the interaction between F0 and aspiration to test the phonetic and phonological implications of this interaction. The categories for aspiration were determined following Keating (1984) and Jansen (2004).

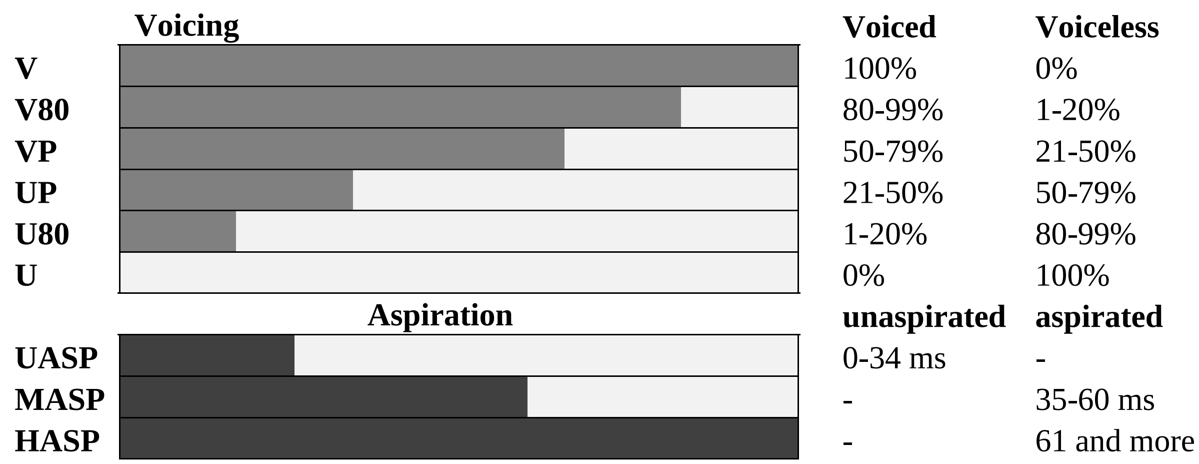


Figure 4.2 The coding symbols and their meaning for describing voicing, devoicing, and aspiration in stops.

4.4.2 Token exclusion

The number of tokens expected at the initial stage of the research was 24,000 tokens (40 speakers x 600 sentences). However, the recordings of 5 male participants and 3 female participants were excluded due to poor recording and noise in the background. The nature of the acoustic signal requires accurateness in the quality of the recording, as well as the performance of the speaker in order to segment the sound files and extract the correlates without losing important details that might affect the results (Roettger, 2019). For stop-stop clusters across word boundaries, tokens where the speakers paused between the two members of the cluster were excluded due to the nature of the process tested in this phonetic context (regressive voicing assimilation). After this exclusion, the number of tokens included in the analysis were 19,200 tokens (32 speakers x 600 sentences).

4.5 Statistical analysis

To analyse the data of the current research, lme4 package (Bates et al., 2014) was used in R (R core team, 2014) to conduct a linear mixed-effect model (LMM) for each of the acoustic correlates that differentiated Voiced and Voiceless stops. The models were conducted first in maximal form by including all the fixed effects, random intercepts for each of the random effects, as well as random slopes for each of the fixed effects (Barr et al., 2013). All the maximal models converged without any issues in the analysis. To get more accurate results, contrast coding was applied on all the fixed effects by assigning each level in each fixed effect an equal distance values between -0.5, 0.5. For instance, the levels in the place of articulation were assigned values using the function *mutate* as follows: velar = 0.5, alveolar = 0, bilabial = -0.5, which yielded the centred fixed effect *place_c*. There were no interactions between fixed effects added to the models for two reasons: 1) by using contrast coding on the fixed effects and centring their levels, the results are generalised over all the interactions (Al-Tamimi and Khattab, 2018), 2) including interactions did not improve the fit of some of the tested models (AIC values increased with interactions). The parallel structure between the models is a requirement in the present study to compare the performance of the acoustic correlates across phonetic contexts. The LMM analysis was performed based on a predictive modelling approach to visualise the predictions of each of the models (Kuhn and Johnson, 2013). After running the LMM model for each correlate, the function *predict* was used to get the predicted values of the models. The predicted values, then, were presented through boxplot figures and tables that show means and standard deviations for Voiceless and Voiced stops at normal and fast speech rates. To compare between the predicted values of the acoustic correlates in each boxplot and table, a pairwise comparison was performed using *pairwise-t-test* in R with FDR (False Discovery Rate) corrections. The *pairwise-t-test* was performed on the fitted values presented in each graphic results. For example, if the table presents the results for the predicted values of F0 onset across voicing, speech rate, and place, the *pairwise-t-test* is performed using the function *pairwise.t.test* as in the following code: *pairwise.t.test (Main data set $ predicted data set, Main data set: voicing, Main data set: rate, Main Data set: place, p.adjust = "fdr")*

The structure of each model differs based on the phonetic context. The two subsections below provide a detailed description of the models.

4.5.1 LMM for initial, medial, and final stops

The main objective of the LMM analysis in initial, medial, and final stops is to examine the acoustic correlates in Voiced and Voiceless stops in Najdi Arabic across positions. Accordingly, a linear mixed-effects model was built to test each acoustic correlate that is expected to cue the phonological opposition between Voiced and Voiceless stops across positions as a function of voicing, speech rate, place of articulation, vowel type, and gender. The models were structured as follows:

> acoustic correlate ~ voicing_c + place_c + rate_c + vowel_c + gender_c + (1 + voicing_c + place_c + rate_c + vowel_c || speaker) + (1 + gender_c| word)

The models included the acoustic correlates as dependent variables, followed by (~) to indicate "as a function of". The fixed effects included voicing (levels: Voiceless/Voiced), place (levels: bilabial/alveolar/velar), rate (levels: normal/fast), vowel (utterance-initial and utterance-final contexts: levels: /i:/e:/a:/o:/u:/, word-medial intervocalic context: levels: /i:/, /a:/, /u:/), and gender (levels: male/female). *Speaker* and *word* were added as random effects. The number '1' next to *word* and *speaker* indicates the addition of by-word and by-speaker random intercepts to the analysis. The model also included voicing, place of articulation, rate, and vowel type as by-speaker random slopes and gender as by-word random slope.

4.5.2 LMM for stop-stop clusters

The acoustic correlates in stop-stop clusters across word boundaries were examined to check for any changes that occur due to voicing assimilation in C1 and C2. The acoustic results for C1 were compared to that of C1 in the baseline context, and the acoustic results for C2 were compared to that of C2 in the baseline context. Accordingly, a linear mixed-effects model was built to test each acoustic correlate in C1 as a function of voicing, place of articulation, cluster, context, and gender. On the other hand, a linear mixed-effects model was built to test each acoustic correlate in C2 as a function of voicing, place of articulation, cluster, context, and gender. The models were conducted as follows:

acoustic correlate ~ voicing_c + place_c + cluster_c + context_c + gender_c +(1 + voicing_c + place_c + cluster_c + context_c | speaker) + (1 + gender_c| word).

The models included the acoustic correlate as a dependent variable, followed by (~) to indicate "as a function of". The fixed effects included voicing (levels: Voiceless/Voiced), place (levels: bilabial/alveolar/velar), cluster (levels: /bk/, /kb/, /tg/, /dg/, /gt/), context (for C1 models: levels: C1 cluster/ C1 baseline, for C2 models: levels: C2 cluster/ C2 baseline), and gender (levels: male/female). *Speaker* and *word* were added as random effects. The number '1' next to *word* and *speaker* indicates the addition of by-word and by-speaker random intercepts to the analysis. The model also included voicing, place of articulation, cluster, context as by-speaker random slopes and gender as by-word random slope.

4.6 Vowel duration as a proxy of speech rate

This section aims to examine vowel duration in the test words under normal/fast conditions and use it as a proxy of speech rate to make sure that the participants performed the experiments accurately and produced the required difference in the two speech rate conditions. The vowels included in the analysis are the stressed vowels in the vicinity of the target stops. An optimal linear mixed effect model was built considering vowel duration as an independent variable in the function of voicing (Voiceless-Voiced), rate (normal-fast), place (bilabial-alveolar-velar), vowel type (i:-e:-a:-o:-u:), and gender (male-female). *Speaker* and *word* were added as random effects with random intercepts. Voicing, place of articulation, rate, and vowel type as by-speaker random slopes and gender as by-word random slope. Following the same procedures mentioned in 4.5 above, contrast coding, the predictive approach, and the pairwise comparison were performed. The results are shown in table 4.3 and plotted in figure 4.3.

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	113.07	4.01	28.15	**<0.0001**
Vowel-c	4.97	5.71	0.86	0.38
Rate-c	-30.38	1.98	-15.3	**<0.0001**
Place-c	21.12	2.46	8.56	**<0.0001**
Voicing-c	-7.45	2.4	-3.09	**0.002**
Gender-c	4.58	5.41	0.84	0.403

Table 4.3. The results of the linear mixed effects model for vowel duration under normal/fast.

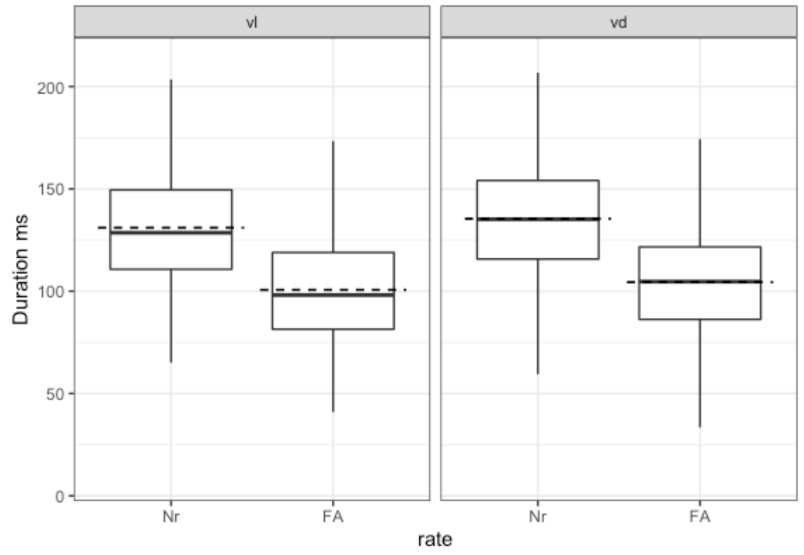


Figure 4.3 Boxplots of the fitted values of vowel duration classified by voicing (vl = Voiceless, vd = Voiced), and speech rate (FA = fast, Nr = Normal). The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	131	27.1	135	26.9
Fast	101	25.4	104	25.2

Table 4.4 Mean and standard deviation of the fitted values of aspiration for Voiceless and Voiced utterance-initial stops grouped by voicing and speech rate.

The model output shows that speech rate is significantly affecting vowel duration in that vowels in fast speech is shorter than in normal speech with an average of 30 ms ($p<0.0001$). It also shows that the effect of gender and vowel type were not obtained. The predicted values in figure (4.3) and table (4.4), generalised over all the interactions, confirm the same difference in the context of both Voiceless and Voiced stops.

4.7 The presentation of the results

The next four chapters present the results of the acoustic analysis for the patterns of voicing and aspiration (chapter 5), the durational correlates (chapter 6), the spectral correlates (chapter7), and regressive voicing assimilation in stop-stop clusters at word boundaries (chapter 8). All the statistical output of LMM models and pairwise comparisons are presented in Appendix C.

In chapter 5,6, and 7, the results of the acoustic analysis for each correlate in each position are presented through figures and tables in two sections: 1) the values as a function of voicing and rate, and 2) the values as a function of voicing, rate, and place of articulation. In terms of the results as a function of vowel and gender, they are presented in Appendix C only if they are significant in the LMM models. Each section ends with a short summary that gives the main results and links them with the previous studies. At the end of each chapter, the results are summarised in a table that shows the acoustic values (mean/standard deviation) under normal and fast speech rate conditions.

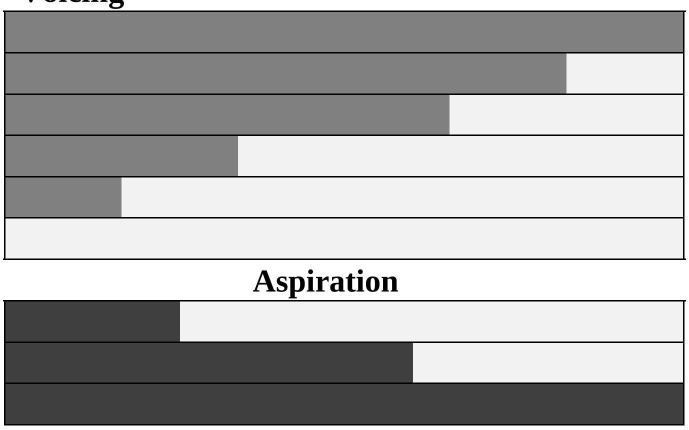
Chapter 5. Results for the patterns of voicing and aspiration

This chapter presents the results of the acoustic analysis of voicing and aspiration in the target stops in the determined phonetic contexts under normal/fast speech rate conditions. As previously noted, the present study focuses on two sides of voicing contrast in Najdi Arabic: the phonetic realisation and the phonological representation. This chapter focuses on the acoustic analysis of voicing and aspiration expected to signal the distinction between Voiced and Voiceless stops in Najdi Arabic in utterance-initial, utterance-medial intervocalic, and utterance-final positions. The acoustic correlates investigated in this chapter includes aspiration, %voicing, prevoicing, and voicing duration. The results of the acoustic analysis of voicing and aspiration are expected to provide the description required to form the foundation for the phonological representation of voicing contrast in Najdi Arabic and the choice of distinctive features specifying this contrast.

This chapter is divided into three main parts: 1) general results, focusing on the results of instances of voicing, devoicing, aspiration following the coding system employed in the present study to characterise these features; 2) the acoustic analysis, focusing on the patterns of voicing and aspiration in Voiced and Voiceless stops in utterance-initial ***C***V:C, utterance-medial CV:***C***V(C)/ CV***C***V:(C), and utterance-final CV:***C***; and 3) the summary of the results for the examined correlates under normal/fast speech rate conditions.

5.1 General results

This section presents the percentages of tokens categorised as instances of voicing, devoicing, and aspiration under normal/fast speech rate conditions. For the purpose of this section, a coding system was employed to characterise the phonetic realisation of each stop in each context (see Figure 5.1, note: reproduced again here to be easier for the reader to follow).



	Voicing	Voiced	Voiceless
V		100%	0%
V80		80-99%	1-20%
VP		50-79%	21-50%
UP		21-50%	50-79%
U80		1-20%	80-99%
U		0%	100%
	Aspiration	**unaspirated**	**aspirated**
UASP		0-34 ms	-
MASP		-	35-60 ms
HASP		-	61 and more

Figure 5.1 The coding symbols and their meaning for categorising instances of voicing, devoicing, and aspiration in stops.

The following figures and tables present general description of voicing, devoicing, and aspiration in Najdi Arabic.

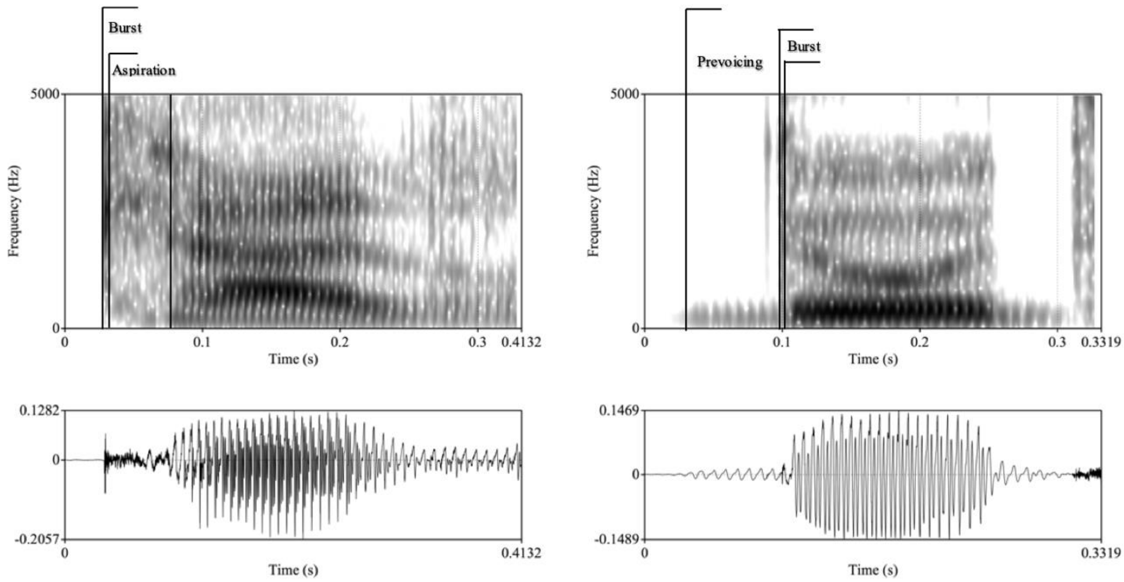


Figure 5.2 Examples of utterance-initial stops produced by male speaker S01, Voiceless aspirated stop on the left (kʰaːl ‘weigh’), prevoiced Voiced stop on the right (duːd ‘worms’).

Voiced					Voiceless				
Prevoicing					Aspiration				
	Normal		Fast			Normal		Fast	
V	(1358)	90.5%	(1235)	86%	**UASP**	(94)	10%	(413)	43.6%
U	(142)	9.5%	(201)	14%	**MASP**	(614)	64.3%	(503)	53.1%
					HASP	(246)	25.7%	(31)	3.3%
Total	1500		1436		**Total**	954		947	

Table 5.1 The percentage of voicing, devoicing, and aspiration for utterance-initial stops (Voiced/Voiceless: phonological voicing. Numbers in parentheses refer to the number of tokens.)

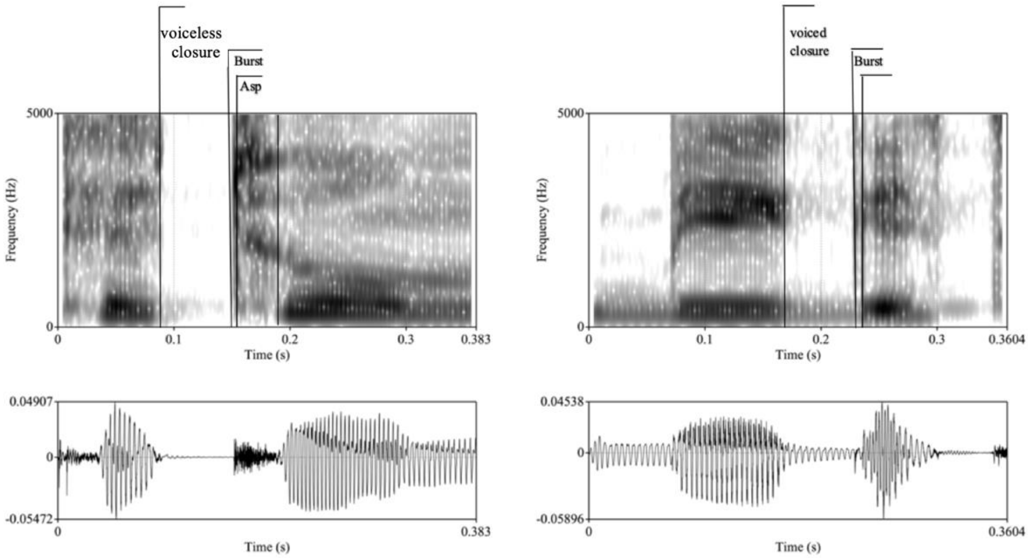


Figure 5.3 Examples of utterance-medial intervocalic stops produced by female speaker S13, Voiceless aspirated stop on the left (*katʰuːm* 'secretive'), prevoiced stop on the right (*biːgat* 'stolen').

Voiced					Voiceless				
Voicing in CD					Voicing in CD				
	Normal		Fast			Normal		Fast	
V	(794)	98%	(749)	98%	**U**	(518)	98%	(433)	94%
VP	(5)	1%	(11)	1.4%	**UP**	(2)	0.37%	(11)	2.41%
UP	(5)	1%	(3)	0.6%	**U80**	(6)	1.26%	(8)	1.75%
Total	804		763		**V**	(2)	0.37%	(4)	1.84%
					Total	528		456	
					Aspiration				
						Normal		Fast	
					UASP	(154)	29.2%	(264)	57.8%
					MASP	(282)	53.6%	(176)	38.5%
					HASP	(90)	17.2%	(16)	3.7%
					Total	526		456	

Table 5.2 The percentage of voicing, devoicing, and aspiration for utterance-medial stops (iambic) (Voiced/Voiceless: phonological voicing. Numbers in parentheses refer to the number of tokens.)

Voiced					Voiceless				
Voicing in CD					Voicing in CD				
	Normal		Fast			Normal		Fast	
V	(1649)	98%	(1583)	97%	**U**	(1050)	93%	(992)	92%
VP	(20)	1.19%	(35)	2.14%	**UP**	(37)	3.2%	(55)	5.1%
UP	(8)	1%	(8)	0.6%	**U80**	(31)	2.7%	(14)	1.3%
V80	(1)	0.47%	(1)	0.49%	**V**	(8)	0.7%	(10)	0.92%
U80	(2)	0.34%	(2)	0.37%	**V80**	(0)	0%	(1)	0.09%
Total	1680		1629		**VP**	(2)	0.17%	(4)	0.59%
					Total	1128		1076	
					Aspiration				
						Normal		**Fast**	
					UASP	(645)	57.1%	(902)	84.1%
					MASP	(466)	41.2%	(170)	15.9%
					HASP	(18)	1.7%	(0)	0%
					Total	1129		1072	

Table 5.3 The percentage of voicing, devoicing, and aspiration for utterance-medial stops (trochaic) (Voiced/Voiceless: phonological voicing. Numbers in parentheses refer to the number of tokens.)

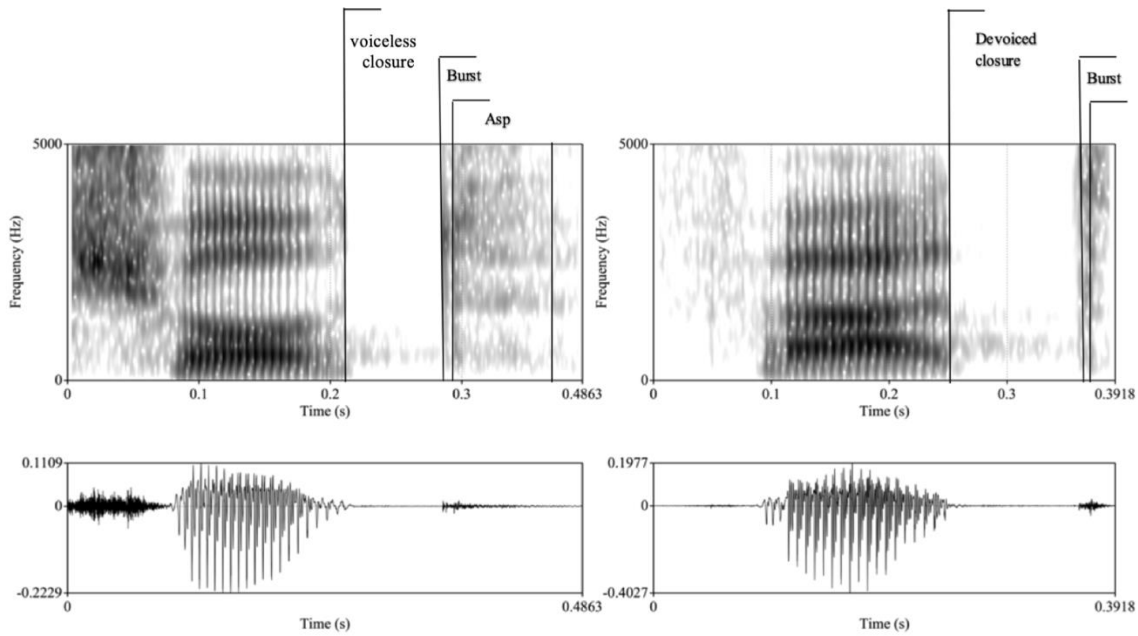


Figure 5.4 Examples of utterance-final stops produced by male speaker S06, Voiceless aspirated stop on the left (*fo:tʰ* 'kick'), Voiced stop on the right (*fa:d* 'benefited').

Voiced					Voiceless				
Voicing in CD					Voicing in CD				
	Normal		Fast			Normal		Fast	
V	(32)	2.4%	(77)	7.1%	**U**	(888)	99%	(778)	97%
VP	(103)	7.8%	(94)	8.7%	**UP**	(3)	0.33%	(13)	1.63%
V80	(1)	0.07%	(12)	1.1%	**U80**	(2)	0.22%	(1)	0.21%
U	(782)	59.2%	(678)	63.2%	**V**	(0)	0%	(3)	0.37%
UP	(292)	22%	(176)	16.4%	**VP**	(1)	0.45%	(0)	0%
U80	(109)	8.2%	(35)	3.2%	**Total**	894		795	
Total	1319		1072		**Aspiration**				
						Normal		**Fast**	
					UASP	(14)	1.7%	(326)	41%
					MASP	(117)	13%	(318)	40%
					HASP	(765)	85.3%	(149)	19%
					Total	896		793	

Table 5.4 The percentage of voicing, devoicing, and aspiration for utterance-final stops (Voiced/Voiceless: phonological voicing. Numbers in parentheses refer to the number of tokens.)

Tables (5.1-4) show the proportion of voicing, devoicing, and aspiration patterns identified in the examined phonetic contexts. Voiced stops were prevoiced in utterance-initial position in the majority of tokens across speech rates (Nr = 90.5%, FA = 86%). The percentage of devoicing is slightly higher in fast speech than in normal speech. Voiceless stops, on the other hand, were aspirated in utterance-initial position in the majority of tokens across speech rates (Nr = 90%, FA = 56%). This result confirms the main expectation about Najdi Arabic in that both prevoicing and aspiration exist in its voicing contrast system. In utterance-medial position, Voiced stops were fully voiced in the majority of the tokens across foot structure and speech rates. The devoicing percentage is slightly higher in the trochaic than in the iambic contexts across speech rates. Voiceless stops, on the other hand, were aspirated in the majority of tokens in the iambic context in the normal speech rate whereas in the trochaic context the majority of tokens were unaspirated. Voiceless stops in utterance-medial position showed a fully voiceless closure in the majority of the data across foot structure and speech rates. In utterance-final position, Voiced stops were devoiced in the majority of the tokens across speech rates (Nr = 59.2%, FA = 63.2%). Voiceless stops were aspirated in the majority of tokens across speech rates (Nr = 98.3%, FA = 59%).

These results suggest that Najdi Arabic has features from both voicing and aspirating languages. At this stage, it is still difficult to describe the manifestation of voicing contrast in Najdi Arabic in a precise manner. A more in-depth analysis of the phonetic aspects of voicing contrast in this dialect will be presented in the coming sections starting with the acoustic correlates that are expected to cue the phonological opposition between Voiced and Voiceless stops.

5.2 Acoustic analysis of the patterns of voicing and aspiration.

5.2.1 Aspiration (ms)

Aspiration is a crucial acoustic correlate that signals the distinction between Voiceless and Voiced stops in aspirating languages. Voiceless stops are expected to be realised with long lag aspiration while Voiced stops are expected to be realised with short lag aspiration. The presence of aspiration in Najdi Arabic reported in some studies in the literature (Flege and Port, 1981; Al-Gamdi et al., 2019) raises questions about the nature of such an acoustic correlate and its robustness taking into account various factors. Aspiration duration is examined in this section in Voiceless and Voiced stops regardless of their phonetic voicing in the closure.

5.2.1.1 Utterance-initial stops CV:C

Aspiration as a function of voicing and rate

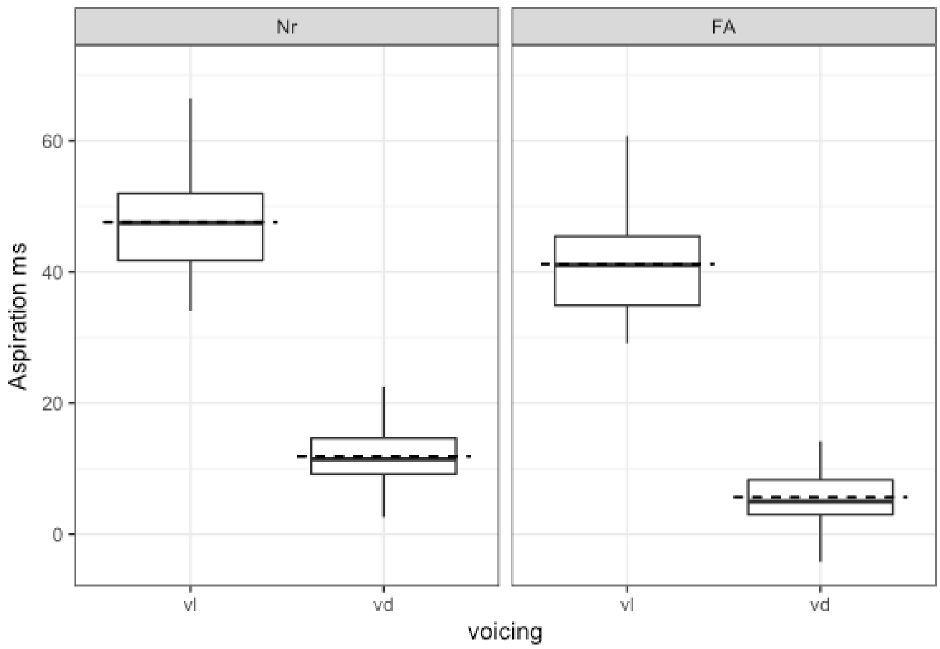


Figure 5.5 Boxplots of the fitted values of aspiration in utterance-initial stops classified by voicing (vl = Voiceless, vd = Voiced), and speech rate (FA = fast, Nr = Normal) and over place, vowel type and gender. The dashed line represents the mean.

	Voiceless		Voiced	
	Mean	SD	Mean	SD
Normal	47.6	7.19	11.9	3.92
Fast	41.2	7.31	5.65	3.42

Table 5.5 Mean and standard deviation of the fitted values of aspiration for Voiceless and Voiced utterance-initial stops grouped by voicing and speech rate.

As we can see in figure and table 5.5, voicing category has a clear impact on aspiration across speech rates. At normal speech rate, aspiration for the Voiceless was significantly longer than for the Voiced stops by an average of 35.5 ms across speech rates ($p<0.0001$). Moving to speech rate impact on aspiration, normal speech rate resulted in longer aspiration for both Voiceless and Voiced stops. For Voiceless stops, there was a significant difference by 6.4 ms in favour of normal speech ($p<0.0001$). A similar pattern was found for Voiced stops; aspiration was significantly longer in normal speech by an average of 6.25 ms ($p<0.0001$).

Aspiration as a function of voicing, rate, and place of articulation

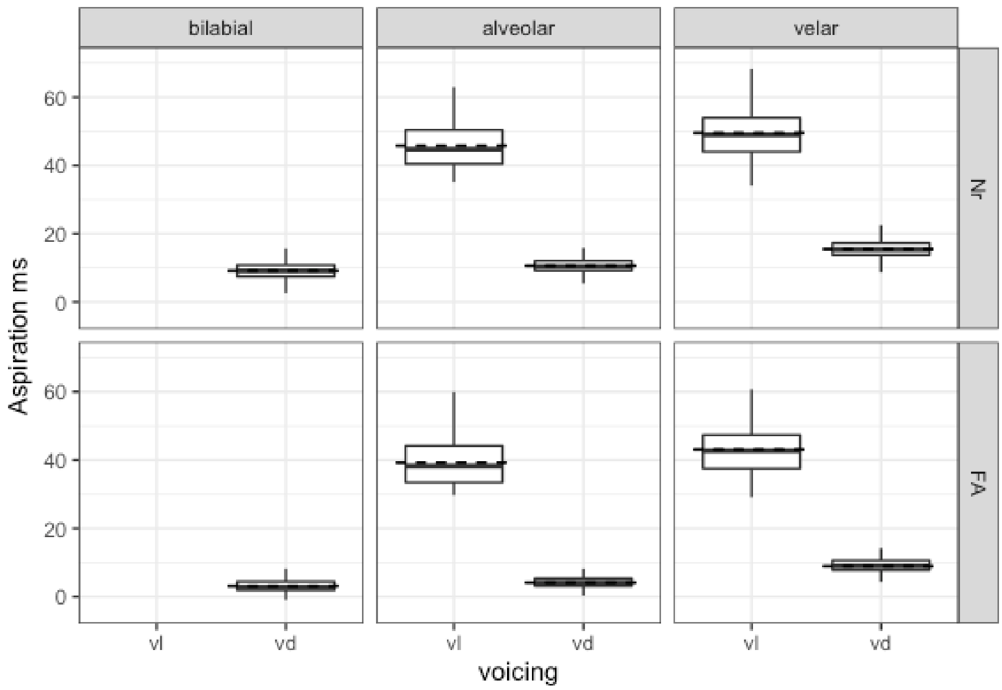


Figure 5.6 Boxplots of the fitted values of aspiration in utterance-initial stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal) and place of articulation. The dashed line represents the mean.

	Voiceless normal		Voiced normal		Voiceless fast		Voiced fast	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	9.09	2.94	-	-	3.15	2.06
Alveolar	45.8	6.69	10.6	2.46	39.3	6.8	4.17	1.91
Velar	49.5	7.22	15.5	2.98	43.1	7.31	8.93	2.72

Table 5.6. Mean and standard deviation of the fitted values of aspiration for Voiceless and Voiced utterance-initial stops grouped by speech rate and place of articulation.

With respect to the impact of place of articulation on aspiration, figure and table 5.6 show that the pattern appeared to be in the order velar > alveolar > bilabial across voicing categories and speech rates. For Voiceless stops, aspiration in velar stops was statistically longer than inalveolar ones by an average of 3.7 ms in both normal and fast speech rates ($p<0.0001$). Within velar vs alveolar for Voiced stops, aspiration was on average about 4.8 ms longer in velar than in alveolar stops in both normal and fast speech rates ($p<0.0001$). The results also showed that aspiration for velar stops were significantly longer than for bilabial stops by an average of 6.41 ms in normal speech ($p<0.0001$), and by an average of 5.78 in fast speech($p<0.0001$). Within alveolar vs bilabial for Voiced stops, aspiration in alveolar was longer by an average of 1.51 ms in normal speech ($p<0.0005$), and by an average of 1.02 ms ($p<0.008$) in fast speech.

5.2.1.2 Utterance-medial stops CV:CVC (Trochaic)

Aspiration as a function of voicing and rate

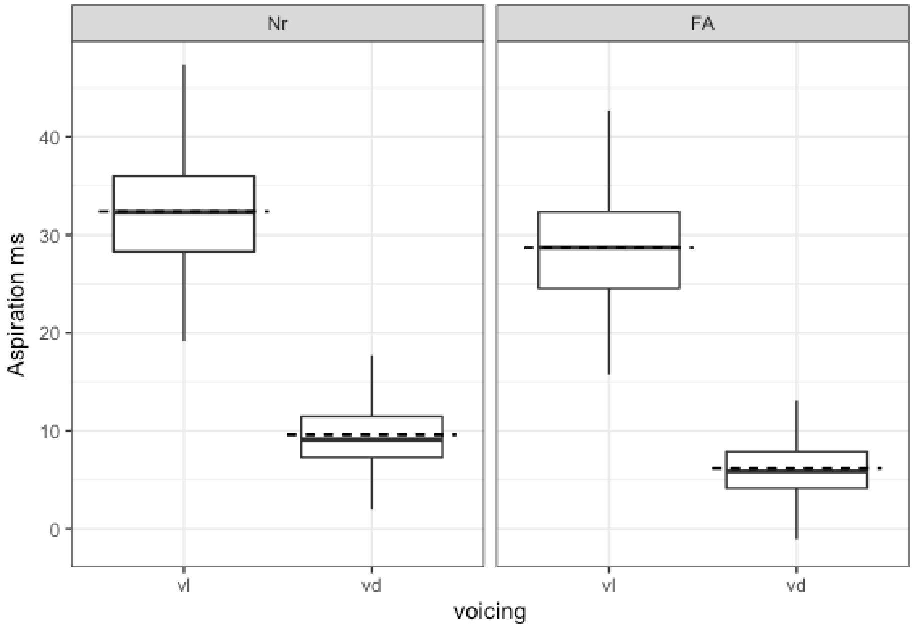


Figure 5.7. Boxplots of the fitted values of aspiration in utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced), and speech rate (FA = fast, Nr = Normal). The dashed line represents the mean.

	Voiceless		Voiced	
	Mean	SD	Mean	SD
Normal	32.4	5.68	9.59	3.38
Fast	28.7	5.62	6.19	3.05

Table 5.7. Mean and standard deviation of the fitted values of aspiration for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

As figure and table 5.7 show, there is a clear separation between aspiration values for Voiced and Voiceless stops. Starting with normal speech, aspiration in Voiceless was significantly longer than in Voiced stops by an average of 22.81 ms ($p<0.0001$). With respect to fast speech, aspiration for the Voiceless was significantly longer than for the Voiced stops by an average of 22.51 ms ($p<0.0001$). It can be seen that normal speech resulted in longer aspiration for both Voiceless and Voiced stops. For Voiceless stops, there was a significant difference by 3.7 ms in favour of normal speech ($p<0.0001$). Similarly, aspiration for Voiced stops was significantly longer in normal speech by an average of 3.4 ms ($p<0.0001$).

Aspiration as a function of voicing, rate, and place of articulation

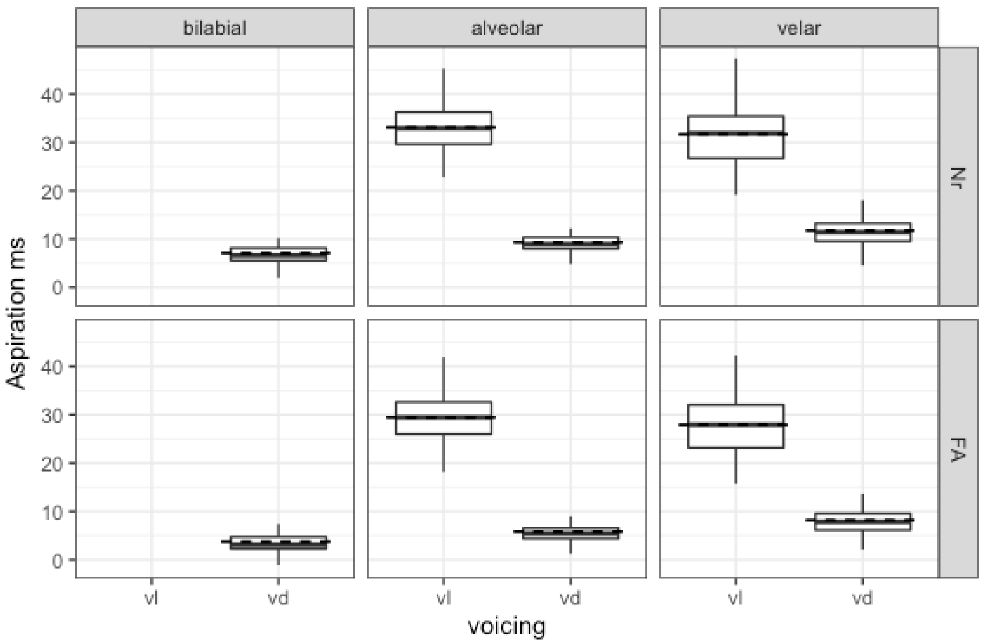


Figure 5.8. Boxplots of the fitted values of aspiration in utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal) and place of articulation. The dashed line represents the mean.

	Voiceless normal		Voiced normal		Voiceless fast		Voiced fast	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	7.14	2.77	-	-	3.74	2.38
Alveolar	33.1	4.97	9.34	2.52	29.5	4.97	5.78	2.11
Velar	31.7	6.23	11.8	3.23	27.9	6.1	8.17	2.95

Table 5.8. Mean and standard deviation of the fitted values of aspiration for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by speech rate and place of articulation.

The results showed that aspiration for alveolar stops were significantly longer than for velar ones by an average of 1.5 ms in both normal and fast speech rates ($p<0.0001$). For Voiced stops, however, the pattern appeared to be in the order velar > alveolar > bilabial across speech rates. Within velar vs alveolar for Voiced stops, aspiration for velar stops showed a significant increase by an average of 2.3 ms ($p<0.0001$) in both normal and fast speech rates. The results also showed that aspiration for velar stops was significantly longer than for bilabial stops by an average of 4.5 ($p<0.0001$) ms in both normal and fast speech. Within alveolar vs bilabial for Voiced stops, aspiration for the former showed a significant increase by an average of 2.1 ms ($p<0.0001$) in both normal and fast speech.

5.2.1.3 Utterance-medial stops CV*C*V:C(Iambic)

Aspiration as a function of voicing and rate

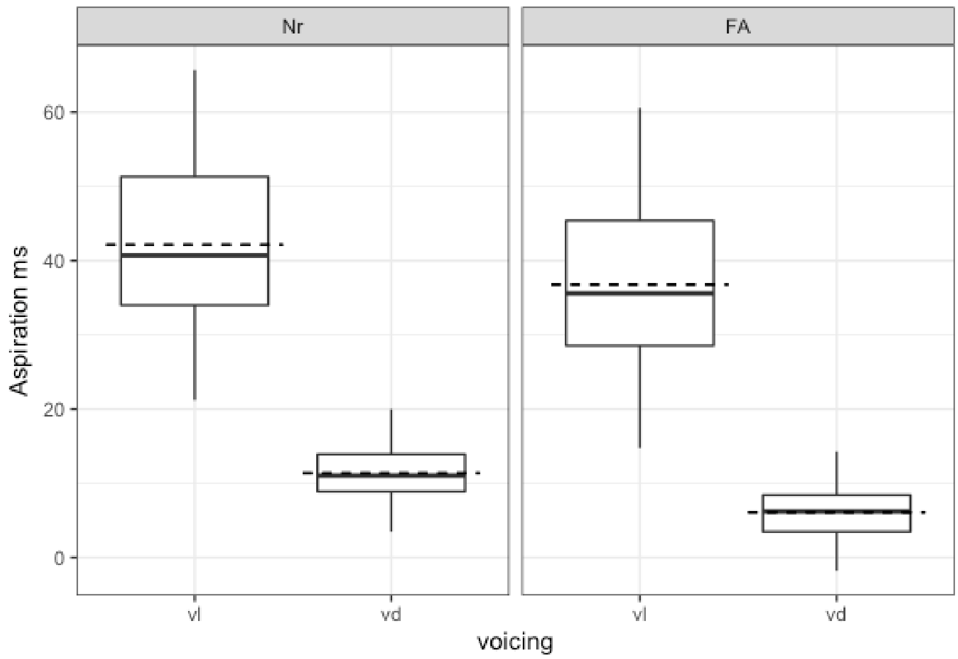


Figure 5.9. Boxplots of the fitted values of aspiration in utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced), and speech rate (FA = fast, Nr = Normal). The dashed line represents the mean.

	Voiceless		Voiced	
	Mean	SD	Mean	SD
Normal	42.1	10.5	11.4	3.59
Fast	36.8	10.4	6.09	3.48

Table 5.9. Mean and standard deviation of the fitted values of aspiration for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

Figure and table 5.9 show that voicing affected aspiration across speech rates. Aspiration in Voiceless was significantly longer than in Voiced stops by an average of 30.7 ms ($p<0.0001$) at both normal and fast speech rates. As for the effect of speech rate on aspiration, normal speech resulted in longer aspiration for both Voiceless and Voiced stops by an average of 5.4 ms ($p<0.0001$).

Aspiration as a function of voicing, rate, and place of articulation

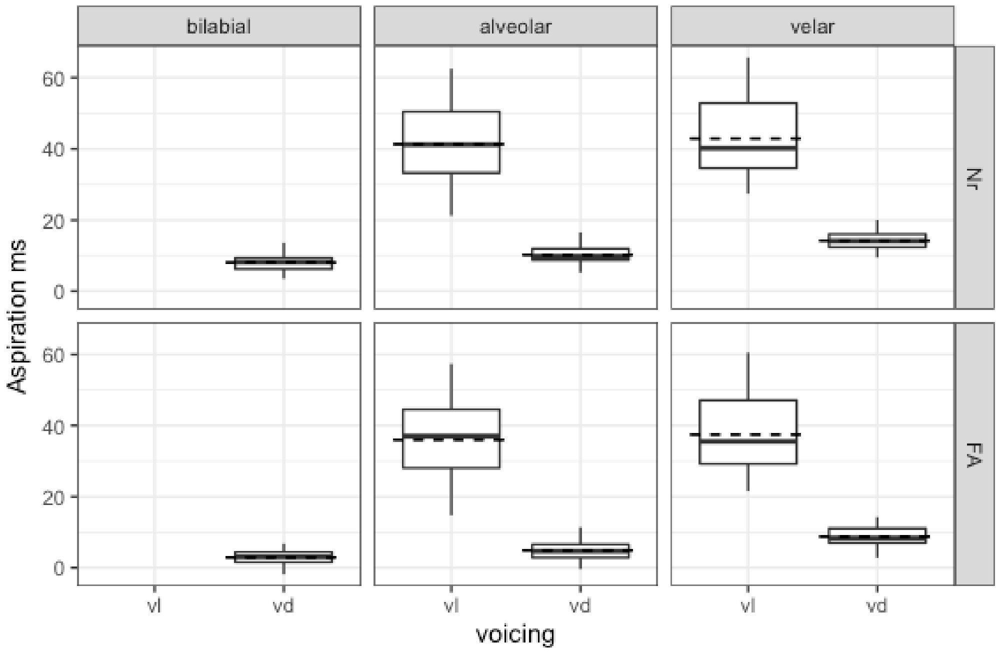


Figure 5.10. Boxplots of the fitted values of aspiration in utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal) and place of articulation. The dashed line represents the mean.

	Voiceless normal		Voiced normal		Voiceless fast		Voiced fast	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	8.04	2.27	-	-	2.92	2.07
Alveolar	41.3	10.5	10.2	2.49	36.1	10.3	4.93	2.59
Velar	43	10.6	14.2	2.71	37.5	10.5	8.79	2.65

Table 5.10. Mean and standard deviation of the fitted values of aspiration for Voiceless and Voiced utterance-medial stops (Iambic) grouped by speech rate and place of articulation.

As shown in figure and table 5.10, aspiration for velar stops was significantly longer than for alveolar ones by an average of 1.7 ms in normal ($p = 0.01204$) and fast speech ($p = 0.04193$)

rates. For Voiced stops, the pattern appeared to be in the order velar > alveolar > bilabial across speech rates. Within velar vs alveolar for Voiced stops, aspiration for velar stops showed a significant increase by an average of 3.9 ms ($p<0.0001$) in both normal and fast speech rates. The results also showed that aspiration for Voiced velar stops was significantly longer than for Voiced bilabial stops by an average of 6.015 ($p<0.0001$) ms in both normal and fast speech. The results showed that aspiration in alveolar stops was on average about 2.085 ms longer than in bilabial stops in both normal ($p = 0.00752$) and fast ($p = 0.0256$) speech rates.

5.2.1.4 Utterance-final stops CV:C

Aspiration as a function of voicing and rate

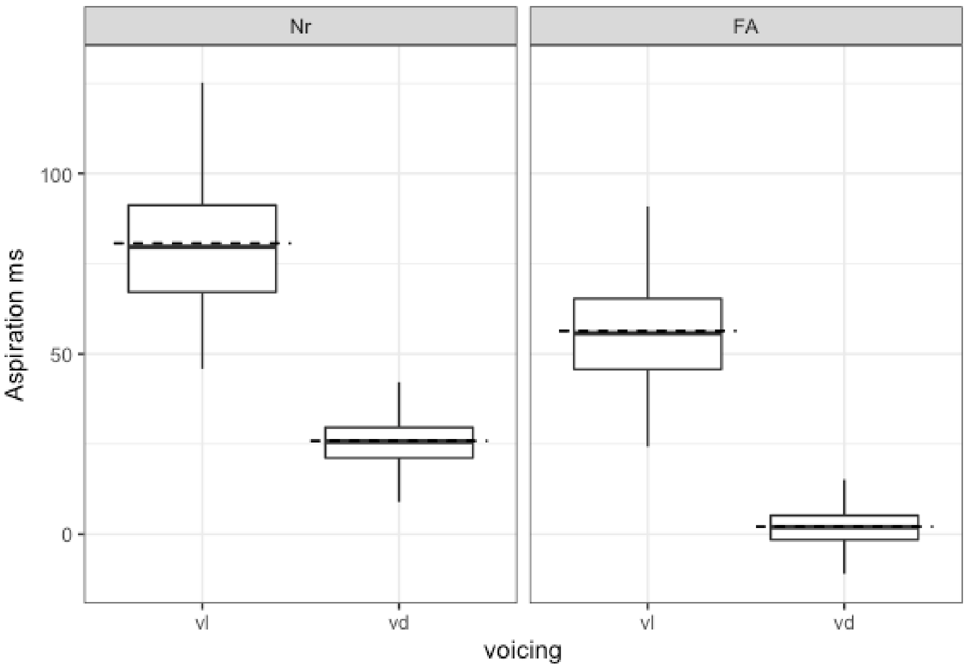


Figure 5.11. Boxplots of the fitted values of aspiration in utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), and speech rate (FA = fast, Nr = Normal). The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	80.7	17.6	25.8	7.12
Fast	56.3	14.2	2.09	5.37

Table 5.11. Mean and standard deviation of the fitted values of aspiration for Voiceless and Voiced utterance-final stops grouped by voicing and speech rate.

As we can see in figure and table 5.11, voicing has a notable impact on aspiration across speech rates. At normal speech, aspiration for Voiceless was significantly longer than for

Voiced stops by an average of 54.9 ms (p<0.0001). As for fast speech, aspiration for the Voiceless was significantly longer than for the Voiced stops by an average of 54.21 ms (p<0.0001). The results also showed large effect of speech rate where aspiration in normal speech was significantly longer than in fast speech by an average of 24.055 for both Voiceless and Voiced stops (p<0.0001).

Aspiration as a function of voicing, rate, and place of articulation

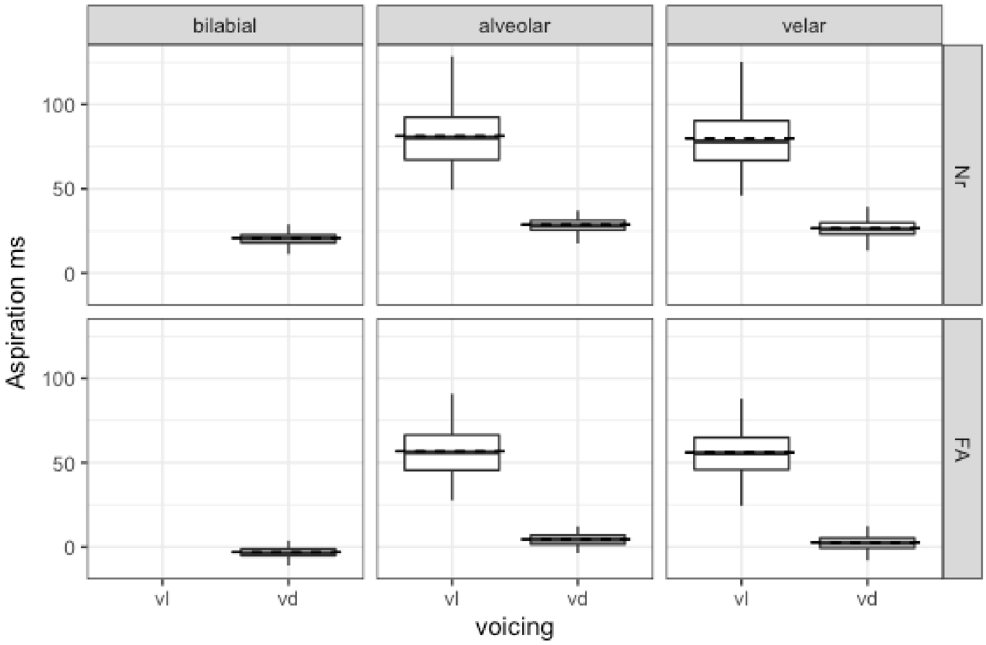


Figure 5.12 Boxplots of the fitted values of aspiration in utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal) and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	20.9	5.98	-	-	3.07	3.88
Alveolar	81.3	17.6	28.3	6.36	56.7	14.6	4.46	4.59
Velar	80	17.6	26.7	6.68	55.9	13.9	2.69	4.80

Table 5.12 Mean and standard deviation of the fitted values of aspiration for Voiceless and Voiced utterance-final stops grouped by speech rate and place of articulation.

As for the impact of place of articulation on aspiration, the pattern appeared to be in the order velar > alveolar > bilabial with marginal differences across voicing status and speech rates. For Voiceless stops, aspiration in velar was longer than in alveolar stops by an average of 1.3 ms in both normal (p =0.0751) and fast (p = 0.3213) rates. Within velar vs alveolar for Voiced stops, aspiration for velar stops showed a significant increase by an average of 1.7 ms in normal (p = 0.0077) and fast (p = 0.0373) rates. Aspiration in velar stops was significantly longer than in bilabial stops by an average of 5 ms in normal speech (p = 002).

5.2.1.5 Summary of results for aspiration

Aspiration results showed that aspiration was a robust acoustic correlate that significantly marked the distinction between Voiceless and Voiced stops in Najdi Arabic with no overlap between the two series of stops across positions and speech rates. The distinction between Voiceless and Voiced stops was maintained across all the included fixed effects.

In normal speech, Voiceless stops were realised with moderate aspiration (MASP) in the majority of the tokens in utterance-initial and utterance-medial iambic stops and to a lesser extent in 41.2% of the tokens in utterance-medial trochaic stops. In utterance-final position, Voiceless stops were realised with heavy aspiration (HASP) in 85.3% of the tokens. The pattern for mean aspiration duration by position in voiceless stops appeared in the order final (m = 80.7 ms) > initial (m = 47.6 ms) > medial iambic (m = 42.1 ms) > medial trochaic (m = 32.4 ms). Voiced stops were realised with short aspiration in all tokens in initial and medial positions and in the majority of tokens in final position.

In fast speech, voiceless stops were realised with moderate aspiration in 57% of the tokens in initial and final stops whereas there were realised with short aspiration in the majority of tokens in medial positions. The positions pattern for mean aspiration duration in Voiceless stops appeared in the order final (m = 56.3 ms) > initial (m = 41.2 ms) > medial iambic (36.8 ms) > medial trochaic (m = 28.7). Voiced stops were realised with short aspiration in all tokens across positions.

The results suggest that Voiceless stops in Najdi Arabic show features of aspirating languages with moderate aspiration in utterance-initial and utterance-medial (iambic) contexts. This pattern of aspiration duration is not as reported for some aspirating languages as German (Jessen, 1998) and English (Kessinger and Blumstein, 1997). The results for aspiration in utterance-final stops show that they are heavily aspirated which is in agreement with the concept of fortition proposed in the work of Iverson and Salmons (1995). The results of speech rate effect showed that aspiration was significantly shortened in response to fast speech in comparison to normal speech and this difference was larger in utterance-final than the other positions.

The results also indicate that Voiceless stops are specified with [spread glottis] because of the presence of moderate aspiration in utterance-initial and utterance-medial (iambic) stops. The impact of speech rate provides support for the activeness of [spread glottis] based on the assumption that aspiration is shortened in specified Voiceless stops in response to fast speech (Beckman et al., 2011).

5.2.2 The proportion of voicing in the closure (% voicing)

This section presents the results of the proportion of voicing in the closure (%voicing) for Voiceless and Voiced stops in utterance-medial and utterance-final positions. This measure is crucial in terms of revealing whether Voiced and Voiceless stops are specified with active features in Najdi Arabic. It is expected for Voiced stops to show robust voicing covering the closure in voicing languages. Voiceless stops, on the other hand, are expected to be actively devoiced in aspirating languages.

As pointed out earlier, 98% of Voiced stops' tokens were fully voiced (100% voicing) in utterance-medial trochaic and iambic contexts whereas 99% of the tokens underwent a devoicing process (100% devoicing) in utterance-final position. A few numbers of Voiceless stops' tokens on the other hand were found to show voicing in the hold phase across positions. The aim of this section is to examine the %voicing in the closure for Voiceless and Voiced stops as a function of rate, place of articulation, vowel type and gender.

5.2.2.1 Utterance-medial stops CV:CVC (Trochaic)

% voicing as a function of voicing and rate

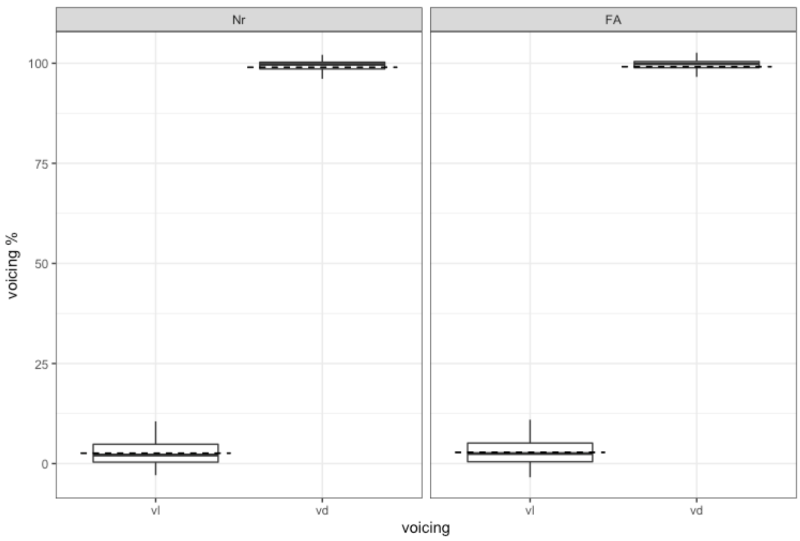


Figure 5.13. Boxplots of the fitted values of %voicing for utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced), and speech rate (FA = fast, Nr = Normal). The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	2.61	2.88	99	2.31
Fast	2.81	2.98	99.2	2.54

Table 5.13. Mean and standard deviation of the fitted values of %voicing for utterance-medial stops (Trochaic) grouped by speech rate.

As we can see in figure and table 5.13, voicing category has a clear impact on the %voicing across speech rates. There was a significant difference in the %voicing for Voiced stops by an average of % 96.39 (p<0.0001) in both normal and fast speech rates. The analysis did not reveal significant impact of speech rate on %voicing.

% voicing as a function of voicing, rate, and place of articulation

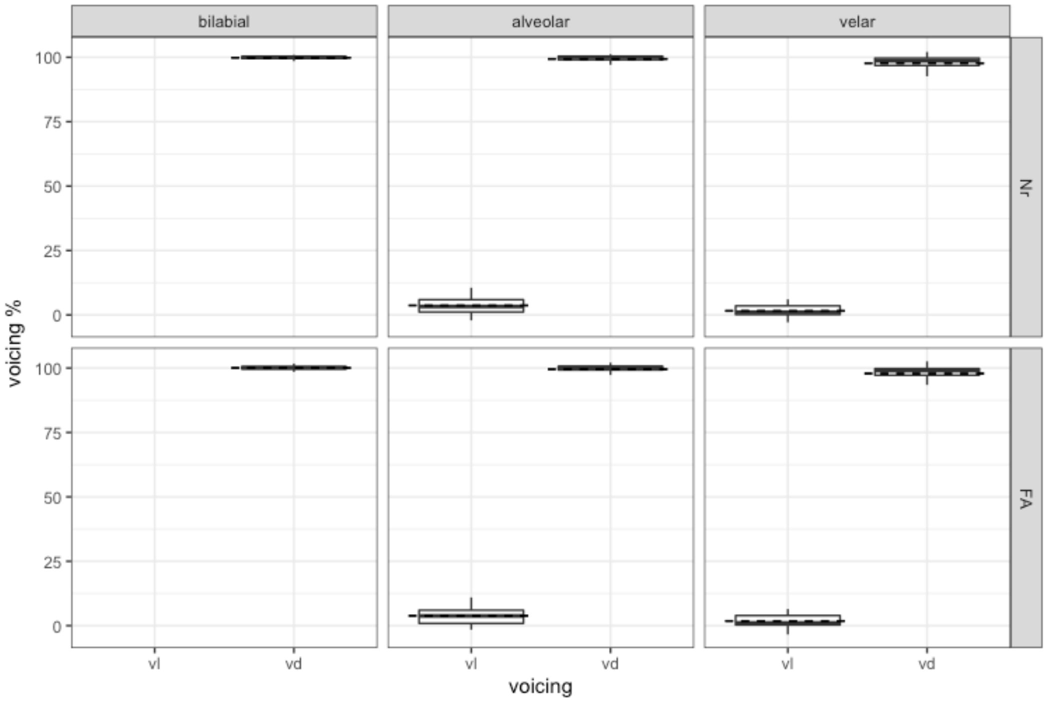


Figure 5.14. Boxplots of the the fitted values of %voicing for utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal) and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-		99.8	0.66	-		100	0.86
Alveolar	3.68	3.04	99.4	1.73	3.84	3.09	99.5	1.79
Velar	1.56	2.25	97.7	3.20	1.81	2.5	97.9	3.47

Table 5.14. Mean and standard deviation of the fitted values of %voicing for utterance-medial stops (Trochaic) grouped by speech rate and place of articulation.

Figure and table 5.14 show that there was a large effect of voicing on the %voicing across speech rates and places of articulation. %voicing was significantly higher for Voiced than for the Voiceless stops by an average of % 95.9 in both normal and fast speech (p<0.0001). With respect to the impact of place of articulation on the %voicing for Voiced stops, the pattern observed was /bilabial/ > /alveolar/ > /velar/ in both normal fast speech rates. The differences between each pair of comparison were statistically significant (pairwise comparison: Appendix C Table C-19).

5.2.2.2 Utterance-medial stops CV<u>C</u>V:C (Iambic)

% voicing as a function of voicing and rate

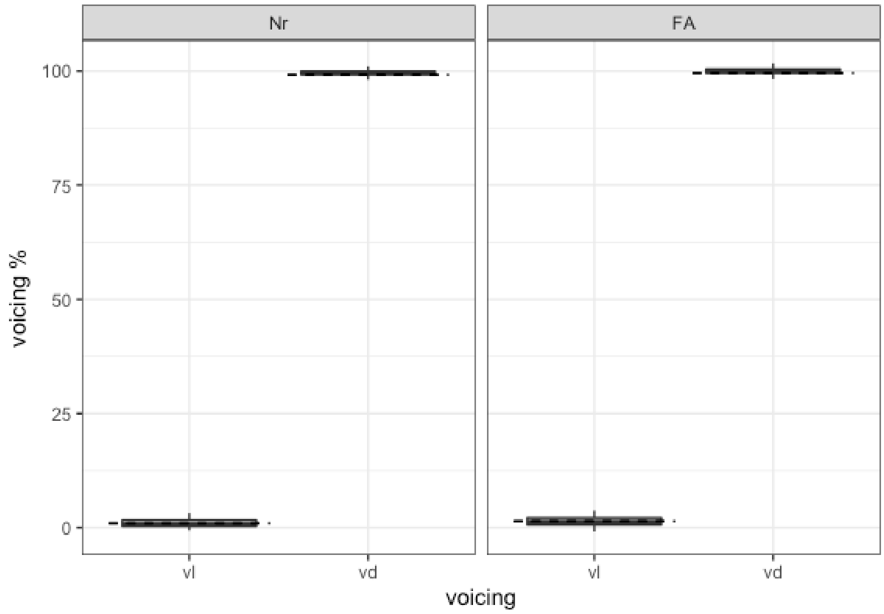


Figure 5.15. Boxplots of the fitted values of %voicing for utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced) and speech rate (FA = fast, Nr = Normal). The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	1.02	0.84	99.2	1.54
Fast	1.43	0.98	99.6	1.63

Table 5.15. Mean and standard deviation of the fitted values of %voicing for utterance-medial stops (Iambic) grouped by speech rate.

Similar to stops in utterance-medial (Trochaic) position, the results in utterance-medial stops (Iambic) show that there was a significant increase of % voicing in Voiced stops by an average of % 98.2 ($p<0.0001$) in both normal and fast speech rates. Moving to the impact of speech rate, there was a significant increase in the %voicing for Voiced stops in fast speech by an average of % 0.4 ($p<0.0001$).

% voicing as a function of voicing, rate, and place of articulation

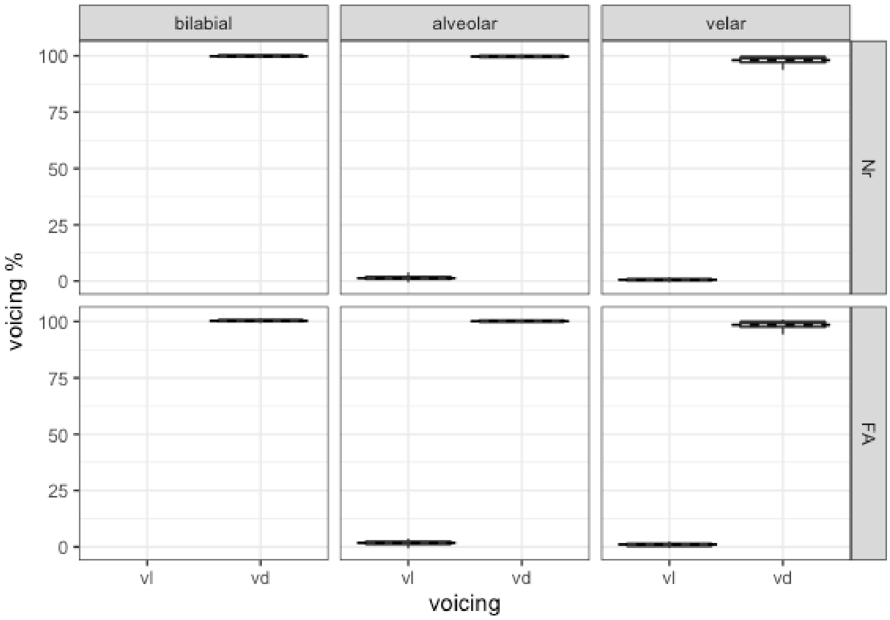


Figure 5.16. Boxplots of the fitted values of %voicing for utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced) and speech rate (FA = fast, Nr = Normal) and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	99.8	0.47	-	-	100	0.63
Alveolar	1.4	0.92	99.7	0.42	1.78	1.08	100	0.57
Velar	0.62	0.5	98	2.15	1.05	0.7	98.4	2.21

Table 5.16. Means and standard deviations of the %voicing for utterance-medial stops (Iambic) grouped by speech rate and place of articulation.

As shown in figure and table 5.16, voicing status has a major effect on the %voicing across speech rates and places of articulation. The %voicing for the Voiced was significantly higher than for the Voiceless stops by an average of % 97.9 in both normal and fast speech ($p<0.0001$). Figure and table 5.16 also show that speech rate marginally affected the %voicing for Voiced and Voiceless stops across places of articulation (Pairwise comparison: Appendix C Table C-22). As for place of articulation, the pattern observed was /bilabial/ > / alveolar/ > /velar/ in normal speech. (Pairwise comparison: Appendix C Table C-22).

5.2.2.3 Utterance-final stops CV:*C*

% voicing as a function of voicing and rate

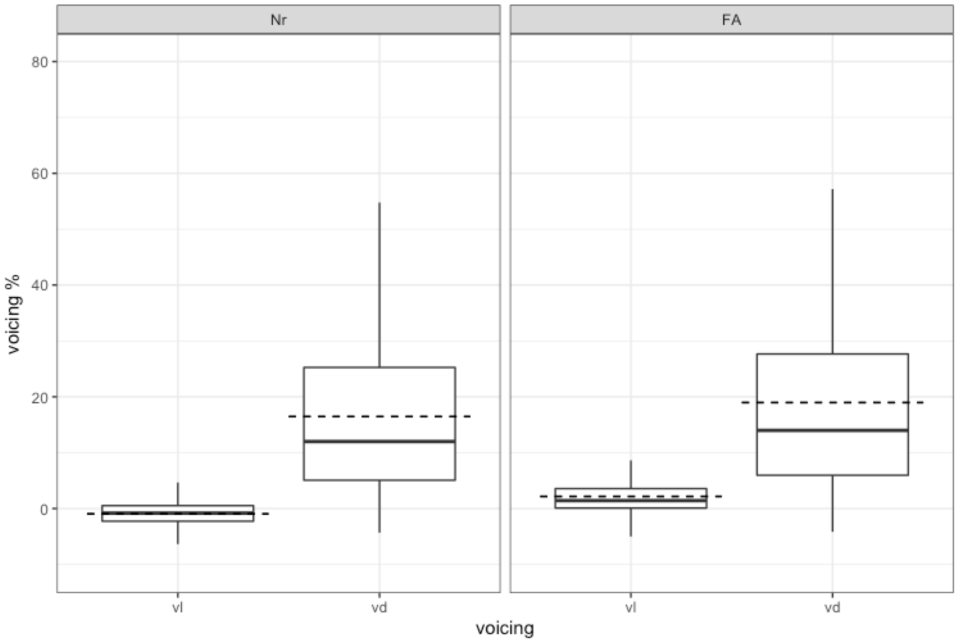


Figure 5.17. Boxplots of the fitted values of %voicing for utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced) and speech rate (FA = fast, Nr = Normal). The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	0.95	2.94	16.5	15.9
Fast	2.18	3.63	19	17.4

Table 5.17. Mean and standard deviation of the fitted values of %voicing for utterance-final stops grouped by speech rate and gender.

Figure and table 5.17 show that there is a significant difference in the %voicing for Voiced stops by an average of 16.2% ($p<0.0001$) in both normal and fast speech rates. Moving to the impact of speech rate on the %voicing, there was a significant increase for Voiced stops in fast speech by an average of 2.5% ($p<0.0001$) and by an average of 1.23% ($p<0.0001$) for Voiceless stops.

% voicing as a function of rate, voicing, and place of articulation

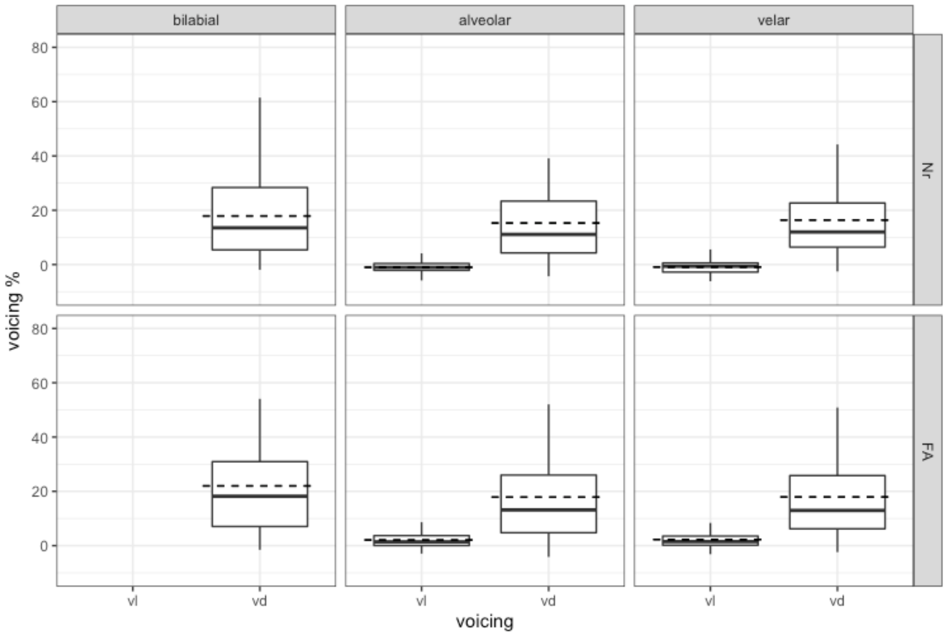


Figure 5.18. Boxplots of the fitted values of %voicing for utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal), and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	17.9	16.5	-	-	22	18.8
Alveolar	1	2.85	15.3	15.6	2.08	3.5	17.9	17
Velar	0.9	3.03	16.4	15.6	2.27	3.75	17.9	16.5

Table 5.18. Mean and standard deviation of the fitted values of %voicing for utterance-final stops grouped by speech rate and place of articulation.

Figure and table 5.18 illustrate the % voicing patterns as a function of voicing, rate, and place of articulation. The analysis revealed that voicing status significantly affected %voicing across speech rates and places of articulation. The %voicing for the Voiced was significantly higher than for the Voiceless stops by an average of 15.3% in both normal and fast speech ($p<0.0001$). Figure 5.29 and table 5.32 also show that speech rate variation significantly affected the %voicing for Voiced bilabial ($p<0.0001$) and alveolar stops ($p = 0.00417$) in which that the fast speech resulted in higher proportion of voicing than the normal speech (pairwise comparison: Appendix C Table C-25). With respect to the impact of place of articulation on the %voicing for Voiced stops, the pattern observed was /bilabial/ > /velar/ > /alveolar/ in normal speech, and /bilabial/ > /velar/ = / alveolar/ in fast speech. Within bilabial vs velar in normal speech, the difference was found to be not significant ($p = 0.10312$) whereas within bilabial vs alveolar the difference was found to be significant ($p = 0.00415$).

In terms of Voiced stops in fast speech, the %voicing for the bilabial was significantly higher than the alveolar and velar stops by an average of 4.1% ($p<0.0001$).

5.2.2.4 Summary of results for %voicing

The results of % voicing indicated that Voiced stops in Najdi Arabic behaved similar to Voiced stops in voicing languages by showing robust voicing in utterance-medial contexts. Voiceless stops on the other hand showed very small percentages of voicing in their closure phase in both utterance-medial and utterance-final stops. Voiced stops in utterance-final context showed a very low percentage of voicing in their hold phase in a small number of tokens (1.5%) which is expected due to the difficulty of maintaining voicing in such a context in both voicing and aspirating languages (Ohala, 1983; Iverson and Salmon, 1995).

With regard to the impact of speech rate, the %voicing tended to be higher in fast speech than in normal speech. This finding was evident for voiced stops more than for voiceless stops. This impact could be interpreted through the fact that reduction taking place in fast speech enabled the extension of voicing to cover larger part of the hold phase.

The results of Voiced stops in utterance-medial contexts indicate an active voicing that implicates the activeness of [voice] in Najdi Arabic based on the predictions proposed in laryngeal realism approach (Beckman et al., 2013; Jessen, 2001). The high percentages of voicelessness in Voiceless stops indicate an active devoicing process that is expected to occur in aspirating languages. This is also supported by the presence of aspiration across contexts which were presented in the previous section.

5.2.3 Prevoicing and voicing duration (ms)

This section presents the results of prevoicing and voicing duration in Najdi Arabic stops with respect to positional variation as well as its interactions with rate, place of articulation, vowel type, and gender. It has been demonstrated earlier in the general results section that 90.5% of Voiced stops' tokens were prevoiced in utterance-initial position and 98% of the tokens showed robust voicing in the closure in utterance-medial positions (V: 100%). Utterance-final stops, however, showed a devoicing process for 59.2% of the tokens (U: 100%). The aim of this section is to check the manifestation of prevoicing and voicing in the closure in each context considering the aforementioned factors. The results for utterance-initial stops are presented only for Voiced stops (no tokens for Voiceless stops showed prevoicing).

5.2.3.1 Utterance-initial stops CV:C

Prevoicing as a function of voicing, rate, and place of articulation

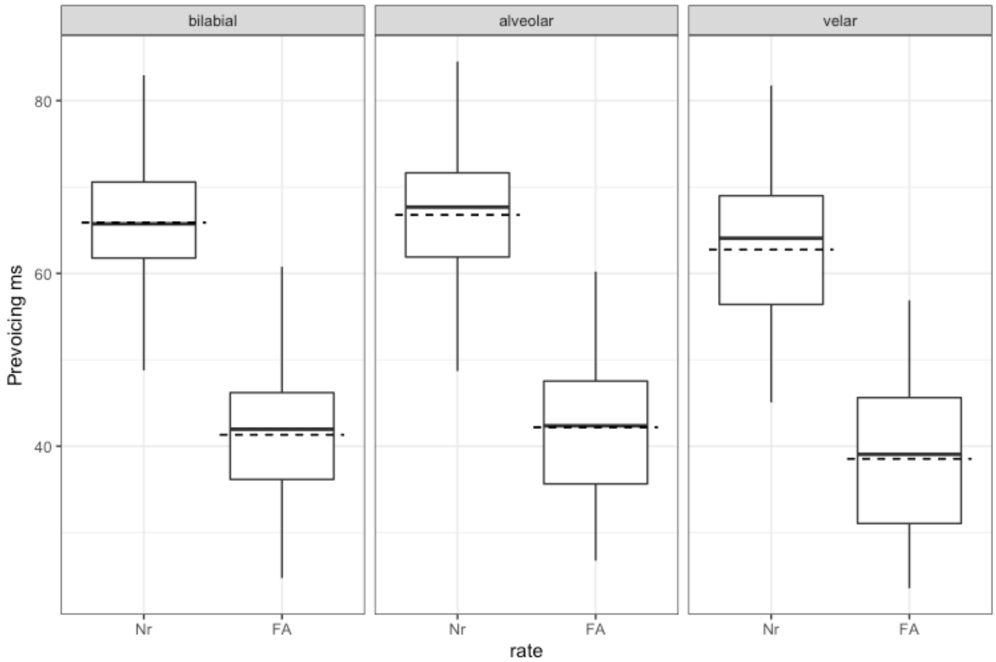


Figure 5.19. Boxplots of the fitted values of prevoicing for Voiced utterance-initial stops classified by speech rate (FA = fast, Nr = Normal), and place of articulation. The dashed line represents the mean.

	Bilabial		**Alveolar**		**Velar**	
	Mean	SD	Mean	SD	Mean	SD
Normal	65.9	8	66.8	7.76	62.7	7.9
Fast	41.3	7.92	42.2	7.72	38.5	8.96

Table 5.19. Mean and standard deviation of the fitted values of prevoicing for Voiced utterance-initial stops grouped by speech rate and place of articulation.

Figure and table 5.19 illustrate the prevoicing values for Voiced stops in the utterance-initial position varied as a function of speech rate across all places of articulation. Beginning with speech rate effect, the analysis revealed that prevoicing was longer in normal speech than in fast speech by an average of 24.4 ms ($p<0.0001$) across places of articulation. As for the impact of place of articulation, the pattern appeared to be in the order alveolar > bilabial > velar in both normal and fast speech rates. Within bilabial vs alveolar, the difference was found to be not significant in both normal ($p = 0.11$) and fast ($p = 0.12$) rates. Within bilabial vs velar, the results showed that prevoicing for bilabial stops were significantly longer than for velar stops by an average of 3.2 ms in normal speech ($p<0.0001$), and by an average of 2.8 ms in fast speech ($p<0.0001$). Within alveolar vs velar for Voiced stops, prevoicing for the former showed a significant increase by an average of 4.1 ms in normal speech (p<0.0001), and by an average of 3.7 ms (p<0.0001) in fast speech.

5.2.3.2 Utterance-medial stops CV:CVC (Trochaic)

Voicing duration as a function of voicing and rate

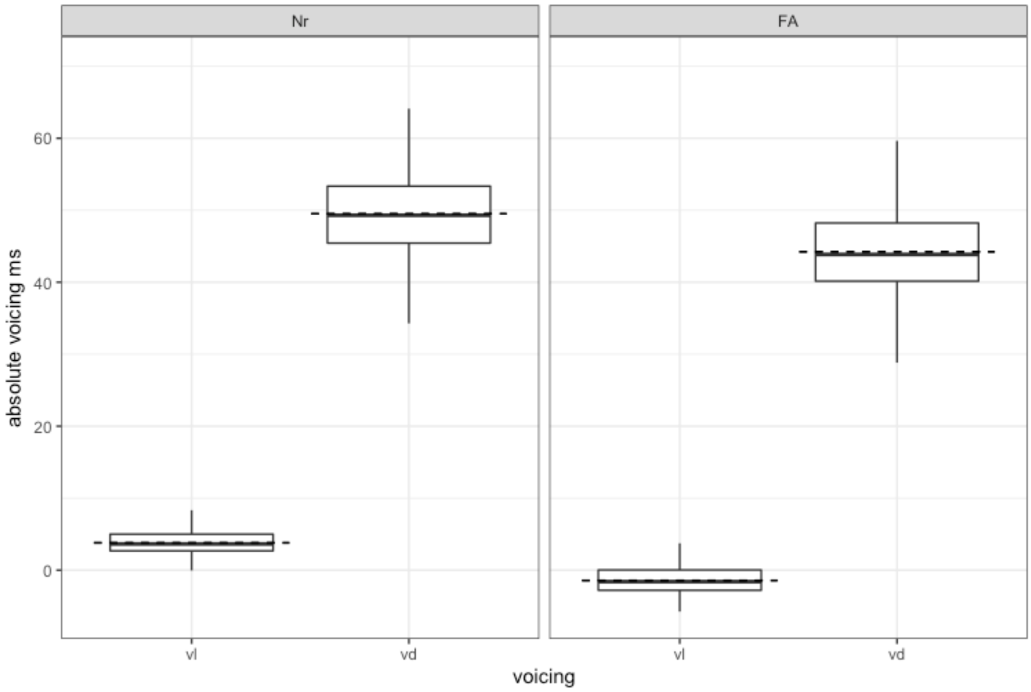


Figure 5.20. Boxplots of the fitted values of voicing duration in utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced) and speech rate (FA = fast, Nr = Normal). The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	3.8	1.6	49.5	5.75
Fast	1.43	1.94	44.2	6.02

Table 5.20. Mean and standard deviation of the fitted values of voicing duration for utterance-medial stops (Trochaic) grouped by speech rate.

The results for voicing duration in utterance-medial stops (trochaic) as a function of voicing and speech rate are shown in figure and table 5.20. Voicing duration for the Voiced was significantly longer than for the Voiceless stops by an average of 45.7 ms (p<0.0001). With respect to fast speech, voicing duration for the Voiced was significantly longer than for the Voiceless stops by an average of 42.77 ms (p<0.0001). Moving to the impact of speech rate on voicing duration, normal speech resulted in longer voicing duration for both Voiceless and Voiced stops. For Voiceless stops within normal vs fast, voicing duration at normal speech showed a significant increase by an average of 2.37 (p<0.0001). A similar pattern was found for Voiced stops whereby voicing duration was significantly longer in normal speech by an average of 5.3 ms ($p<0.0001$).

Voicing duration as a function of voicing, rate, and place of articulation

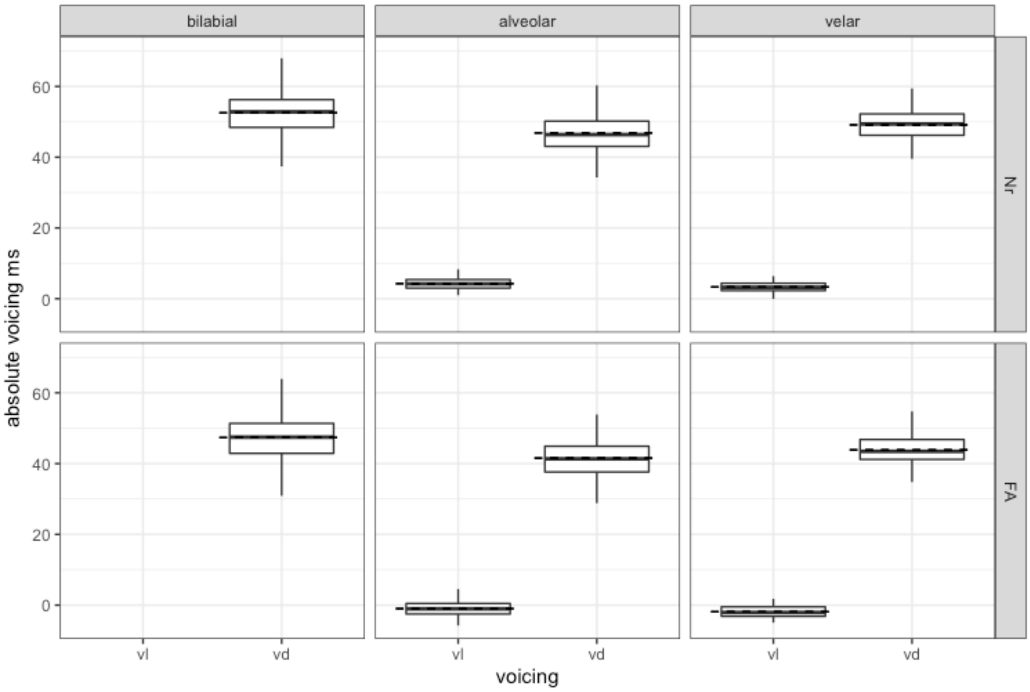


Figure 5.21. Boxplots of the fitted values of voicing duration in utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal), and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-		52.5	6.03	-		47.4	6.47
Alveolar	4.28	1.56	46.9	5.35	0.98	2.12	41.5	5.59
Velar	3.32	1.49	49.1	4.19	1.86	1.64	43.9	4.35

Table 5.21. Mean and standard deviation of the fitted values of voicing duration for utterance-medial stops (Trochaic) grouped by speech rate and place of articulation.

As shown in figure and table 5.21, voicing has a large effect on voicing duration across speech rates and places of articulation. For alveolar stops, voicing duration for the Voiced was significantly longer than for the Voiceless stops by an average of 41.57 ms in both normal and fast speech ($p<0.0001$). For velar stops, voicing duration for the Voiced was significantly longer than for the Voiceless stops by an average of 43.91 ms in both normal and fast speech ($p<0.0001$). Figure and table 5.21 also show that speech rate variation affected voicing duration for Voiced stops across places of articulation in which that the normal speech resulted in longer voicing duration than the fast speech. That is, voicing duration in normal speech was significantly longer than in fast speech by an average of 5.2 ms across places of articulation (p<0.0001). As for the effect of place of articulation, the pattern observed was /bilabial/ > /velar/ > /alveolar/ in both normal and fast speech rates. The differences between each pair of comparison were statistically significant (pairwise comparison: Appendix C Table C-30).

5.2.3.3 Utterance-medial stops CV<u>C</u>V:C (Iambic)

Voicing duration as a function of voicing and rate

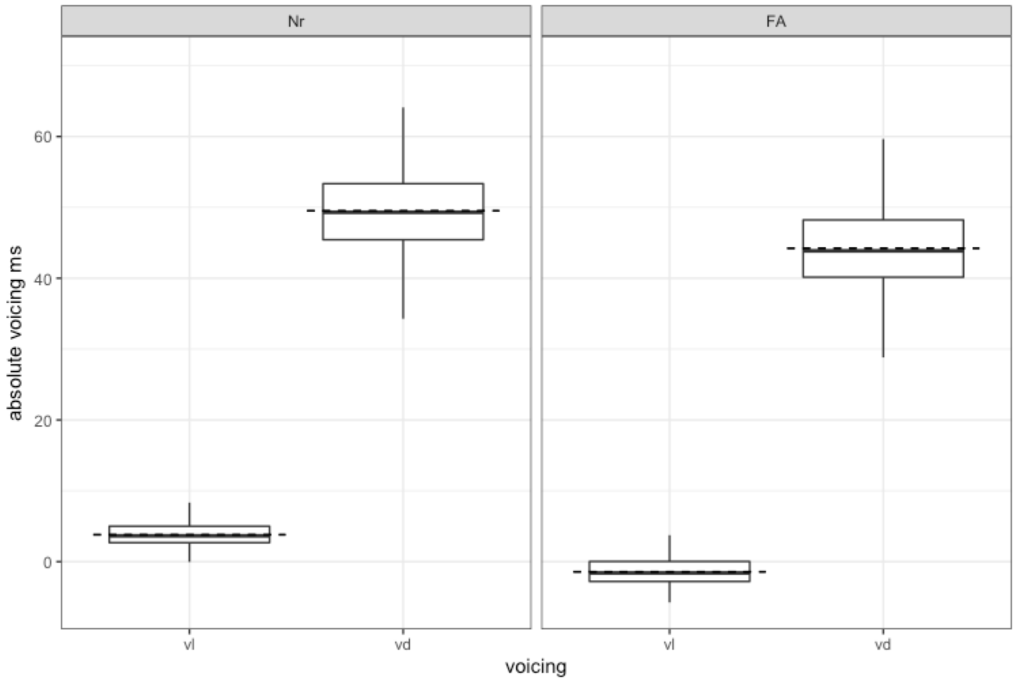


Figure 5.22. Boxplots of the fitted values of voicing duration for utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced) and speech rate (FA = fast, Nr = Normal). The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	1.1	5.29	51.3	11.1
Fast	1.4	5.59	42.4	10.7

Table 5.22. Mean and standard deviation of the fitted values of voicing duration for utterance-medial stops (Iambic) grouped by speech rate.

The analysis of voicing duration revealed that Voiceless and Voiced stops are distinct with no overlap between the two categories. Figure and table 5.22 show that voicing duration for the Voiced was significantly longer than for the Voiceless stops by an average of 50.2 ms in normal speech (p<0.0001) and by an average of 41ms (p<0.0001) in fast speech. Moving to the effect of speech rate on voicing duration, normal speech resulted in longer voicing duration in Voiced stops by an average of 8.9 ms (p<0.0001). The difference between voicing duration for Voiceless stops in normal and fast speech was found to be not significant (p = 0.98345).

Voicing duration as a function of voicing, rate, and place of articulation

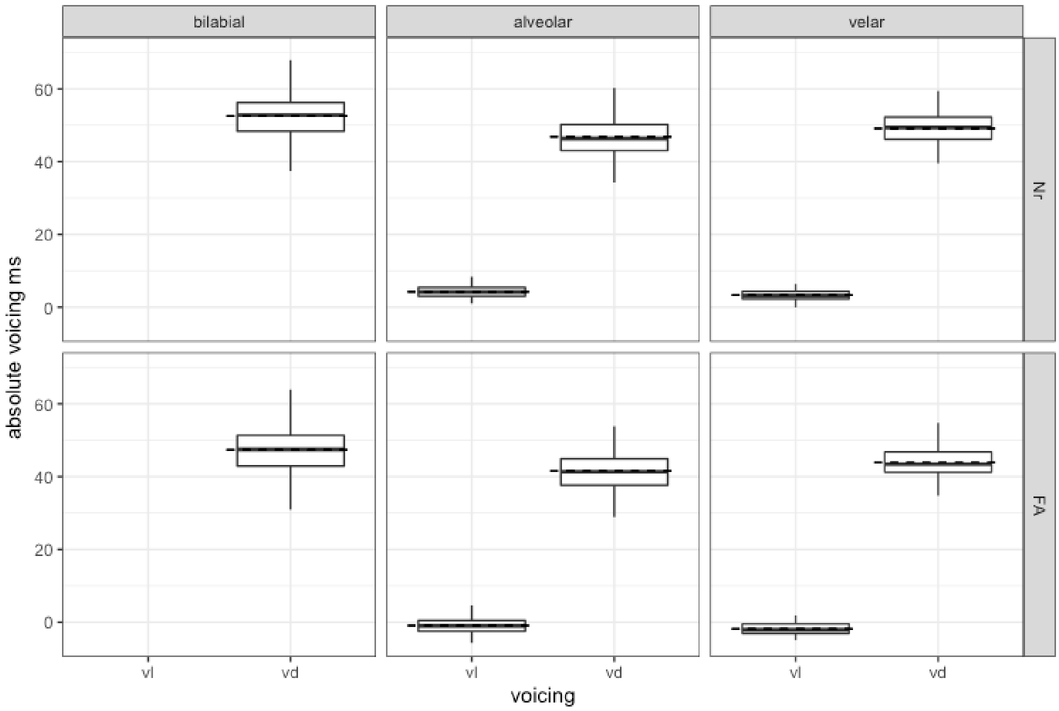


Figure 5.23. Boxplots of the fitted values of voicing duration for utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal), and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-		55.1	11.6	-		44.8	11.6
Alveolar	1.41	5.01	48.6	10.5	2.06	6.57	39.6	9.97
Velar	0.785	5.55	50.3	10.1	0.772	4.36	42.8	9.82

Table 5.23. Mean and standard deviation of the fitted values of voicing duration for utterance-medial stops (Iambic) grouped by speech rate and place of articulation.

As we can see in figure and table 5.23, voicing has a major effect on voicing duration across speech rates and places of articulation. For alveolar stops, voicing duration for the Voiced was significantly longer than for the Voiceless stops by an average of 47.19 ms in normal speech (p<0.0001) and by an average of 37.54 ms in fast speech (p<0.0001). For velar stops, voicing duration for the Voiced was significantly longer than for the Voiceless stops by an average of 49.515 ms in normal speech (p<0.0001) and by an average of 42.028 ms in fast speech (p<0.0001). In terms of the effect of speech rate on voicing duration, figure 5.40 and table

5.43 also show that speech rate affected voicing duration for Voiced stops across places of articulation in which that the normal speech resulted in longer voicing duration than the fast speech. Voicing duration in normal speech was significantly longer than in fast speech by an average of 8.9 ms ($p<0.0001$) across places of articulation. With respect to the impact of place of articulation on voicing duration, the pattern appeared to be in the order bilabial > velar > alveolar across speech rates. All the differences between each pair of comparison for voicing duration as a function of place of articulation were found to be significant (pairwise comparison: Appendix C Table C-32).

5.2.3.4 Utterance-final stops CV:*C*

Voicing duration as a function of voicing and rate

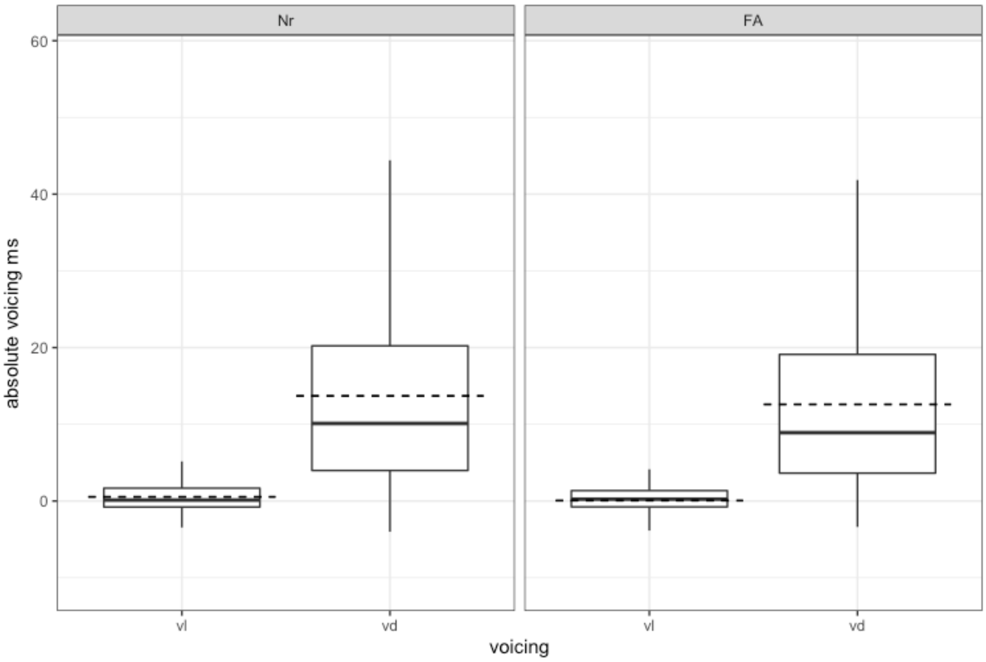


Figure 5.24. Boxplots of the fitted values of voicing duration for utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced) and speech rate (FA = fast, Nr = Normal). The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	0.54	2.13	13.7	13
Fast	0.058	2.45	12.6	11.8

Table 5.24. Mean and standard deviation of the fitted values of voicing duration for utterance-final stops grouped by speech rate.

Figure and table 5.24 show that voicing duration for the Voiced was significantly longer than for the Voiceless stops by an average of 12.8 ms ($p<0.0001$) in both normal and fast speech rates. Moving to speech rate impact on voicing duration for Voiced stops, voicing duration in normal speech was significantly longer than in fast speech by an average of 1.1 ms (p = 0.0069).

Voicing duration as a function of voicing, rate and place of articulation

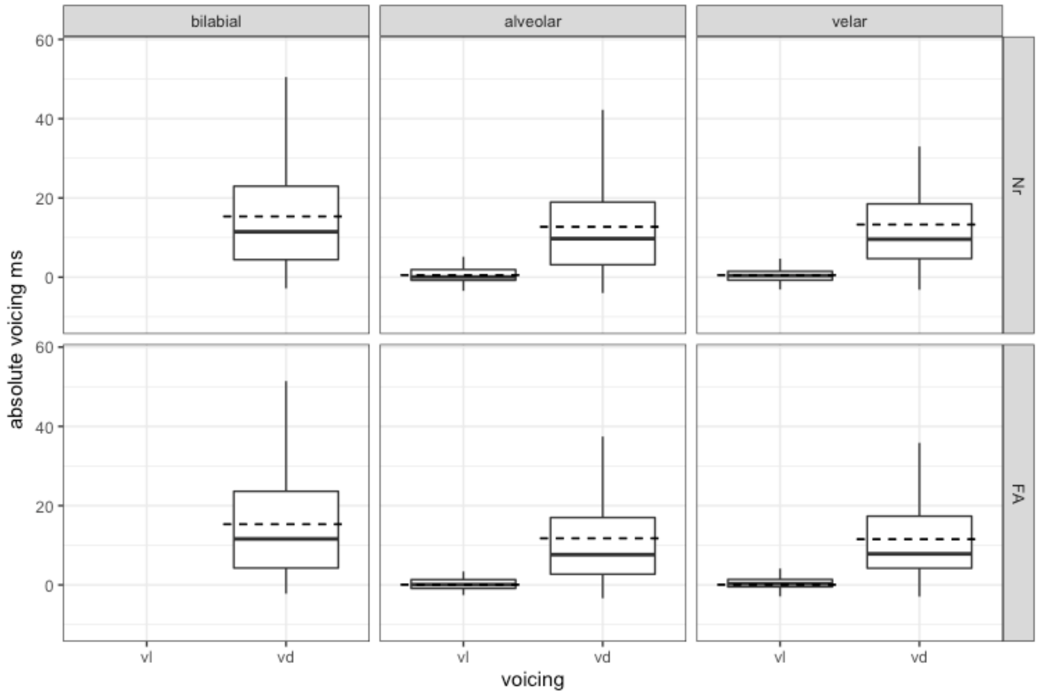


Figure 5.25. Boxplots of the fitted values of voicing duration for utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal), and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	15.3	14	-	-	15.3	13.3
Alveolar	0.55	2.22	12.7	12.7	0.066	2.39	11.8	11.6
Velar	0.54	2.03	13.3	12.3	0.051	2.52	11.5	10.6

Table 5.25. Mean and standard deviation of the fitted values of voicing duration for utterance-final stops grouped by speech rate and place of articulation.

As shown in figure and table 5.25, the analysis revealed that voicing duration for the Voiced was significantly longer than for the Voiceless by an average of 12.1 ms in both normal and fast speech ($p<0.0001$). The change of speech rate did not significantly affect voicing duration for Voiceless and Voiced stops across places of articulation; there were marginal differences in favour of normal speech (pairwise comparison: Appendix C Table C-36). Moving to the impact of place of articulation on voicing duration for Voiced stops, the pattern appeared to be in the order bilabial > velar > alveolar. Voicing duration values for the bilabial were significantly longer than for the velar (p = 0.00242) and the alveolar (p = 0.00011) stops. Within velar vs alveolar for Voiced stops, the difference was found to be not significant in both normal (p = 0.4544) and fast (p = 0.78741) rates. The differences between voicing duration for Voiceless stops were found to be not significant across places of articulation (pairwise comparison: Appendix C Table C-36).

5.2.3.5 Summary of results for prevoicing and voicing duration in the closure

Starting with prevoicing, it has been demonstrated that prevoicing is a robust acoustic correlate that significantly signals the distinction between Voiced and Voiceless stops in Najdi Arabic with no overlap between the two categories in utterance-initial position across speech rates. Voiced stops were prevoiced in 90.5% of the tokens in normal speech and in 86% of the tokens in fast speech. In terms of speech rate effect on prevoicing, normal speech resulted in significantly longer prevoicing in utterance-initial stops. The mean duration for prevoicing in normal speech was 65.1 ms, and it was 40.6 ms in fast speech.

As we have seen, voicing duration for utterance-medial stops behaved similarly in both trochaic and Iambic structure in which that Voiced stops were clearly distinguished from Voiceless stops by voicing duration in the majority of the tokens across places of articulation, speech rates, genders, and vowel types. Voiceless stops were realised with no voicing in the majority of tokens in utterance-medial and utterance-final positions. Voicing duration for utterance-final stops was notably less effective in marking the voicing distinction due to the devoicing process expected in such a context.

These findings are in agreement with what has been reported in the literature regarding the voicing contrast in voicing languages in terms of prevoicing in utterance-initial stops, voicing duration in utterance-medial stops, and final devoicing in utterance-final stops. Prevoicing results are similar to that of Russian (Ringen and Kulikov, 2012) and Dutch (Van Alphen and Smits, 2004). The results of speech rate effect supports the assumption that Voiced stops are specified with [voice] because of the significant increase in response to normal speech.

5.3 Conclusion

This chapter has presented the results of the LMM statistical analysis of the patterns of voicing and aspiration in Voiceless and Voiced stops in Najdi Arabic in utterance-initial, utterance-medial, and utterance-final contexts under normal/fast conditions. The acoustic correlates investigated in this chapter included aspiration, %voicing, prevoicing, and voicing duration.

The following table is an overview of the findings.

		Utterance-initial			Utterance-medial (trochaic)			Utterance-medial (iambic)			Utterance-final		
Correlate	**Voicing**	**rate**	**M**	**SD**	**rate**	**M**	**SD**	**rate**	**M**	**SD**	**rate**	**M**	**SD**
Aspiration	Vl	Nr	47.9	7.19	Nr	32.4	5.68	Nr	42.1	10.5	Nr	80.7	17.6
	Vd	Nr	11.9	3.92	Nr	9.59	3.38	Nr	11.4	3.59	Nr	25.8	7.12
% Voicing	Vl	Nr	…..	…..	Nr	2.61	2.88	Nr	1.02	0.84	Nr	0.95	2.94
	Vd	Nr	…..	…..	Nr	99	2.31	Nr	99.2	1.54	Nr	16.5	15.9
Prevoicing	Vl	Nr	…..	…..	Nr	…..	…..	Nr	…..	…..	Nr	…..	…..
	Vd	Nr	65.1	8	Nr	…..	…..	Nr	…..	…..	Nr	…..	…..
Voicing duration	Vl	Nr	…..	…..	Nr	3.8	1.6	Nr	1.1	5.29	Nr	0.54	2.13
	Vd	Nr	…..	…..	Nr	49.5	5.72	Nr	51.3	11.1	Nr	13.7	13
Aspiration	Vl	FA	41.2	7.31	FA	28.7	5.62	FA	36.8	10.4	FA	56.3	14.2
	Vd	FA	5.56	3.42	FA	6.19	3.05	FA	6.09	3.48	FA	2.09	5.37
% Voicing	Vl	FA	…..	…..	FA	2.81	2.98	FA	1.43	0.98	FA	2.18	3.63
	Vd	FA	…..	…..	FA	99.2	2.54	FA	99.6	1.63	FA	19	17.4
Prevoicing	Vl	FA	…..	…..	FA	…..	…..	FA	…..	…..	FA	…..	…..
	Vd	FA	40.6	8.2	FA	…..	…..	FA	…..	…..	FA	…..	…..
Voicing duration	Vl	FA	…..	…..	FA	1.43	1.94	FA	1.4	5.59	FA	0.05	2.45
	Vd	FA	…..	…..	FA	44.2	6.02	FA	42.4	10.7	FA	12.6	11.8

Table 5.26. Overview of the acoustic measures (Mean and standard deviation) of the patterns of voicing and aspiration in Voiced and Voiceless stops under normal/fast speech rates.

The main purpose of investigating the acoustic measures for voicing and aspiration was to find out how Voiced and Voiceless stops behaved in different phonetic contexts in terms of the manifestations of the voicing contrast. This investigation enabled us to provide an answer to the question related to the phonetic side proposed in the realm of laryngeal realism in which [voice] requires active voicing whereas [spread glottis] requires long lag aspiration.

The results for aspiration and voicing show that both Voiceless and Voiced stops are fully specified. Voiceless stops were realised with long lag aspiration whereby Voiced stops were realised with active voicing in utterance-initial and utterance-medial positions. The distinction between Voiceless and Voiced stops was maintained across all sources of variability including place of articulation, vowel type and gender. The impact of place of articulation, vowel type and gender on the acoustic measure of voicing and aspiration were small although it is true some of the differences were statistically significant.

The results in this chapter also demonstrate that both voicing and aspiration were influenced in response to the change in speech rate. That is, fast speech rate resulted in shorter aspiration and shorter voicing which strengthens the argument for the assumption that specified stops are expected to be affected by the change in speech rate but not the unspecified ones.

The results are consistent with the primary assumption that both [voice] and [spread glottis] are active in the voicing system of Najdi Arabic. The presence of active voicing indicates an active production gesture which includes the vibration of the vocal folds and the articulatory adjustments that lead to initiating and maintaining this vibration in a challenging context such as utterance-initial. On the other hand, the presence of long lag aspiration is an indication for an active glottal opening that allowed for the delay of the voicing of the following vowel to be initiated.

Chapter 6. Results for the durational correlates of voicing contrast

In the previous chapter, the acoustic analysis revealed evidence for the robust presence of voicing and aspiration in Voiced and Voiceless stops, respectively. In connection with the results in the previous chapter, the investigation in this chapter is a continuation of the acoustic analysis of the correlates that are expected to cue the opposition between Voiceless and Voiced stops in Najdi Arabic. As confirmed by many studies cross-linguistically, the voicing contrast in stops is cued by multiple durational and spectral acoustic correlates in addition to VOT. It has also been found that voicing and aspirating languages differ in their use of these correlates based on the phonetic characteristics of Voiced and Voiceless stops in these languages. The manifestation of all the acoustic correlates enhances our understanding of the phonetic and phonological aspects of voicing contrast in Najdi Arabic which shows features from both voicing and aspirating languages.

This chapter presents the results for the durational acoustic correlates that include closure duration, burst duration, preceding vowel duration, and following vowel duration. The results of the LMM analysis are presented for each acoustic correlate in each position considering different factors. The chapter is divided into five parts: 1) closure duration, 2) burst duration, 3) preceding vowel duration, 4) following vowel duration, and 5) summary of the results for all the durational correlates under normal/fast speech rates.

6.1 Closure Duration (CD)

This section reports the closure duration values for Voiceless and Voiced stops in Najdi Arabic considering various contexts and factors. Closure duration has been employed in the acoustic literature to differentiate Voiceless and Voiced stops. The hold phase is typically longer in Voiceless stops than in Voiced stops across positions (Lisker, 1957; Kohler, 1984; Kluender et al., 1988). Some studies showed that this pattern is more consistent in voicing languages than in aspirating languages (Jessen, 2001). The main purpose of investigating closure duration in Najdi Arabic is to find out whether CD differs as a function of voicing category and how this differentiation is implemented.

6.1.1 Utterance-medial stops CV:<u>C</u>VC (Trochaic)

Closure duration as a function of voicing and rate

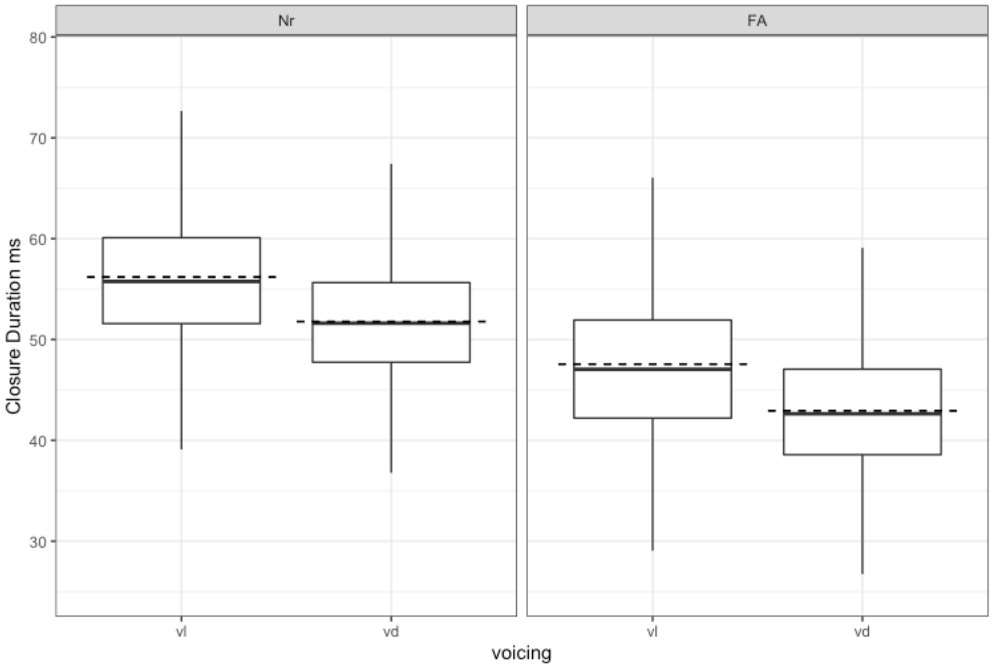


Figure 6.1. Boxplots of the fitted values of closure duration for utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced) and speech rate (FA = fast, Nr = Normal). The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	56.2	6.94	51.8	5.81
Fast	47.6	7.74	42.9	6.49

Table 6.1. Mean and standard deviation of the fitted values of closure duration for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

The results for closure duration in utterance-medial stops (Trochaic) as a function of voicing and rate are given in figure and table 6.1. Within Voiceless vs Voiced, there was a significant increase in the closure duration for Voiceless stops by an average of 4.55 ms ($p<0.0001$) in both normal and fast speech rates with some overlap between the two categories. With regard to the impact of speech rate on the closure duration, closure duration in the normal speech was significantly longer than in the fast speech by an average of 8.7 ms ($p<0.0001$) for both Voiceless and Voiced stops.

Closure duration as a function of voicing, rate, and place of articulation

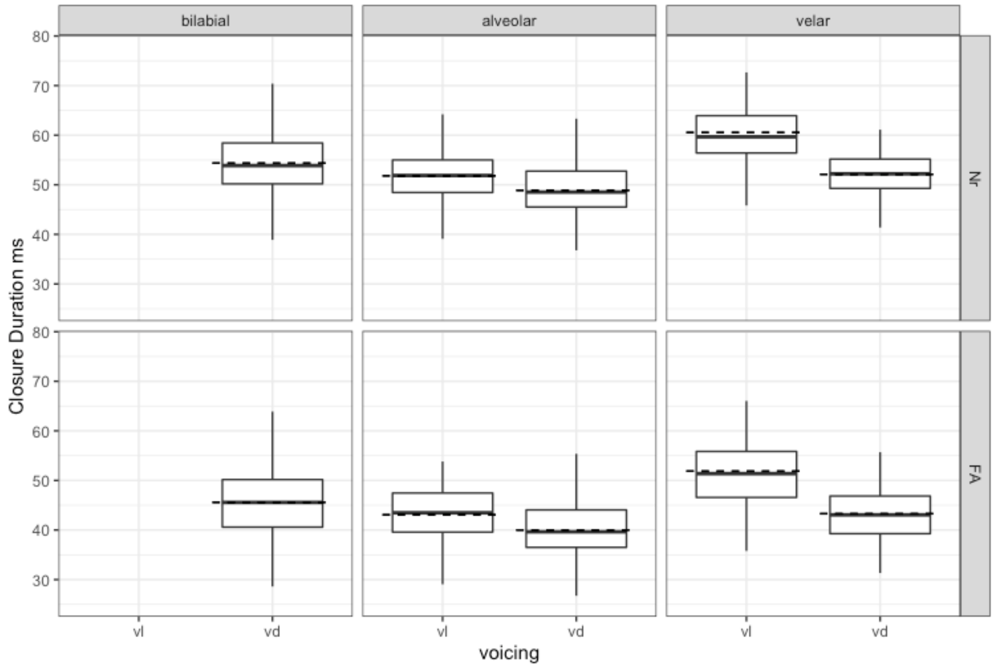


Figure 6.2. Boxplots of the fitted values of closure duration for utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal), and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	54.4	6.33	-	-	45.6	7.03
Alveolar	51.8	4.69	48.9	5.13	43.1	5.61	40	5.74
Velar	60.6	6.01	52.1	4.39	51.9	7.03	43.3	5.34

Table 6.2. Mean and standard deviation of the fitted values of closure duration for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate, and place of articulation.

As shown in figure and table 6.2, the voicing category affected the closure duration across speech rates and places of articulation. For alveolar stops, the closure duration for the Voiceless was significantly longer than for the Voiced by an average of 3 ms in both normal and fast speech (p<0.0001). For velar stops, the closure duration for the Voiceless was significantly longer than for the Voiced by an average of 8.55 ms in both normal and fast speech (p<0.0001). The results also show that speech rate variation significantly affected the closure duration for Voiced and Voiceless stops across places of articulation in which that the normal speech resulted in longer closure duration than the fast speech by an average of 8.78 ms (p<0.0001). With regard to the impact of place of articulation on the closure duration for Voiceless stops, the pattern observed was /velar/ > /alveolar/ in both normal fast speech by an average of 8.8 ms (p<0.0001). The pattern observed for Voiced stops on the other hand was /bilabial/ >/velar/ > /alveolar/ in normal speech and the differences between each pair of comparison were statistically significant (pairwise comparison: Appendix C Table C-39). As for fast speech, the pattern observed was /bilabial/ >/alveolar/ > /velar/. The differences between each pair of comparison were statistically significant (pairwise comparison: Appendix C Table C-39).

6.1.2 Utterance-medial stops CV<u>C</u>V:C (Iambic)

Closure duration as a function of voicing and rate

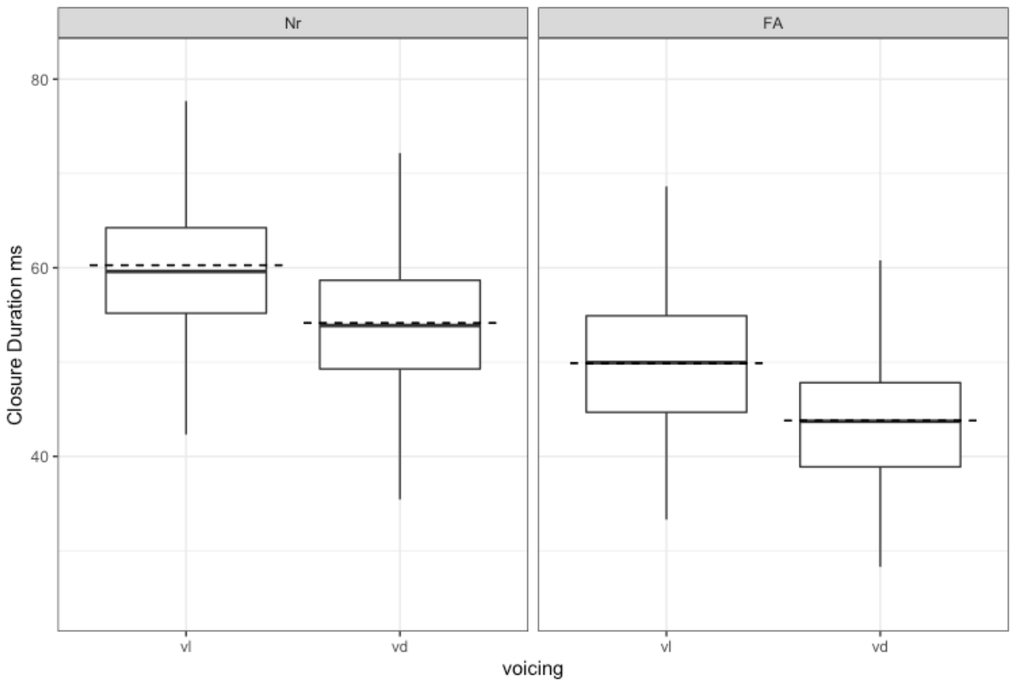


Figure 6.3. Boxplots of the fitted values of closure duration for utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced), and speech rate. The dashed line represents the mean.

	Voiceless		Voiced	
	Mean	SD	Mean	SD
Normal	60.3	7.39	54.1	6.55
Fast	49.9	7.4	43.8	6.8

Table 6.3. Means and standard deviation of the fitted values of closure duration for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

As figure and table 6.3 show, the analysis revealed that voicing category significantly affected the closure duration across speech rates with some overlap between the two categories. Within Voiceless vs Voiced, there was a significant difference in the closure duration for Voiceless stops by an average of 6.2 ms ($p<0.0001$) in both normal and fast speech rates. With regard to the impact of speech rate on the closure duration, closure duration in the normal speech was significantly longer than in the fast speech by an average of 10.2 ms ($p<0.0001$) for both Voiceless and Voiced stops.

Closure duration as a function of voicing, rate, and place of articulation

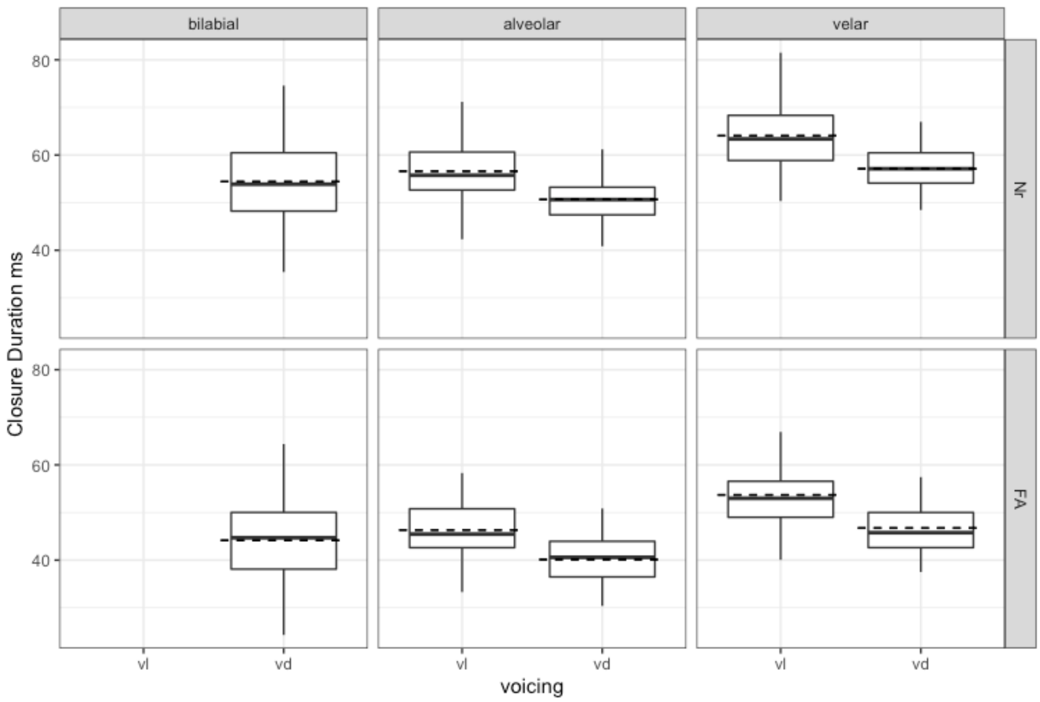


Figure 6.4. Boxplots of the fitted values of closure duration for utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and place of articulation. The dashed line represents the mean.

	Voiceless normal		Voiced normal		Voiceless fast		Voiced fast	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	54.4	8.36	-	-	44.2	8.17
Alveolar	56.6	5.95	50.7	4.42	46.3	5.98	40.1	4.63
Velar	64.1	6.81	57.1	4.19	53.7	6.87	46.7	5.3

Table 6.4. Mean and standard deviation of the fitted values of closure duration for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and place of articulation.

As shown in figure and table 6.4, the voicing category affected the closure duration across speech rates and places of articulation. The closure duration for the Voiceless was significantly longer than for the Voiced by an average of 6.5 ms in both normal and fast speech ($p<0.0001$). Speech rate variation significantly affected the closure duration for Voiced and Voiceless stops across places of articulation in that normal speech resulted in longer closure duration than fast speech by an average of 10.4 ms ($p<0.0001$). With regard to the impact of place of articulation on the closure duration for Voiceless stops, the pattern observed was /velar/ > /alveolar/ in both normal fast speech by an average of 7.45 ms ($p<0.0001$). The pattern observed for Voiced stops on the other hand was /velar/ >/bilabial/ > /alveolar/ in both normal and fast speech rates. The differences between each pair of comparison were statistically significant (pairwise comparison: Appendix C Table C-44).

6.1.3 Utterance-final stops CV:C

Closure duration as a function of voicing and rate

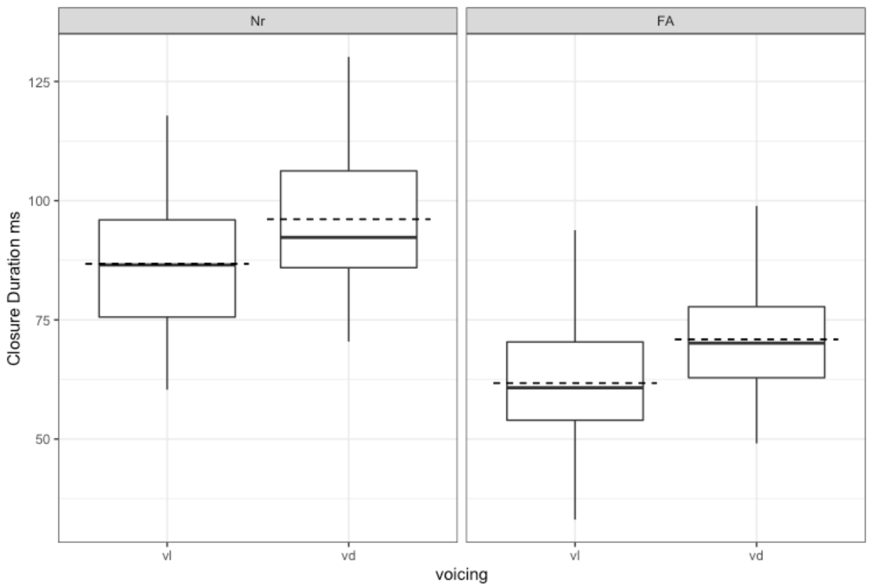


Figure 6.5. Boxplots of the fitted values of closure duration for utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), and speech rate. The dashed line represents the mean.

	Voiceless		Voiced	
	Mean	SD	Mean	SD
Normal	86.8	13.1	96.1	13.2
Fast	61.8	11.8	70.9	11.2

Table 6.5. Mean and standard deviation of the fitted values of closure duration for Voiceless and Voiced utterance-final stops grouped by voicing and speech rate.

Unlike the distinction in utterance-medial stops, closure duration in utterance-final stops showed the opposite pattern. There was a significant increase in the closure duration for Voiced stops by an average of 9.2 ms (p<0.0001) in both normal and fast speech rates. With regard to the impact of speech rate on the closure duration, closure duration in normal speech was significantly longer than in fast speech by an average of 25.1 ms (p<0.0001) for both Voiceless and Voiced stops.

Closure duration as a function of voicing, rate, and place of articulation

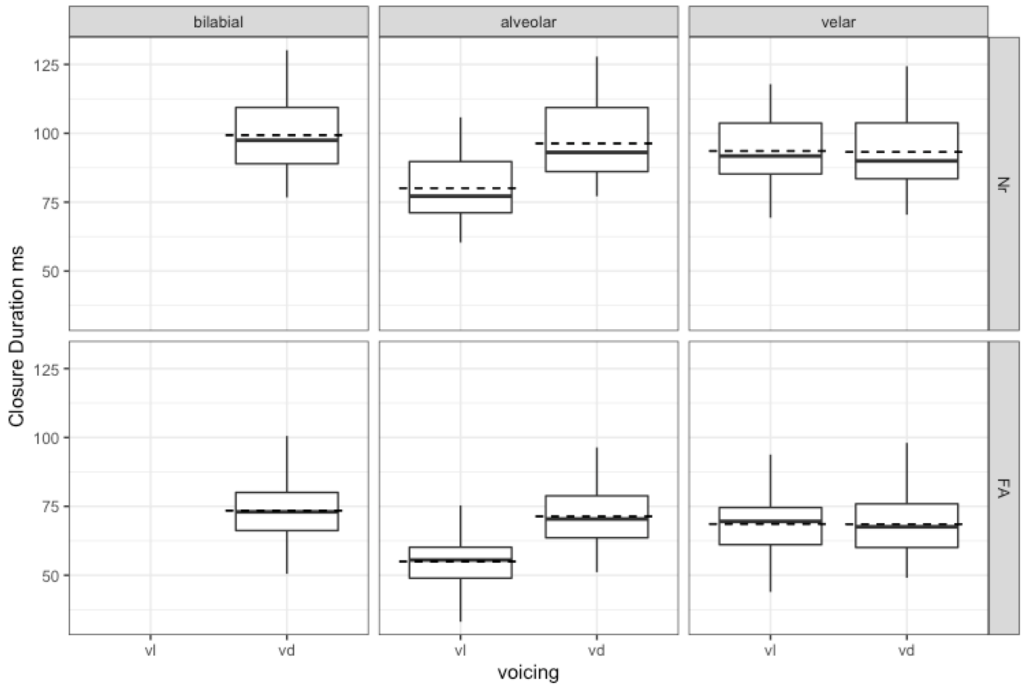


Figure 6.6. Boxplots of the fitted values of closure duration for utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and place of articulation. The dashed line represents the mean.

	Voiceless normal		Voiced normal		Voiceless fast		Voiced fast	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	99.3	13	-	-	73.5	11.8
Alveolar	80.1	11	96.3	13.3	55	9.26	71.4	11.2
Velar	93.6	11.6	93.2	12.6	68.6	10.1	68.6	10.5

Table 6.6. Mean and standard deviation of the fitted values of closure duration for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and place of articulation.

As figure and table 6.6 show, the voicing category affected closure duration across speech rates in alveolar stops. That is, the closure duration for the Voiced was significantly longer than for the Voiceless by an average of 16.3 ms in both normal and fast speech ($p<0.0001$). For velar stops, however, the difference between closure duration for Voiceless and Voiced stops was found to be not significant in both normal ($p = 0.64204$) and fast ($p = 0.95378$) speech rates. The results also show that speech rate variation significantly affected the closure duration for Voiced and Voiceless stops across places of articulation. Normal speech resulted in longer closure duration than the fast speech by an average of 25.08 ms ($p<0.0001$) for Voiced and Voiceless stops across places of articulation. With regard to the impact of place of articulation on the closure duration for Voiceless stops, the pattern observed was /velar/ > /alveolar/ in both normal and fast speech by an average of 13.55 ms ($p<0.0001$). The pattern observed for Voiced stops on the other hand was /bilabial/ >/alveolar/ > /velar/ in both normal and fast speech rates. The differences between each pair of comparison were statistically significant (pairwise comparison: Appendix C Table C-47).

6.1.4 Summary of results for closure duration

The results showed that **closure duration** in Voiceless stops was significantly longer than closure duration in Voiced stops in utterance-medial contexts in both normal and fast speech rates with some overlap between the two categories. Furthermore, closure duration appeared to be longer in normal speech than in fast speech across positions and voicing categories.

Surprisingly, closure duration for utterance-final stops showed the opposite pattern: closure duration was longer in Voiced stops than in Voiceless stops across speech rates. This distinction was larger in alveolar stops with 16 ms difference in favour of Voiced stops. In addition, the results showed a longer closure duration for utterance-final stops than for utterance-medial stops across voicing categories. Moreover, more variability was noted in utterance-final stops with relatively higher standard deviations. Looking more closely at the values of closure duration for utterance-final stops, it could be suggested that the shortening of closure duration for voiceless stops in comparison to voiced stop was a by-product of lengthening of aspiration that occurred in the majority of Voiceless stops' tokens assuming the enhancement mechanism proposed by Jessen (2001).

6.2 Burst duration (ms)

This section presents the results of burst duration for Voiceless and Voiced stops in Najdi Arabic with respect to rate, place of articulation, vowel type, and gender. It has been proposed that burst duration signals the distinction between Voiced and Voiceless stops. That is, burst duration for Voiceless is longer than for Voiced stops (Halle et al., 1957; Zue, 1976; Lavoie, 2001). The results in this section shed light on burst duration patterns in Najdi Arabic taking into account various contexts.

6.2.1 Utterance-initial stops C̲V:C

Burst duration as a function of voicing and rate

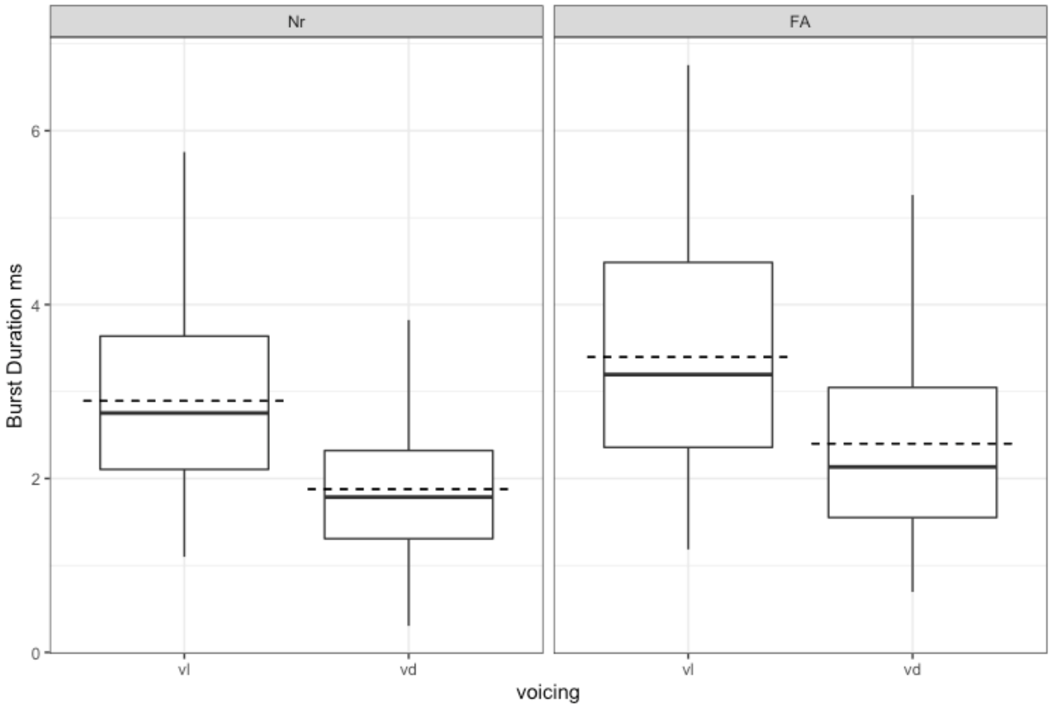


Figure 6.7. Boxplots of the fitted values of burst duration for utterance-initial stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate. The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	2.89	1.08	1.88	0.748
Fast	3.40	1.33	2.4	1.12

Table 6.7. Mean and standard deviation of the fitted values of burst duration for Voiceless and Voiced utterance-initial stops grouped by voicing and speech rate.

Figure and table 6.7 show that voicing category affected burst duration across speech rates. There was a significant increase in burst duration for Voiceless stops by an average of 1.005 ms ($p<0.0001$) in both normal and fast speech rates. Surprisingly, burst duration for in fast speech was significantly longer than in normal speech by an average of 0.52 ms ($p<0.0001$) across voicing categories.

Burst duration as a function of voicing, rate, and place of articulation

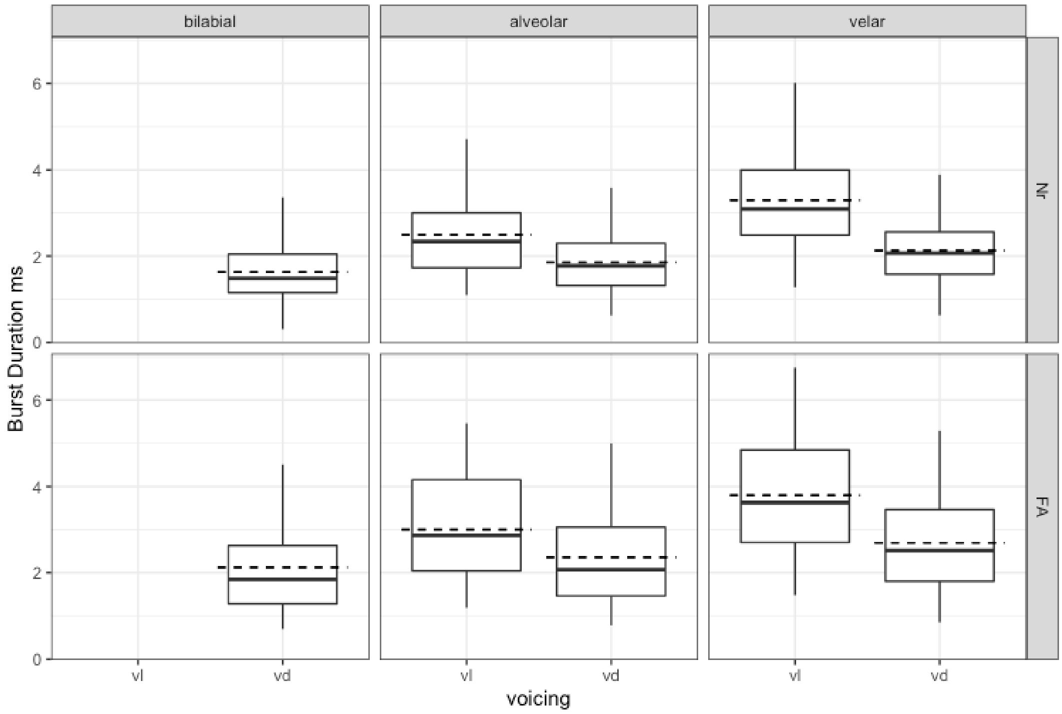


Figure 6.8. Boxplots of the fitted values of burst duration for utterance-initial stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	1.63	0.693	-	-	2.13	1.05
Alveolar	2.49	0.921	1.86	0.712	3	1.2	2.36	1.11
Velar	3.29	1.07	2.13	0.754	3.8	1.34	2.69	1.14

Table 6.8. Mean and standard deviation of the fitted values of burst duration for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate and place of articulation.

Figure and table 6.8 show that burst duration for the Voiceless was significantly longer than for the Voiced stops by an average of 0.89 ms in both normal and fast speech ($p<0.0001$). The results also show that speech rate variation affected burst duration for Voiced and Voiceless stops across places of articulation in which that the fast speech resulted in significantly longer burst duration than the normal speech by an average of 0.52 ms ($p<0.0001$). With respect to the impact of place of articulation on burst duration for Voiceless stops, the pattern observed was /velar/ > /alveolar/ by an average of 0.8 ms ($p<0.0001$) in both normal and fast speech. With respect to the impact of place of articulation on burst duration for Voiced stops, the pattern observed was /velar/ > /alveolar/ > /bilabial/ in both normal and fast speech. The differences between each pair of comparison were statistically significant (pairwise comparison: Appendix C Table C-50).

6.2.2 Utterance-medial stops CV:<u>C</u>VC (Trochaic)

Burst duration as a function of voicing and rate

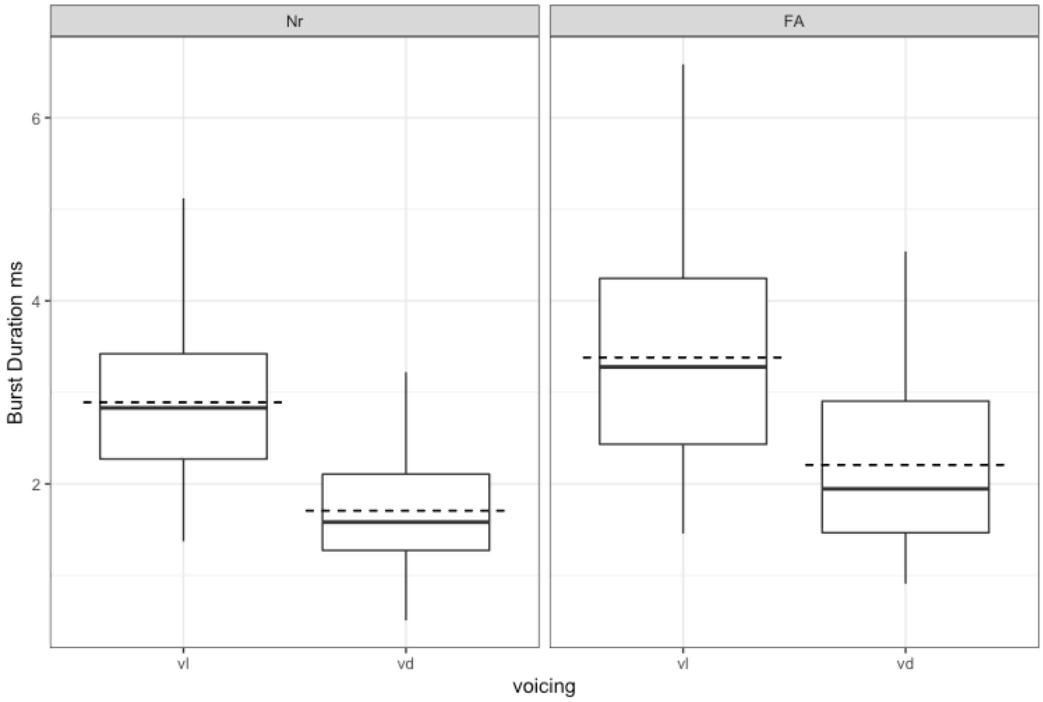


Figure 6.9. Boxplots of the fitted values of burst duration for utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced) and speech rate. The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	2.89	0.799	1.71	0.607
Fast	3.38	1.07	2.21	0.865

Table 6.9 Mean and standard deviation of the fitted values of burst duration for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

Figure and table 6.9 show that voicing category affected burst duration across speech rates. Burst duration was significantly longer in Voiceless stops by an average of 1.175 ms ($p<0.0001$) in both normal and fast speech rates. With regard to the impact of speech rate on burst duration, burst duration for in fast speech was significantly longer than in normal speech by an average of 0.5 ms ($p<0.0001$) across voicing categories.

Burst duration as a function of voicing, rate and place of articulation

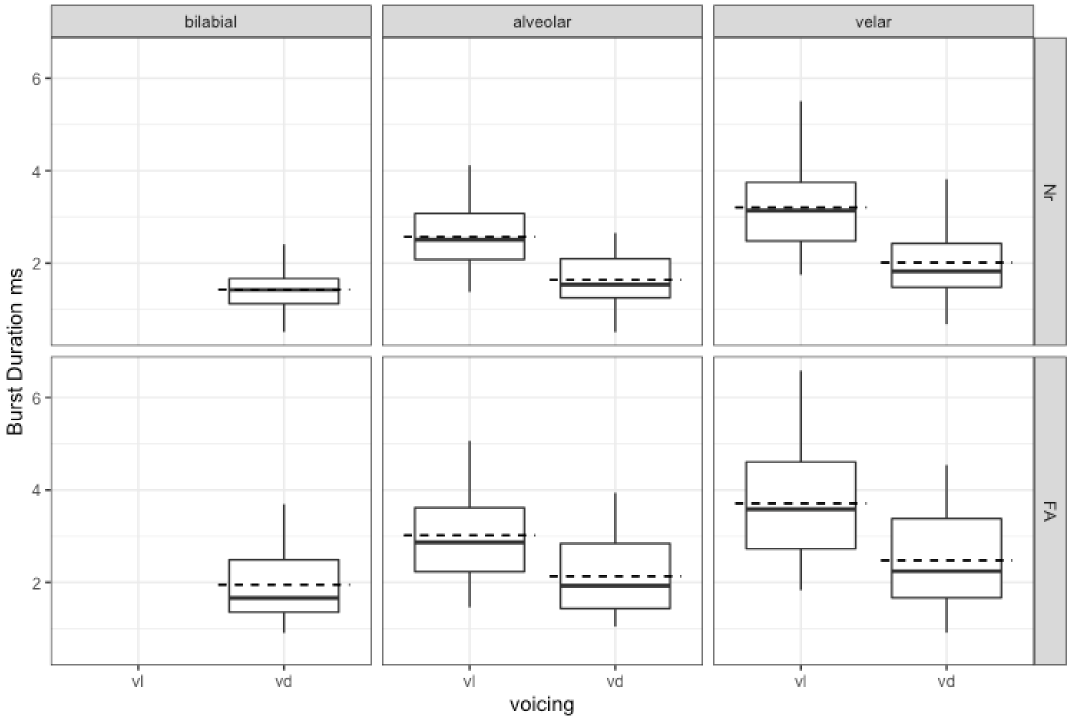


Figure 6.10. Boxplots of the fitted values of burst duration for utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	1.43	0.445	-	-	1.95	0.756
Alveolar	2.57	0.606	1.64	0.508	3.02	0.9	2.14	0.802
Velar	3.21	0.84	2.01	0.685	3.71	1.1	2.47	0.93

Table 6.10. Mean and standard deviation of the fitted values of burst duration for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate, and place of articulation.

Figure and table 6.10 show that burst duration for Voiceless stops was significantly longer than for Voiced stops by an average of 1.06 ms in both normal and fast speech ($p<0.0001$). The results also showed that fast speech resulted in significantly longer burst duration than the normal speech by an average of 0.49 ms ($p<0.0001$). With respect to the impact of place of articulation on burst duration for Voiceless stops, the pattern observed was /velar/ > /alveolar/ by an average of 0.67 ms ($p<0.0001$) in both normal and fast speech. With respect to the impact of place of articulation on burst duration for Voiced stops, the pattern observed was /velar/ > /alveolar/ > /bilabial/ in both normal and fast speech. The differences between each pair of comparison were found to be statistically significant (pairwise comparison: Appendix C Table C-53).

6.2.3 Utterance-medial stops CV<u>C</u>V:C (Iambic)

Burst duration as a function of voicing and rate

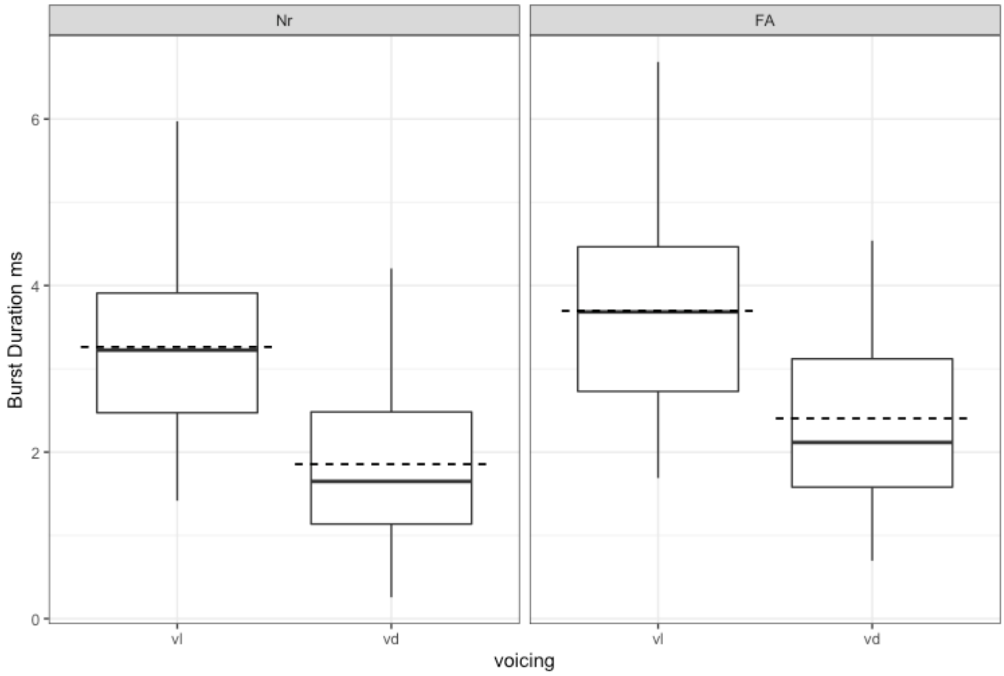


Figure 6.11. Boxplots of the fitted values of burst duration for utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate. The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	3.26	0.977	1.85	0.837
Fast	3.7	1.1	2.4	1.03

Table 6.11. Mean and standard deviation of the fitted values of burst duration for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

Figure and table 6.11 show that voicing category affected burst duration across speech rates with an overlap between the two categories. There was a significant increase in burst duration for Voiceless stops by an average of 1.36 ms ($p<0.0001$) in both normal and fast speech rates. With regard to the impact of speech rate on burst duration, burst duration for in fast speech was significantly longer than in normal speech by an average of 0.5 ms ($p<0.0001$) across voicing categories.

Burst duration as a function of voicing, rate and place of articulation

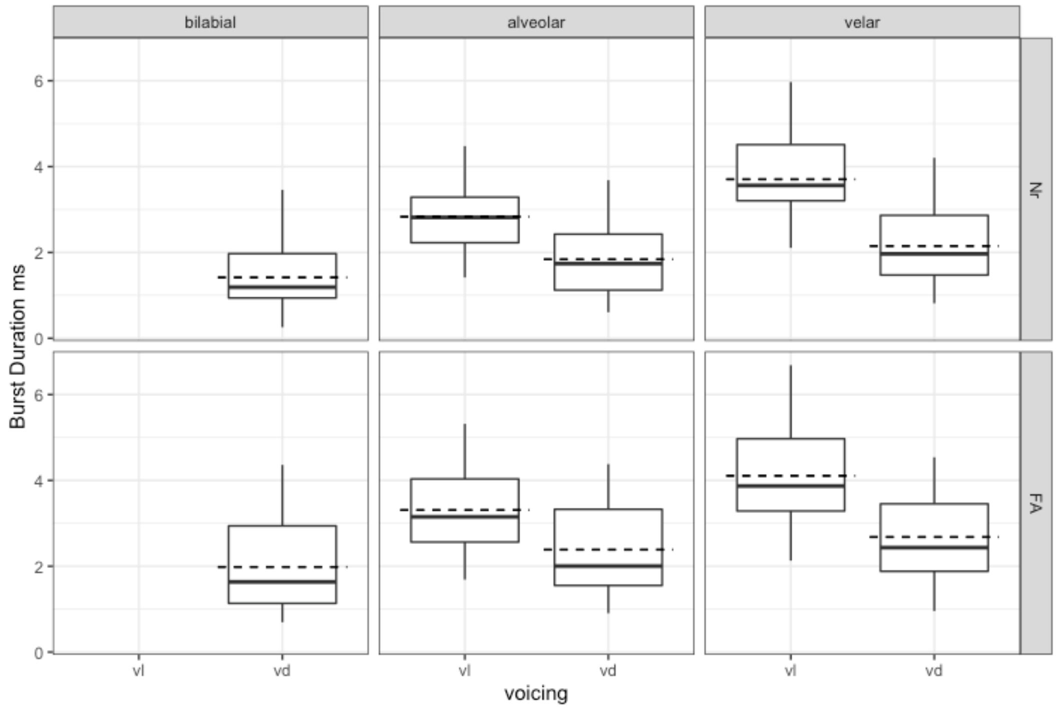


Figure 6.12. Boxplots of the the fitted values of burst duration for utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	1.42	0.693	-	-	1.98	1.01
Alveolar	2.83	0.78	1.84	0.801	3.31	0.942	2.38	0.988
Velar	3.7	0.963	2.15	0.832	4.11	1.1	2.68	0.988

Table 6.12 Mean and standard deviation of the fitted values of burst duration for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and place of articulation.

The results in figure and table 6.12 show that burst duration for the Voiceless was significantly longer than for the Voiced stops by an average of 1.26 ms across places of articulation in both normal and fast speech ($p<0.0001$). The results also showed that fast speech resulted in significantly longer burst duration than normal speech by an average of 0.5 ms ($p<0.0001$) across places of articulation and voicing categories. As for the impact of place of articulation on burst duration for Voiceless stops, the pattern observed was /velar/ > /alveolar/ by an average of 0.84 ms ($p<0.0001$) in both normal and fast speech. With respect to the impact of place of articulation on burst duration for Voiced stops, the pattern observed was /velar/ > /alveolar/ > /bilabial/ in both normal and fast speech. The differences between each pair of comparison were found to be statistically significant (pairwise comparison: Appendix C Table C-56).

6.2.4 Utterance-final stops CV:C

Burst duration as a function of voicing and rate

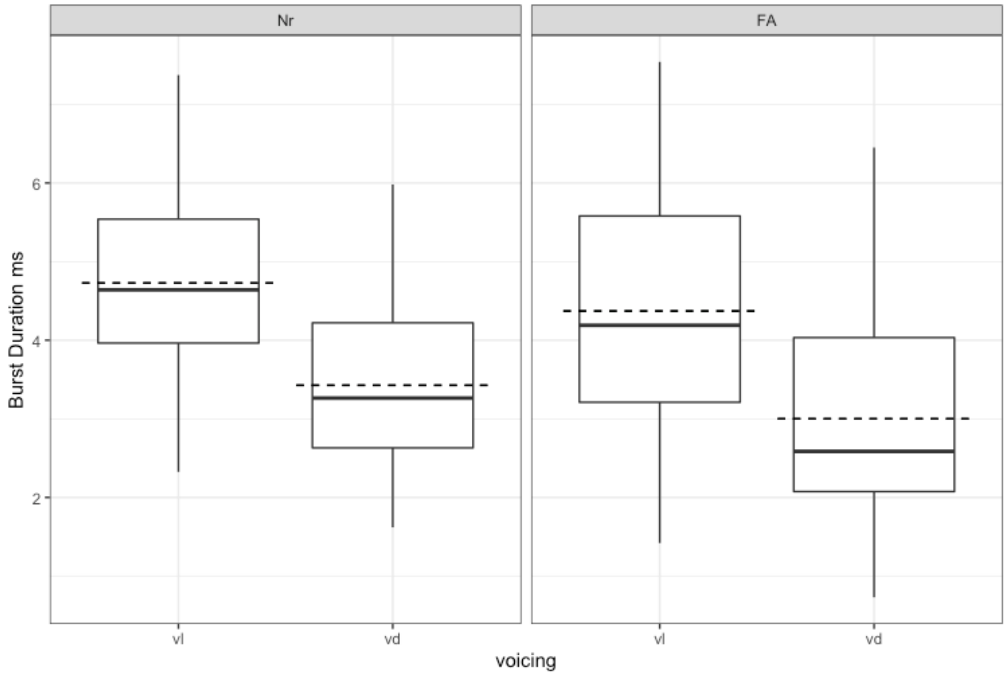


Figure 6.13 Boxplots of the fitted values of burst duration for utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate. The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	4.73	1.09	3.43	0.914
Fast	4.37	1.48	3	1.37

Table 6.13 Mean and standard deviation of the fitted values of burst duration for Voiceless and Voiced utterance-final stops grouped by voicing and speech rate.

Figure and table 6.13 show that voicing category affected burst duration across speech rates. There was a significant increase in burst duration for Voiceless stops by an average of 1.34 ms ($p<0.0001$) in both normal and fast speech rates. With regard to the impact of speech rate on burst duration, the analysis revealed opposite results to what has been found in utterance-initial and utterance-medial stops. That is, burst duration in normal speech was significantly longer than in fast speech by an average of 0.4 ms ($p<0.0001$) across voicing categories.

Burst duration as a function of voicing, rate and place of articulation

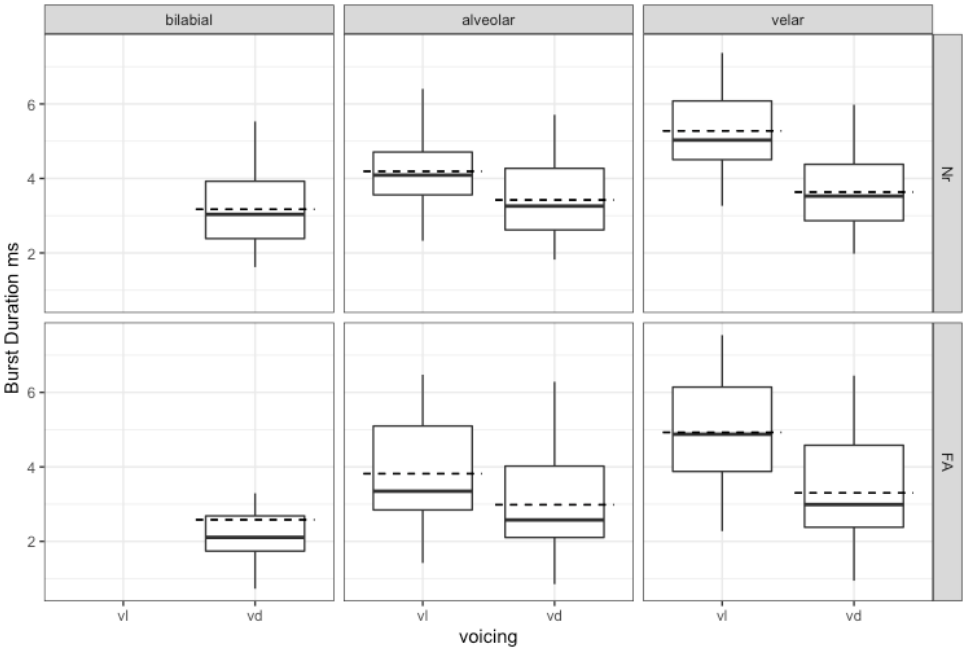


Figure 6.14 Boxplots of the fitted values of the burst duration for utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	3.17	0.9	-	-	2.58	1.37
Alveolar	4.19	0.939	3.42	0.896	3.82	1.37	2.98	1.33
Velar	5.28	0.956	3.63	0.893	4.93	1.37	3.3	1.35

Table 6.14 Mean and standard deviation of burst duration for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and place of articulation.

The results in figure and table 6.14 show that burst duration for the Voiceless was significantly longer than for the Voiced stops by an average of 1.22 ms across places of articulation in both normal and fast speech ($p<0.0001$). The results also showed that normal speech resulted in significantly longer burst duration than fast speech by an average of 0.42 ms ($p<0.0001$) across places of articulation and voicing categories. With respect to the impact of place of articulation on burst duration for Voiceless stops, the pattern observed was /velar/ > /alveolar/ by an average of 1.1 ms ($p<0.0001$) in both normal and fast speech. With respect to the impact of place of articulation on burst duration for Voiced stops, the pattern observed was /velar/ > /alveolar/ > /bilabial/ in both normal and fast speech. The differences between each pair of comparison were found to be statistically significant (pairwise comparison: Appendix C Table C-59).

6.2.5 Summary of results for burst duration

Burst duration for Voiceless stops is longer than for voiced stops across positions, places of articulation, vowel type, gender, and speech rates with overlap between the two categories. Longer burst duration is noted in utterance-final stops than in the rest of the positions. This is supported by the articulatory high pressure during the constriction in utterance-final stops that lead to longer release burst (Van Alphen and Smits, 2004). In addition, more variability occurs in fast speech than in normal speech across positions and voicing categories.

Interestingly, unlike the previous durational correlates, burst duration was longer in fast speech than in normal speech in utterance-initial and utterance-medial stops. This difference held across all the factors. In utterance-final stops, however, the difference appeared opposite to the rest of the positions in that burst duration was longer in normal speech than in fast speech. One explanation with regard to longer burst duration in fast speech may be that it is affected by the duration of aspiration. That is, short aspiration results in longer burst duration and vice versa.

6.3 Preceding vowel duration *PVD* (ms)

The purpose of this section is to present the results for the duration of the vowels preceding Voiceless and Voiced stops in Najdi Arabic with respect to various factors including voicing, position, rate, place of articulation, vowel type, and gender. As pointed out in the literature, it has been proposed that vowels preceding Voiced stops are longer than vowels preceding Voiceless stops in aspirating and voicing languages (Chen, 1970; Alghamdi, 1990; Kluender et al., 1982; Luce and Charles Luce, 1985; Al-Tamimi and Khattab, 2018). Some studies showed that a voicing effect on preceding vowel duration is not found in some Arabic dialects (Mitleb, 1984; De Jong and Zawaydeh, 2002; Al-Gamdi, 2013).

6.3.1 Utterance-medial stops (Trochaic) CV:CVC

PVD as a function of voicing and rate.

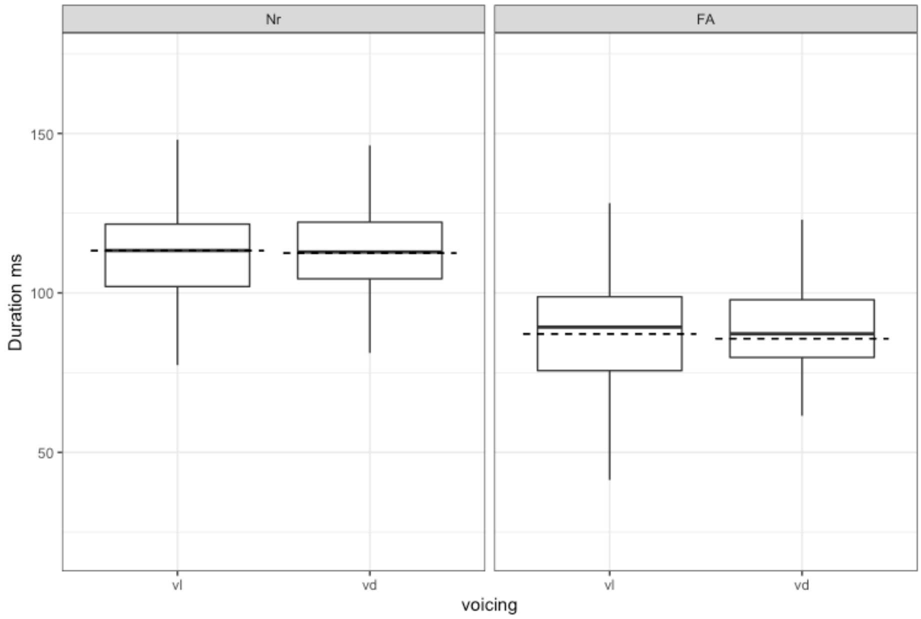


Figure 6.15 Boxplots of the fitted values of PVD for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	113	17.7	112	19.8
Fast	87.1	16	85.7	18.9

Table 6.15 Mean and standard deviation of the fitted values of PVD for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

Figure and table 6.15 show the values of PVD for Voiceless and Voiced utterance-medial stops (Iambic) as a function of voicing and rate. The results show that PVD for the Voiceless showed a marginal increase by an average of 1.2 ms in both normal (p=0.47) and fast (p=0.17) rates. With regard to speech rate effect, the results show a significant increase of PVD in normal speech by an average of 26.02 ms across voicing categories (p<0.0001).

PVD as a function of voicing, rate, and place of articulation

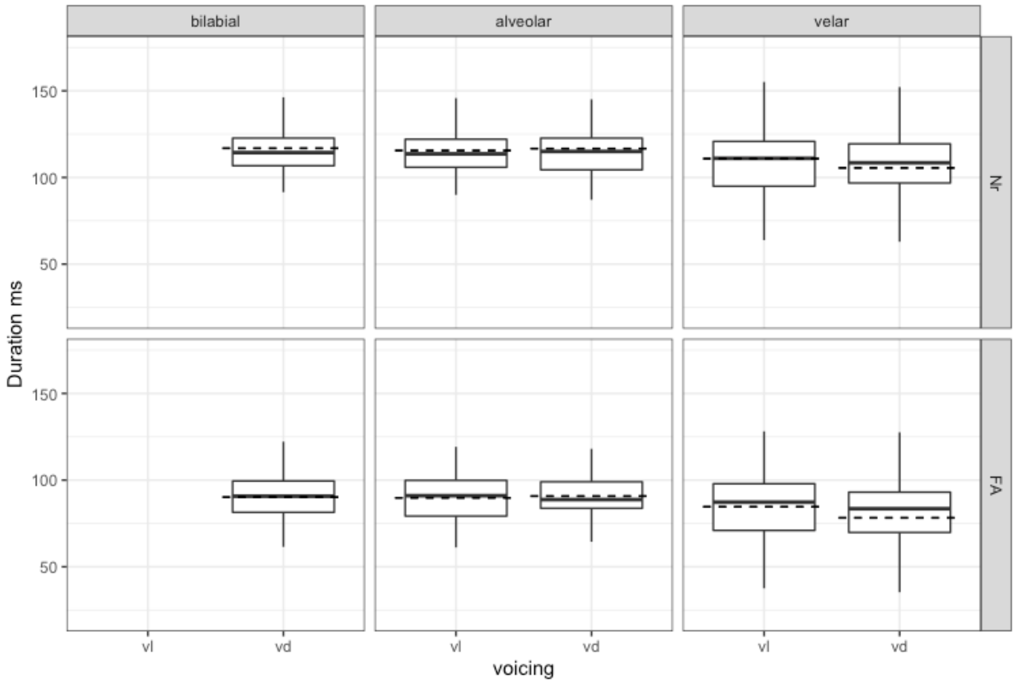


Figure 6.16. Boxplots of the fitted values of PVD for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	117	13.8	-	-	90.2	12.6
Alveolar	116	14.9	117	15.3	89.8	13	90.8	12.5
Velar	111	19.9	106	24.5	84.6	18.1	78.2	24

Table 6.16. Mean and standard deviation of the fitted values of PVD for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and place of articulation.

Figure and table 6.16 show PVD for Voiced and Voiceless utterance-medial stops (Iambic) as a function of voicing, rate, and place of articulation. With regard to voicing effect on PVD, PVD for alveolar stops was marginally longer in the Voiced context than in the Voiceless by an average of 1 ms across speech rates ($p=0.5$). PVD for velar stops was significantly longer in the Voiceless context than in the Voiced by an average of 5.7 across speech rates ($p<0.0001$). With regard to speech rate effect, PVD was significantly longer in the normal speech than in the fast by an average of 26.7 ms across voicing categories and places of articulation ($p<0.0001$). With respect to the impact of place of articulation on PVD, the pattern observed was / alveolar / >/velar/ by an average of 5.1 ms in the Voiceless context across speech rates ($p<0.0001$). With respect to the impact of place of articulation on PVD for

Voiced stops, the pattern observed was /bilabial/ = /alveolar/ > /velar/ across speech rates (pairwise comparison: Appendix C Table C-72).

6.3.2 Utterance-final stops CV:C

PVD as a function of voicing and rate

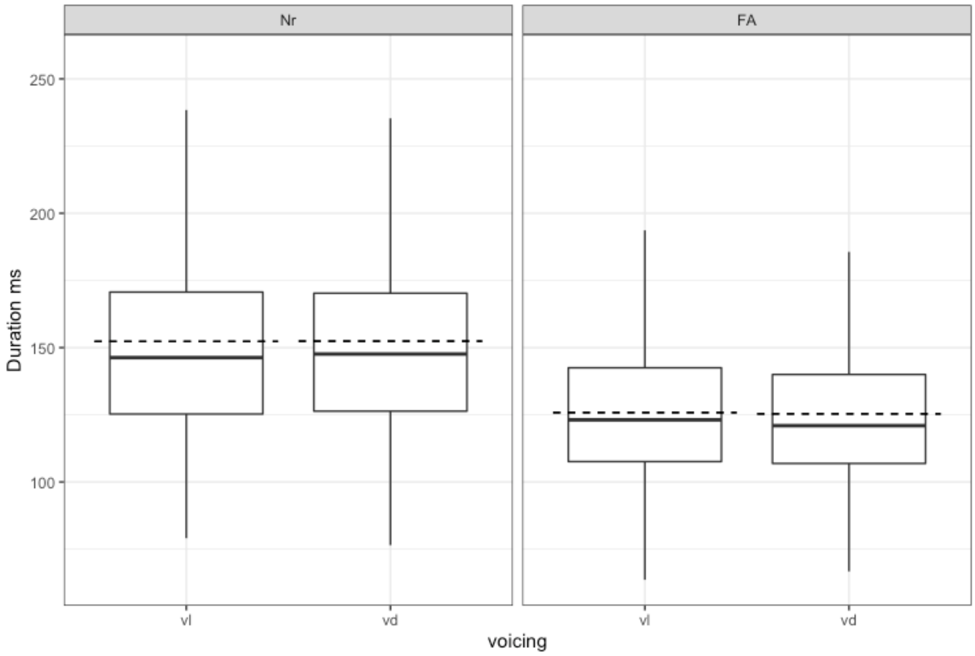


Figure 6.17. Boxplots of the fitted values of PVD for Voiceless and utterance-final stops grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	152	36.4	152	37.3
Fast	126	25.9	125	26.5

Table 6.17. Mean and standard deviation of the fitted values of PVD for Voiceless and Voiced utterance-final stops grouped by voicing and speech rate.

Figure and table 6.17 present PVD values for Voiceless and Voiced utterance-final stops as a function of voicing and speech rate. The results show no difference between PVD for Voiceless and Voiced stops in the normal speech whereas there was a marginal difference in the fast speech. With regard to speech rate effect, the results show a significant increase of PVD in normal speech by an average of 26.5 ms across voicing categories ($p<0.0001$).

PVD as a function of voicing, rate, and place of articulation

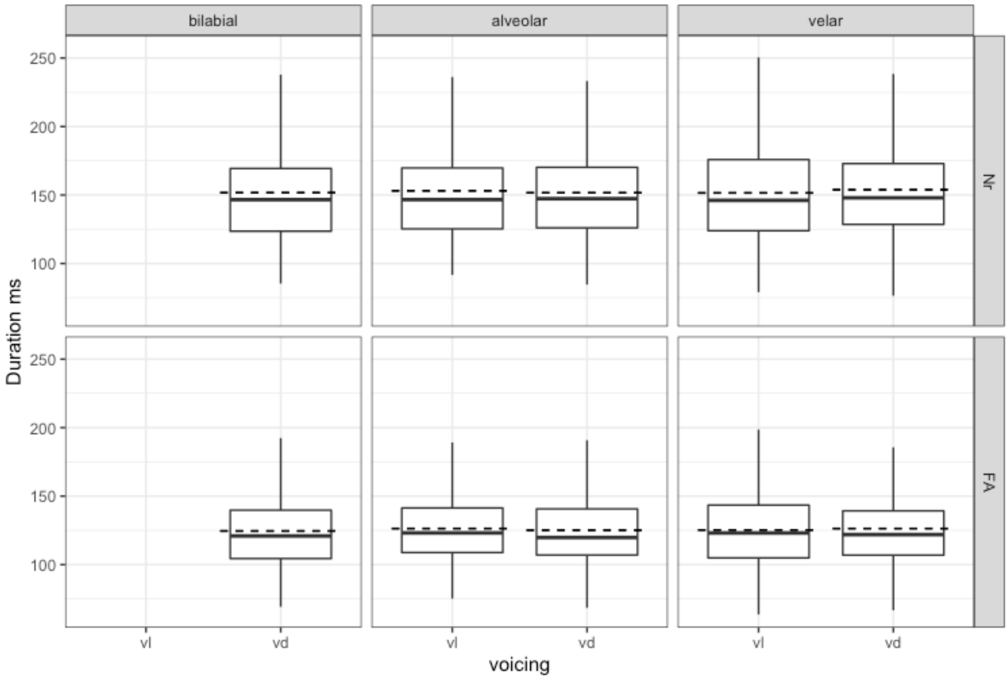


Figure 6.18. Boxplots of the fitted values of PVD for Voiceless and utterance-final stops grouped by voicing, speech rate, and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	152	37.7	-	-	125	26.5
Alveolar	153	35.3	152	36.9	126	23.6	125	25.8
Velar	152	37.5	154	38.4	125	28	126	27.4

Table 6.18. Mean and standard deviation of the fitted values of PVD for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and place of articulation.

Figure and table 6.18 show PVD for Voiced and Voiceless utterance-final stops as a function of voicing, rate, and place of articulation. With regard to voicing effect on PVD, PVD for alveolar stops was marginally longer in the Voiceless context than in the Voiced by an average of 1 ms in normal ($p = 0.7$) and fast ($p = 0.75$) rates. PVD for velar stops was marginally longer in the Voiced context than in the Voiceless by an average of 1.5 ms in normal ($p = 0.47$) and fast ($p = 0.75$) speech rates. With regard to speech rate effect, PVD was significantly longer in the normal speech than in the fast by an average of 27.2 ms across places of articulation and voicing categories (pairwise comparison: Appendix C Table C-77). With respect to the impact of place of articulation on PVD for Voiceless stops, the pattern observed was /alveolar/ >/velar/ by an average of 1 ms across speech rates (pairwise comparison: Appendix C Table C-77). With respect to the impact of place of articulation on PVD for Voiced stops, the pattern observed was / velar / > /alveolar/ = /bilabial/ across speech rates. All the differences were found to be not significant across speech rate (pairwise comparison: Appendix C Table C-77).

6.3.3 Summary of results for PVD

It has been established that **PVD** for Voiceless and Voiced stops in Najdi Arabic is not a robust acoustic correlate that signals the distinction between the two voicing categories. The results show a variation in PVD values with some cases that contradict the expected pattern; vowels preceding Voiced stops are longer than vowels preceding Voiceless stops. Similar results have been found in some dialects of Arabic which might suggest that voicing has no impact on precding vowel duration in these dialects (Jordanian Arabic: Mitleb, 1984 ; Jordanian Arabic: De Jong and Zawaydeh, 2002; Saudi Arabic: Al-Gamdi, 2013). As posited by De Jong and Zawaydeh (2002), vowel quantity in the vowel system of Arabic is primary and essential for the vocalic contrast and consequently could be more important than the voicing which is suppressed and reduced.

6.4 Following vowel duration *FVD* (ms)

The aim of this section is to present the results for the duration of the vowels following Voiceless and Voiced stops in Najdi Arabic with respect to various factors including voicing, position, rate, place of articulation, vowel type, and gender. As pointed out in the literature, it has been proposed that vowels following Voiced stops are longer than vowels following Voiceless stops in aspirating and voicing languages (Alghamdi, 1990; Allen and Miller, 1999; Jessen, 2001; Al-Tamimi and Khattab, 2018).

6.4.1 Utterance-initial stops CV:C

FVD as a function of voicing and rate

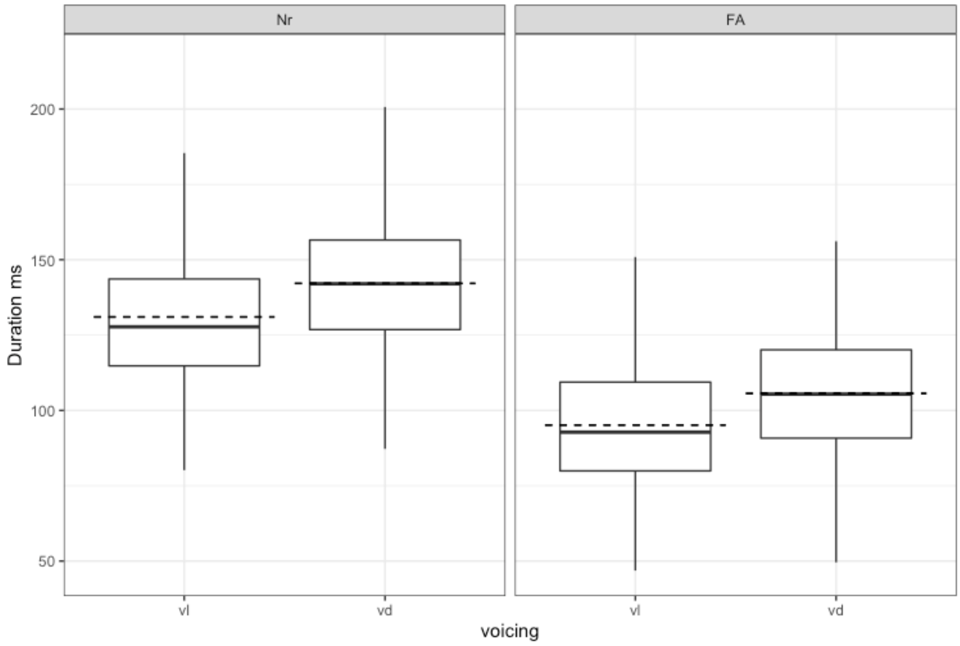


Figure 6.19. Boxplots of the fitted values of FVD for Voiceless and Voiced utterance-initial stops grouped by voicing and speech.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	131	23.5	142	23.5
Fast	95.1	21.7	106	22.2

Table 6.19. Mean and standard deviation of the fitted values of FVD for Voiceless and Voiced utterance-initial stops grouped by voicing and speech.

Figure and table 6.19 show the values of FVD for Voiceless and Voiced stops as a function of voicing and rate. FVD for the Voiced was significantly longer than for the Voiceless stops by an average of 10.95 ms across speech rates ($p<0.0001$). For speech rate effect, the results show significant increase in normal speech by an average of 35.95 ms across voicing categories ($p<0.0001$).

FVD as a function of voicing, rate, and place of articulation

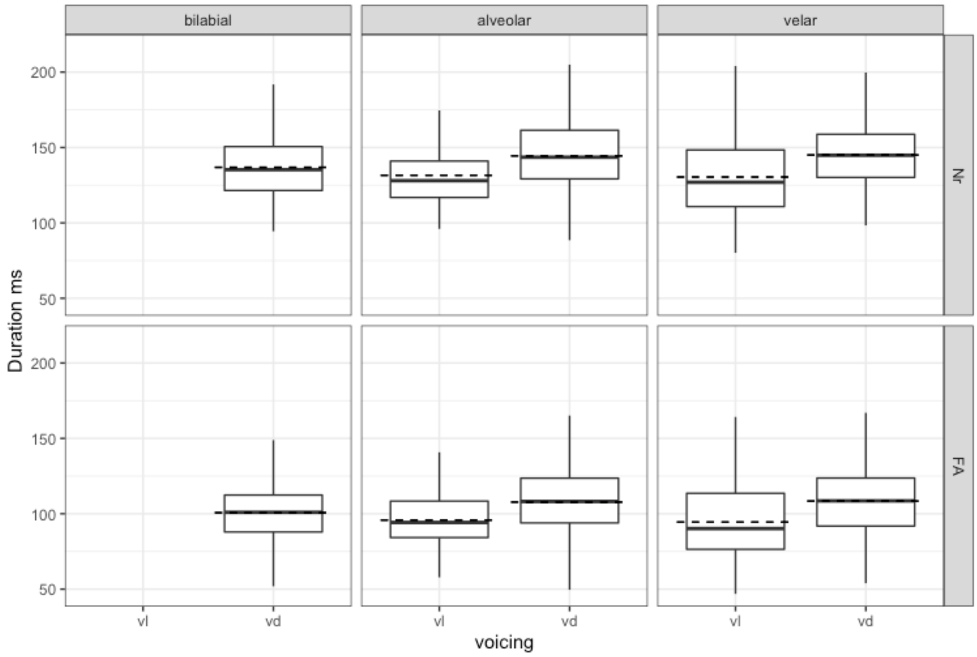


Figure 6.20. Boxplots of the fitted values of FVD for Voiceless and Voiced utterance-initial stops grouped voicing, speech rate, and place of articulation.

	Voiceless normal		Voiced normal		Voiceless fast		Voiced fast	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	137	21.2	-	-	101	18.9
Alveolar	132	20.3	144	25.1	95.8	17.5	108	23.7
Velar	130	26.4	145	23	94.5	25.3	108	22.9

Table 6.20. Mean and standard deviation of the fitted values of FVD for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate, and place of articulation.

Figure and table 6.20 summarise FVD values for Voiced and Voiceless utterance-initial stops as a function of voicing category, rate, and place of articulation. With regard to the voicing effect on FVD, FVD for Voiced stops was significantly longer than for Voiceless stops by an average of 13.2 ms across speech rates and places of articulation ($p<0.0001$). With regard to speech rate effect, FVD in normal speech was significantly longer than in fast speech by an average of 36.1 ms across voicing categories and places of articulation (pairwise comparison: Appendix C Table C-62). With respect to the impact of place of articulation on FVD, the pattern observed was /alveolar/ >/velar/ by an average of 1.7 ms in the Voiceless context in both normal ($P<0.45$) and fast ($p<0.38$) rates. With respect to the impact of place of articulation on FVD for Voiced stops, the pattern observed was /velar/ > /alveolar/ > /bilabial/. The results showed that FVD in the context of bilabial stops was significantly shorter than in the contexts of alveolar and velar stops across speech rates (pairwise comparison: Appendix C Table C-62).

6.4.2 Utterance-medial stops CVCV:C (Iambic)

FVD as a function of voicing and rate

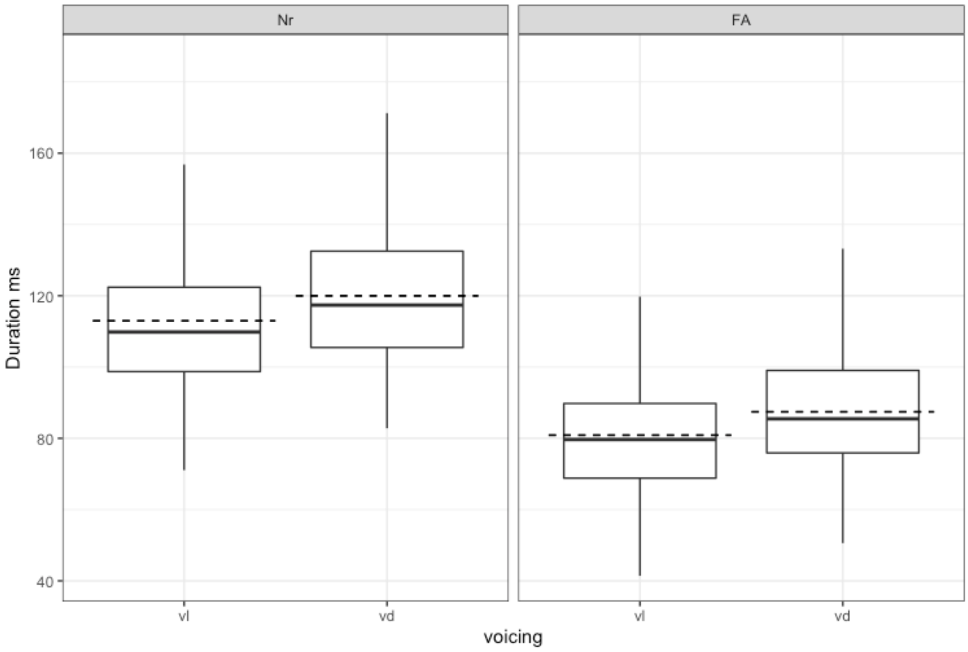


Figure 6.21. Boxplots of the fitted values of FVD for Voiceless Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	113	20.6	120	20.1
Fast	80.9	15.8	87.4	16.2

Table 6.21. Mean and standard deviation of the fitted values of FVD for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

Figure and table 6.21 show the values of FVD for Voiceless and Voiced utterance-medial stops (Iambic) as a function of voicing and rate. FVD for the Voiced was significantly longer than for the Voiceless stops by an average of 6.8 ms across speech rates ($p<0.0001$). For speech rate effect, the results show a significant increase in normal speech by an average of 32.4 ms across voicing categories ($p<0.0001$).

FVD as a function of voicing, rate, and place of articulation

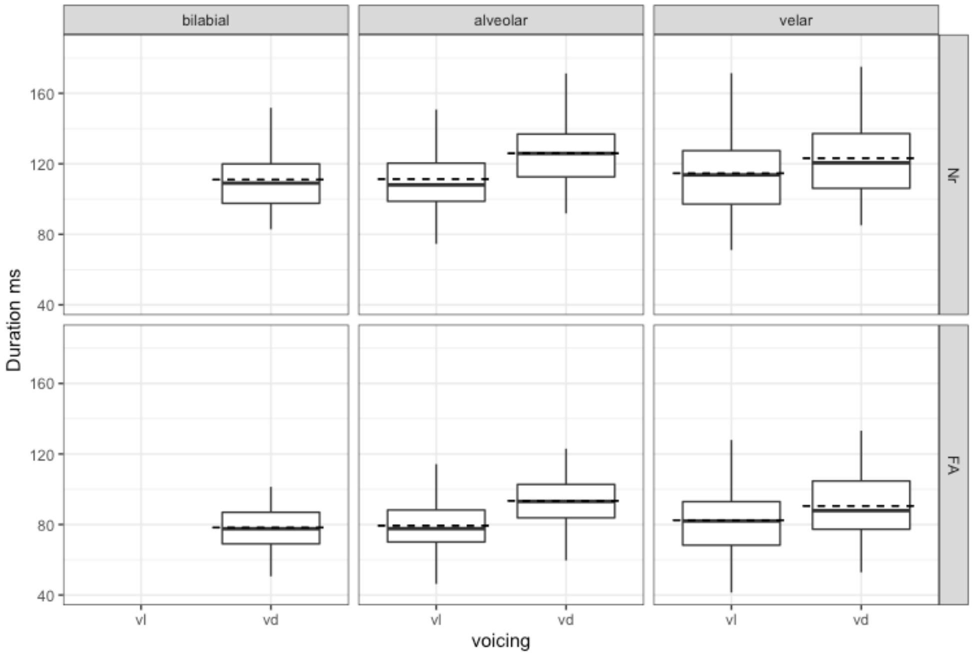


Figure 6.22. Boxplots of the fitted values of FVD for Voiceless Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	111	17.1	-	-	78.3	12.7
Alveolar	111	19.2	126	18.3	79.3	14.1	93.5	14.3
Velar	115	21.8	123	21.3	82.5	17.3	90.5	17.2

Table 6.22. Mean and standard deviation of the fitted values of FVD for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and place of articulation.

Figure and table 6.22 show FVD values for Voiced and Voiceless utterance-medial stops (Iambic) as a function of voicing, rate, and place of articulation. With regard to voicing effect on FVD, FVD was significantly longer in the Voiced context than in the Voiceless by an average of 11.3 ms across speech rates and places of articulation ($p<0.0001$). With regard to speech rate effect, FVD in normal speech was significantly longer than in fast speech by an average of 32.4 ms across voicing categories and places of articulation ($p<0.0001$). With respect to the impact of place of articulation on FVD, the pattern observed was /velar/ >/alveolar/ by an average of 3.6 ms in the Voiceless context in both normal ($p=0.032$) and fast ($p=0.047$). With respect to the impact of place of articulation on FVD for Voiced stops, the pattern observed was / alveolar / > / velar / > /bilabial/ across speech rates (pairwise comparison: Appendix C Table C-67).

6.4.3 Summary of results for FVD

The results showed that **FVD** was a robust acoustic correlate that signals the distinction between Voiceless and Voiced stops in Najdi Arabic across positions, speech rates, places of articulation, vowel type, and genders. It has been established that FVD for Voiced stops is significantly longer than for Voiceless stops across the examined contexts. It could be noticed that FVD is more effective in utterance-initial stops when considering the size of the difference between the two voicing categories.

The results suggest a robust presence of aspiration in which the long aspiration duration leads to shorter following vowel as reported for aspirating languages (Jessen, 2001; Allen and Miller, 1999). The results also match the pattern found in Ghamidi Arabic (Alghamdi, 1990) and Lebanese Arabic (Al-Tamimi and Khattab, 2018).

6.5 Conclusion

The results of the acoustic analysis of the durational correlates presented in this chapter emphasise the importance of the phonetic details in characterising the phonological aspects of voicing contrast in Najdi Arabic. The following table displays the summary of the results of the durational acoustic correlates under normal/fast speech rates.

		Utterance-initial			**Utterance-medial (trochaic)**			**Utterance-medial (iambic)**			**Utterance-final**		
Correlate	**Voicing**	**rate**	**M**	**SD**	**rate**	**M**	**SD**	**rate**	**M**	**SD**	**rate**	**M**	**SD**
Closure duration	Voiceless	Nr	…..	…..	Nr	56.2	6.94	Nr	60.3	7.39	Nr	86.8	13.1
	Voiced	Nr	…..	…..	Nr	51.8	5.81	Nr	54.1	6.55	Nr	96.1	13.2
Burst duration	Voiceless	Nr	2.89	1.08	Nr	2.89	0.79	Nr	3.26	0.97	Nr	4.73	1.09
	Voiced	Nr	1.88	0.74	Nr	1.71	0.6	Nr	1.85	0.83	Nr	3.43	0.91
FVD	Voiceless	Nr	131	23.5	Nr	23.5	23.5	Nr	113	20.6	Nr	…..	…..
	Voiced	Nr	142	23.5	Nr	23.5	23.5	Nr	120	20.1	Nr	…..	…..
PVD	Voiceless	Nr	…..	…..	Nr	113	17.7	Nr	…..	…..	Nr	152	36.4
	Voiced	Nr	…..	…..	Nr	112	19.8	Nr	…..	…..	Nr	152	37.3
Closure duration	Voiceless	FA	…..	…..	FA	47.6	7.74	FA	49.9	7.4	FA	61.8	11.8
	Voiced	FA	…..	…..	FA	42.9	6.49	FA	43.8	6.8	FA	70.9	11.2
Burst duration	Voiceless	FA	3.4	1.33	FA	3.38	1.07	FA	3.7	1.1	FA	4.37	1.48
	Voiced	FA	2.4	1.12	FA	2.21	0.86	FA	2.4	1.03	FA	3	1.37
FVD	Voiceless	FA	95.1	21.7	FA	…..	…..	FA	80.9	15.8	FA	…..	…..
	Voiced	FA	106	22.2	FA	…..	…..	FA	87.4	16.2	FA	…..	…..
PVD	Voiceless	FA	…..	…..	FA	87.1	16	FA	…..	…..	FA	126	25.9
	Voiced	FA	…..	…..	FA	85.7	18.9	FA	…..	…..	FA	125	26.5

Table 6.23. Overview of the acoustic measures (Mean and standard deviation) of the durational correlates in Voiced and Voiceless stops under normal/fast speech rates.

It has been demonstrated that the durational correlates in Voiceless and Voiced stops in Najdi Arabic are generally consistent with what has been proposed in the literature in terms of the distinction between voicing and aspirating languages. This includes longer closure in Voiceless stops in utterance-medial and longer following vowel in Voiced stops contexts. The results also show that these patterns occur across places of articulation and speech rates.

The pattern that appeared to be different from the expectation is that closure duration in utterance-final stops shows the opposite results in which closure duration is longer for Voiced than for Voiceless stops. The results also show that voicing has no significant effect on the preceding vowel duration. this pattern has been found in some Arabic dialects as well.

The overall picture of the durational correlates results demonstrates that the voicing contrast in Najdi Arabic has features from both aspirating and voicing languages. The presence of long lag aspiration in utterance-initial and utterance-medial Voiceless stops has a clear impact on the duration of the following vowel as reported in aspirating languages. Additionally, the presence of heavy aspiration (HASP) in utterance-final Voiceless stops might have affected the duration of their closure (Jessen 1998). That is, the heavy aspiration resulted in the shortening of closure duration in Voiceless stops which provided explanation for the similarity between Voiceless and Voiced stops in terms of closure duration (Jessen 1998). In terms of the features from voicing languages, Najdi Arabic shows similar patterns to some Arabic dialects including Lebanese Arabic (Al-Tamimi and Khattab, 2018) in terms of burst duration and Jordanian Arabic (Mitleb, 1984) in terms of preceding vowel duration.

Chapter 7. Results for the spectral correlates of voicing contrast

As mentioned earlier, the main goal of investigating the acoustic correlates that signal the distinction between Voiceless and Voiced stops in Najdi Arabic is to have deep understanding of the phonetic and phonological aspects of voicing contrast in this variety, which shows uncommon patterns by contrasting aspirated and prevoiced stops. Spectral correlates have been investigated in numerous studies cross-linguistically to characterise their manifestation in Voiceless and Voiced stops in both aspirating and voicing languages. This chapter reports on the results of the acoustic analysis for the spectral correlates that are expected to cue the opposition between Voiceless and Voiced stops in Najdi Arabic. The correlates presented in this chapter include burst intensity, F0 onset, F0 offset, F1 onset, F1 offset, H1-H2 onset, and H1-H2 offset.

The results of the LMM analysis are presented for each of the acoustic correlates considering multiple factors including speech rate, place of articulation, vowel type, and gender. The chapter is divided into eight parts: 1) burst intensity, 2) F0 onset, 3) F0 offset, 4) F1 onset, 5) F1 offset, 6) H1-H2 onset, 7) H1-H2 offset, and 8) summary of the results under normal and fast speech rate conditions.

7.1 Burst intensity (dB)

It was found in previous literature that burst intensity is higher for Voiceless than for Voiced stops in aspirating and voicing languages (Halle et al., 1957; Zue, 1976; Lousada et al., 2010; Lavoie, 2001). The articulatory explanation for this distinction is that Voiceless stops are produced with longer constriction which leads to high airflow in the burst while Voiced stops produced with low frequency voice bar in the constriction which leads to low intensity in the burst (van Alphen and Smits, 2004).

7.1.1 Utterance-initial stops C̲V:C

Burst intensity as a function of voicing and rate

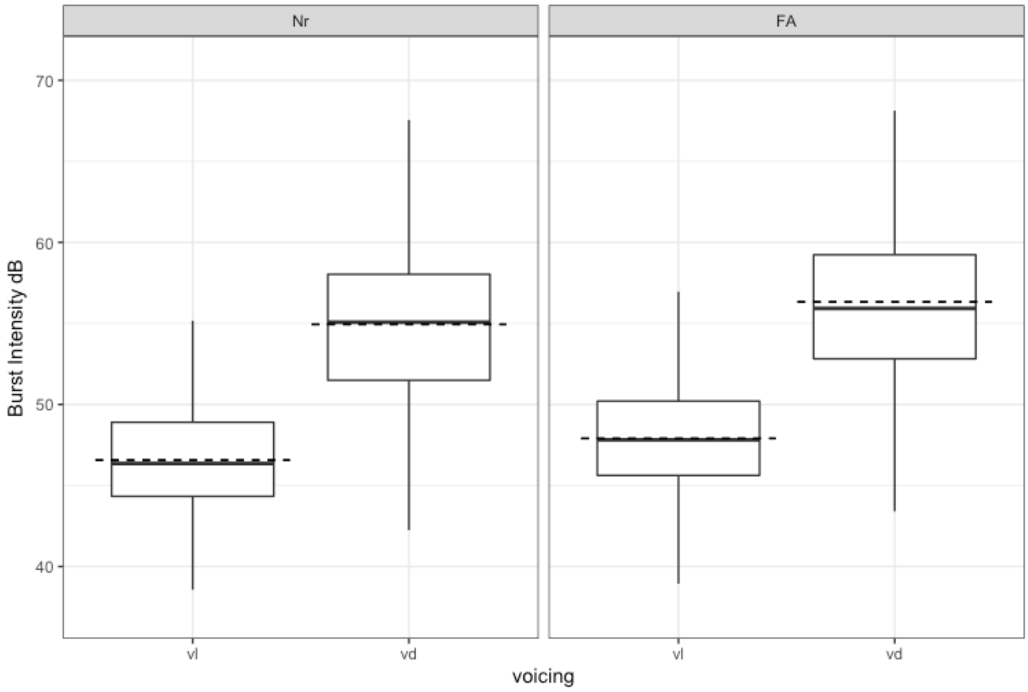


Figure 7.1. Boxplots of the fitted values of burst intensity for utterance-initial stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate. The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	46.6	3.53	54.9	5
Fast	47.9	3.68	56.3	4.85

Table 7.1. Mean and standard deviation of the fitted values of burst intensity for Voiceless and Voiced utterance-initial stops grouped by voicing and speech rate.

Figure and table 7.1 show that voicing category affected burst intensity across speech rates. The values of burst intensity for Voiceless and Voiced stops showed that there was a significant difference in burst intensity for the Voiced stops by an average of 8.35 dB (p<0.0001) in both normal and fast speech rates. With regard to the impact of speech rate on burst intensity, burst intensity in fast speech was significantly higher than in normal speech by an average of 1.35 dB (p<0.0001) across voicing categories.

Burst intensity as a function of voicing, rate, and place of articulation

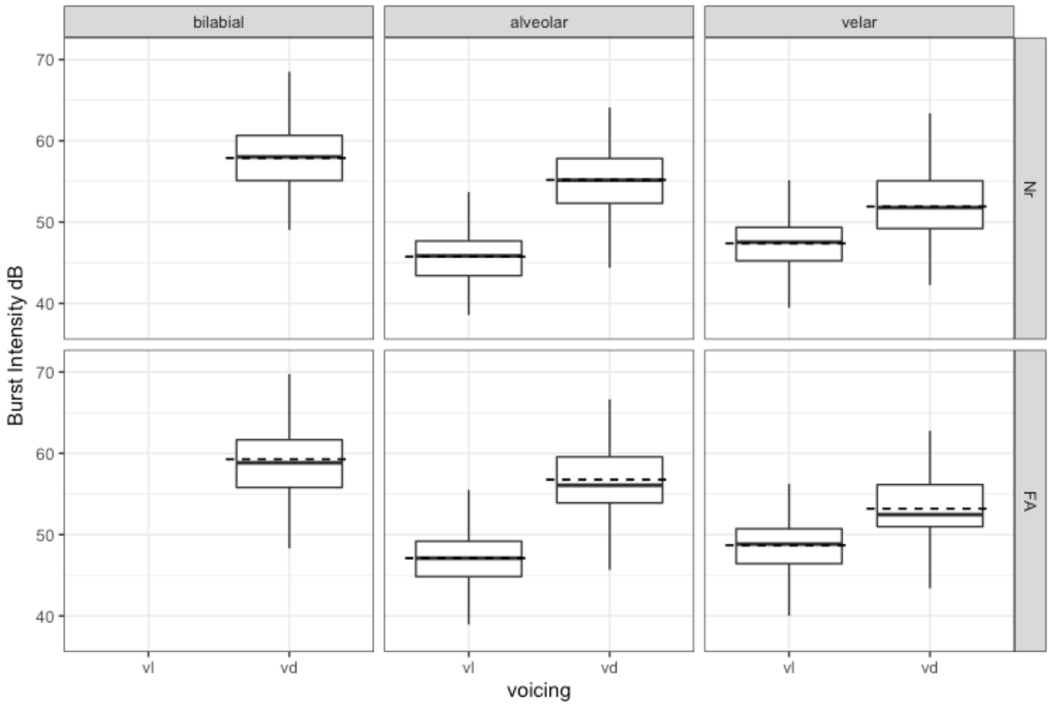


Figure 7.2. Boxplots of the fitted values of burst intensity for utterance-initial stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	57.9	4.63	-	-	59.3	4.56
Alveolar	45.7	3.35	55.2	4.38	47.1	3.52	56.8	4.15
Velar	47.4	3.52	51.9	4.12	48.7	3.67	53.2	3.82

Table 7.2. Mean and standard deviation of the fitted values of burst intensity for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate and place of articulation.

The results in figure and table 7.2 display that burst intensity for the Voiced was significantly higher than for the Voiceless stops by an average of 9.6 dB for alveolar stops in both normal and fast speech (p<0.0001). The results also showed that burst intensity for the Voiced was significantly higher than for the Voiceless stops by an average of 4.5 dB for velar stops in both normal and fast speech (p<0.0001). As for speech rate effect, fast speech resulted in significantly higher burst intensity than normal speech by an average of 1.4 dB (p<0.0001) across places of articulation and voicing categories. With respect to the impact of place of articulation on burst intensity for Voiceless stops, the pattern observed was /velar/ > /alveolar/ by an average of 1.65 dB (p<0.0001) in both normal and fast speech. With respect to the impact of place of articulation on burst intensity for Voiced stops, the pattern observed was /bilabial/ > /alveolar/ > /velar/ in both normal and fast speech. The differences between each pair of comparison were found to be statistically significant (pairwise comparison: Appendix C Table C-80).

7.1.2 Utterance-medial stops CV:CVC (Trochaic)

Burst intensity as a function of voicing and rate

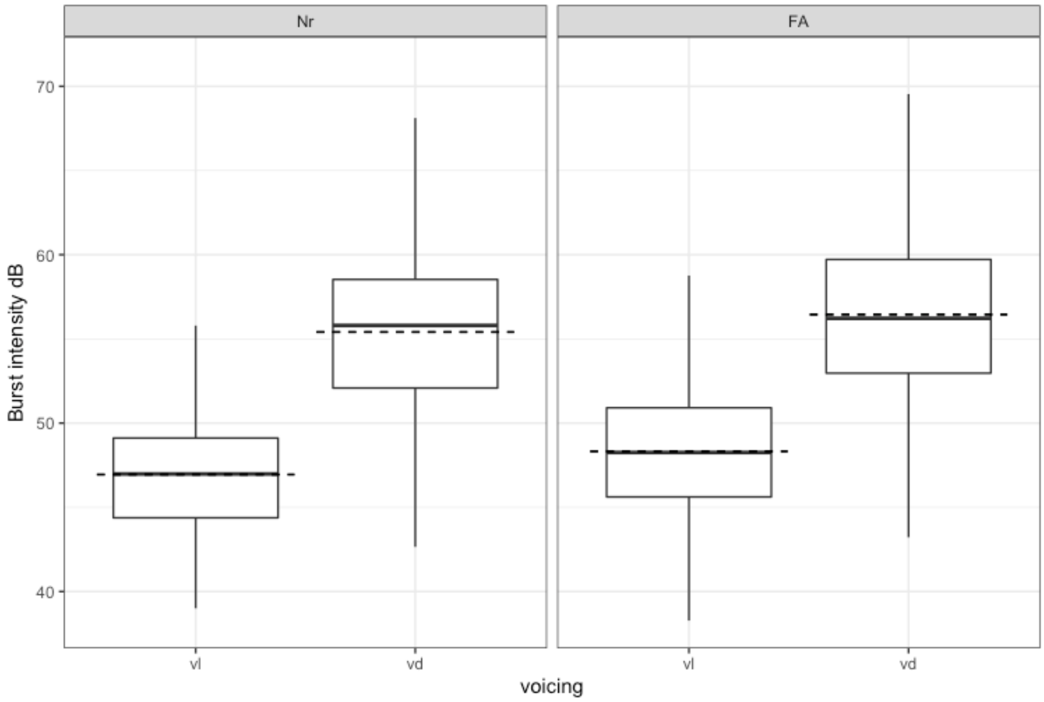


Figure 7.3. Boxplots of the fitted values of burst intensity for utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate. The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	46.9	0.799	55.4	0.607
Fast	48.3	1.07	56.4	0.865

Table 7.3. Mean and standard deviation of the fitted values of burst intensity for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

Figure and table 7.3 show that voicing category affected burst intensity for utterance-medial stops (Trochaic) across speech rates. The values of burst intensity for Voiceless and Voiced stops showed that there was a significant increase in burst intensity for the Voiced stops by an average of 8.3 dB (p<0.0001) in both normal and fast speech rates. Burst intensity in fast speech was significantly higher than in normal speech by an average of 1.2 dB (p<0.0001) across voicing categories.

Burst intensity as a function of voicing, rate, and place of articulation

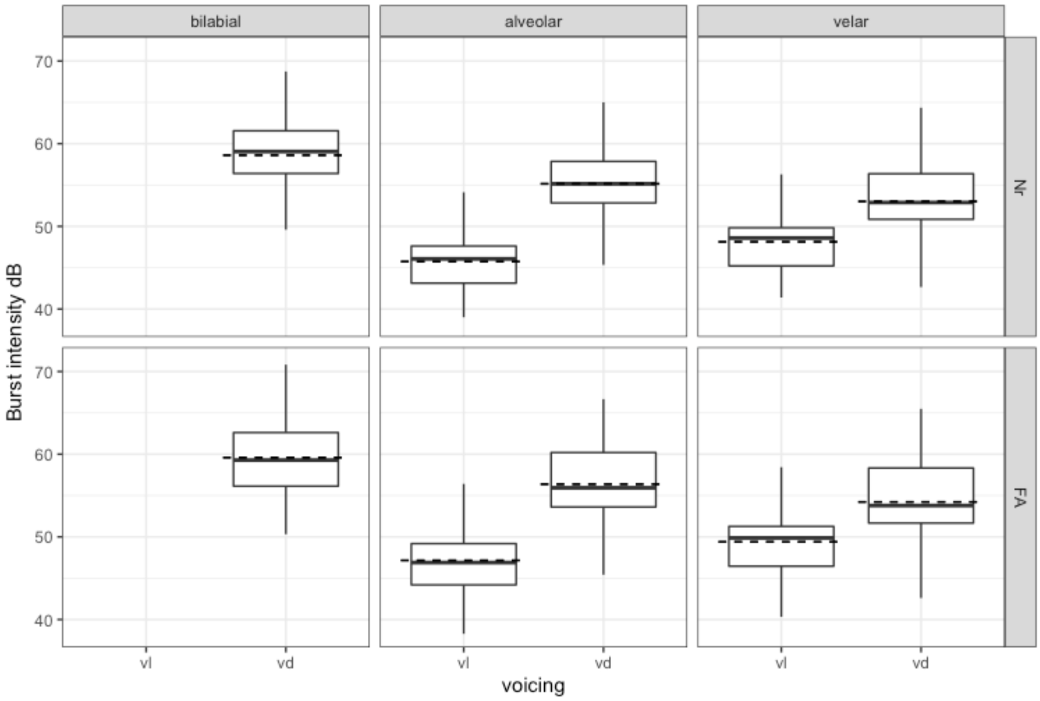


Figure 7.4. Boxplots of the fitted values of burst intensity for utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	58.6	0.445	-	-	59.6	0.756
Alveolar	45.7	0.606	55.1	0.508	47.2	0.9	56.4	0.802
Velar	48.1	0.84	53	0.685	49.4	1.1	54.2	0.93

Table 7.4. Mean and standard deviation of the fitted values of burst intensity for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and place of articulation.

Figure and table 7.4 display that burst intensity for the Voiced was significantly higher than for the Voiceless stops by an average of 9.3 dB for alveolar stops in both normal and fast speech (p<0.0001). The results also showed that burst intensity for the Voiced was significantly higher than for the Voiceless stops by an average of 4.85 dB for velar stops in both normal and fast speech (p<0.0001). Moving to speech rate effect, fast speech resulted in significantly higher burst intensity than normal speech by an average of 1.3 dB (p<0.0001) across places of articulation and voicing categories. With respect to the impact of place of articulation on burst intensity for Voiceless stops, the pattern observed was /velar/ > /alveolar/

by an average of 2.3 dB (p<0.0001) in both normal and fast speech. With respect to the impact of place of articulation on burst intensity for Voiced stops, the pattern observed was /bilabial/ > /alveolar/ > /velar/ in both normal and fast speech. The differences between each pair of comparison were found to be statistically significant (pairwise comparison: Appendix C Table C-85).

7.1.3 Utterance-medial stops CVCV:C (Iambic)

Burst intensity as a function of voicing and rate

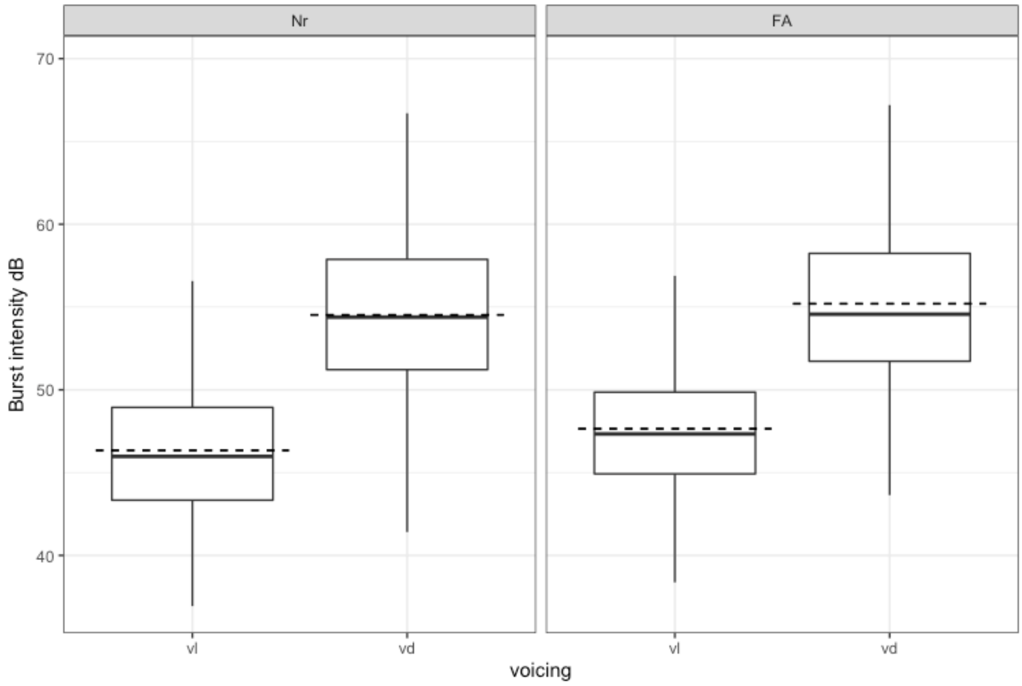


Figure 7.5. Boxplots of the fitted values of burst intensity for utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate. The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	46.3	4.11	54.5	5.27
Fast	47.6	4.14	55.2	497

Table 7.5. Mean and standard deviation of the fitted values of burst intensity for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

The analysis revealed that there was a significant impact of voicing category on burst intensity in utterance-medial stops (Iambic). The values of burst intensity for Voiceless and Voiced stops showed that there was a significant increase in burst intensity for the Voiced stops by an average of 7.9 dB (p<0.0001) in both normal and fast speech rates. With regard to the impact of speech rate on burst intensity, burst intensity in fast speech was significantly higher than in normal speech by an average of 1 dB (p<0.0001) across voicing categories.

Burst intensity as a function of voicing, rate, and place of articulation

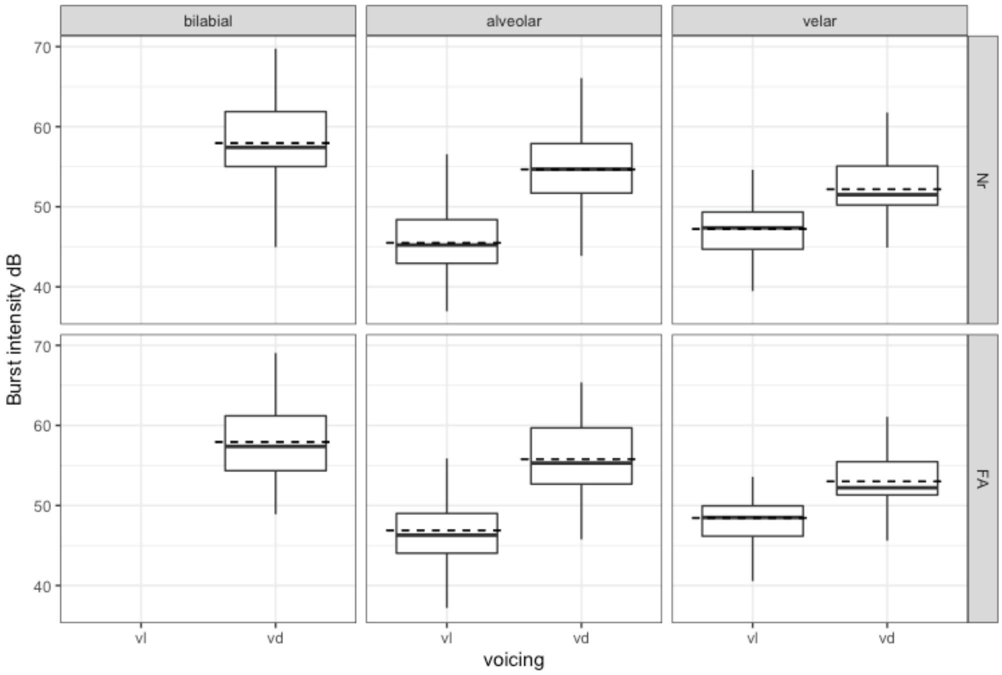


Figure 7.6. Boxplots of the fitted values of burst intensity for utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	58	5.24	-	-	57.9	4.9
Alveolar	45.5	4.16	54.7	4.75	46.9	4.16	55.8	4.7
Velar	47.2	3.88	52.2	4.46	48.4	3.97	53	4.23

Table 7.6. Mean and standard deviation of the fitted values of burst intensity for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and place of articulation.

Figure and table 7.6 display that burst intensity for the Voiced was significantly higher than for the Voiceless stops by an average of 9.05 dB for alveolar stops in both normal and fast speech (p<0.0001). The results also showed that burst intensity for the Voiced was significantly higher than for the Voiceless stops by an average of 4.8 dB for velar stops in both normal and fast speech (p<0.0001). Moving to speech rate effect, fast speech resulted in significantly higher burst intensity than normal speech by an average of 0.9 dB (p<0.0001) across places of articulation and voicing categories. With respect to the impact of place of articulation on burst intensity for Voiceless stops, the pattern observed was /velar/ > /alveolar/ by an average of 1.6 dB in both normal (p<0.0001) and fast (p = 0.00025) speech. With respect to the impact of place of articulation on burst intensity for Voiced stops, the pattern observed was /bilabial/ > /alveolar/ > /velar/ in both normal and fast speech. The differences between each pair of comparison were found to be statistically significant (pairwise comparison: Appendix C Table C-90).

7.1.4 Utterance-final stops CV:<u>C</u>

Burst intensity as a function of voicing and rate

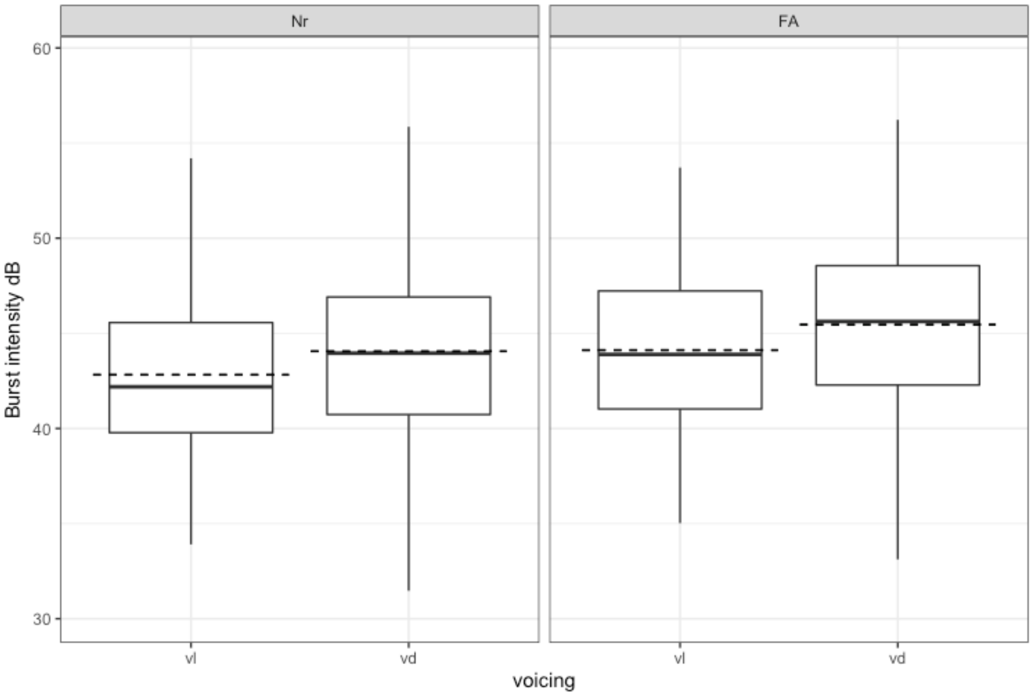


Figure 7.7. Boxplots of the fitted values of burst intensity for Voiced utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate. The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	42.8	4.13	44.1	4.78
Fast	44.1	3.69	45.5	4.56

Table 7.7. Mean and standard deviation of the fitted values of burst intensity for Voiceless and Voiced utterance-final stops grouped by voicing and speech rate.

Figure and table 7.7 show that voicing category affected burst intensity across speech rates. Burst intensity values for Voiceless and Voiced stops showed that there was a significant increase in the Voiced stops by an average of 1.35 dB (p<0.0001) in both normal and fast speech rates. As for the impact of speech rate on burst intensity, burst intensity in fast speech was significantly higher than in normal speech by an average of 1.35 dB (p<0.0001) across voicing categories.

Burst intensity as a function of voicing, rate, and place of articulation

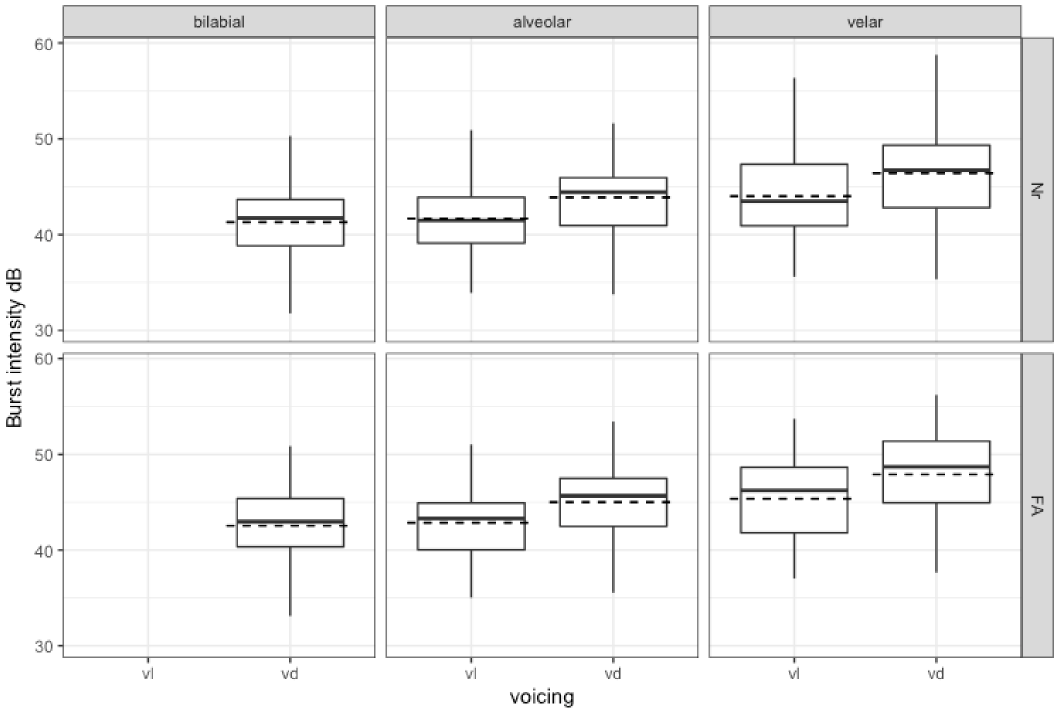


Figure 7.8. Boxplots of the fitted values of burst intensity for Voiced utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	41.3	4.13	-	-	42.6	4.03
Alveolar	41.7	3.62	43.9	4.1	42.9	3.43	45	3.81
Velar	44	4.3	46.4	4.66	45.4	4.06	47.9	4.32

Table 7.8. Mean and standard deviation of the fitted values of burst intensity for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and place of articulation.

Figure and table 7.8 display that burst intensity for the Voiced was significantly higher than for the Voiceless stops by an average of 2.2 dB across places of articulation in both normal and fast speech (p<0.0001). With regard to speech rate effect, fast speech resulted in significantly higher burst intensity than normal speech by an average of 1.3 dB (p<0.0001) across places of articulation and voicing categories. With respect to the impact of place of articulation on burst intensity for Voiceless stops, the pattern observed was /velar/ > /alveolar/ by an average of 1.6 dB (p<0.0001) in both normal and fast speech. As for the impact of place of articulation on burst intensity in Voiced stops, the pattern observed was /velar/ > /alveolar/ > /bilabial/ in both normal and fast speech. The differences between each pair of comparison were found to be statistically significant (pairwise comparison: Appendix C Table C-95).

7.1.5 Summary of results for burst intensity

It has been shown that burst intensity for Voiced stops is significantly higher than for Voiceless stops which makes it a clear acoustic correlate that signals the distinction between Voiceless and Voiced stops in Najdi Arabic. This difference appeared across positions, speech rates, places of articulation, vowel types, and genders. Furthermore, the difference in burst intensity between Voiced and Voiceless stops was relatively smaller in the utterance-final position than in the rest of the positions. A possible reason may be related to the devoicing of utterance-final stops that may lead to the lowering of burst intensity for Voiced stops. The results also show that fast speech increases the burst intensity in all tested contexts. The results additionally show a gender effect in which burst intensity in males' speech is higher than in females' speech across contexts.

It can be observed that the results for burst intensity were against the expectation presented in the previous literature in which Voiceless stops showed higher values for burst intensity (van Alphen and Smits, 2004). The results were in agreement with burst intensity values presented in the work of Al-Tamimi and Khattab (2018) in their study of voicing contrast in Lebanese Arabic. This might suggest that the intensity of the burst in Lebanese and Najdi Arabic shows a language specific feature. More studies are needed to confirm this pattern.

7.2 F0 onset (Hz)

The present analysis focuses on F0 onset in various phonetic contexts and in different speech rates to test its robustness to signal the distinction between Voiced and Voiceless stops. F0 onset has been investigated in a plethora of studies that focus on the voicing contrast in many languages. The majority of these studies found that F0 onset has higher values following Voiceless stops than Voiced stops in aspirating and voicing languages (House and Fairbanks, 1953; Lehiste and Peterson, 1961; Kulikov, 2012). The importance of F0 onset for the present study stems from the fact that voicing contrast in Najdi Arabic, as discussed in chapter 2, has been described as a dialect that has the features of both voicing and aspirating languages, which make these results potentially interesting due to the articulatory justifications that link F0 onset raising to stiffening of the vocal cords and F0 lowering to slackness of vocal cords (Halle and Stevens, 1971; Jessen, 2001).

7.2.1 Utterance-initial stops C̲V:C

F0 onset as a function of voicing and rate

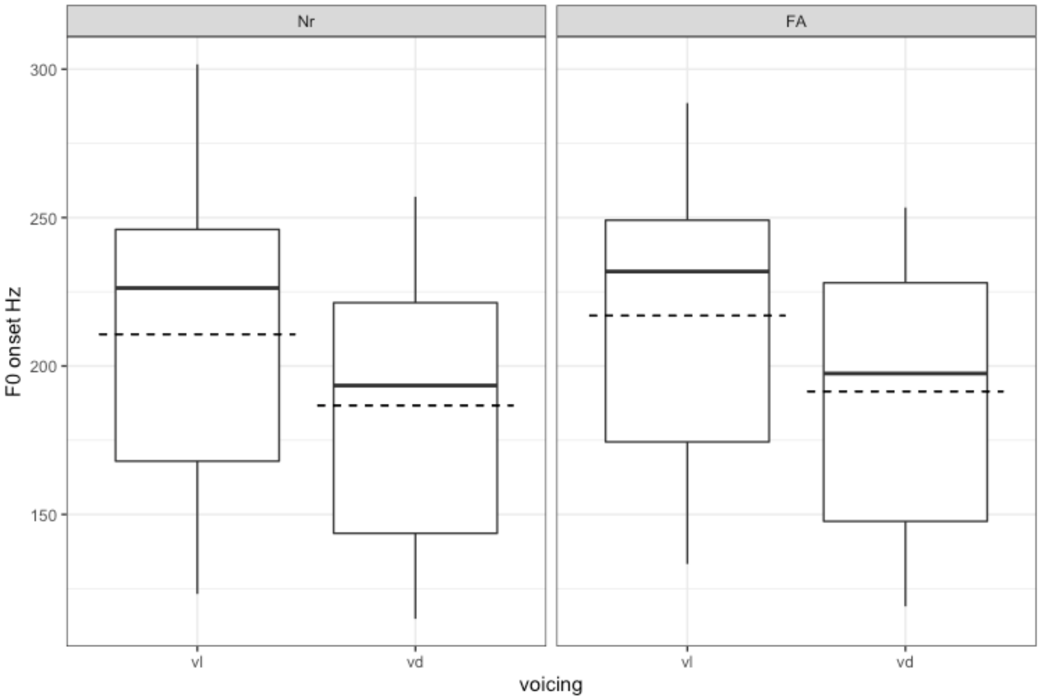


Figure 7.9. Boxplots of the fitted values of F0 onset for Voiceless and Voiced utterance-initial stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate. The dashed line represents the mean.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	211	45.5	187	41.5
Fast	217	43	191	40

Table 7.9. Mean and standard deviation of the fitted values of F0 onset for Voiceless and Voiced utterance-initial stops grouped by voicing and speech rate.

The results for the effect of voicing category and speech rate on F0 onset are given in figure and table 7.9. F0 onset values following Voiceless and Voiced stops showed that there was a significant increase in the Voiceless stops by an average of 25 Hz ($p<0.0001$) in both normal and fast speech rates. For speech rate effect, the results showed that F0 onset in fast speech was significantly higher than in normal speech by an average of 5 Hz in both Voiceless ($p = 0.0024$) and Voiced ($p = 0.0041$) stops.

F0 onset as a function of voicing, rate, and place of articulation

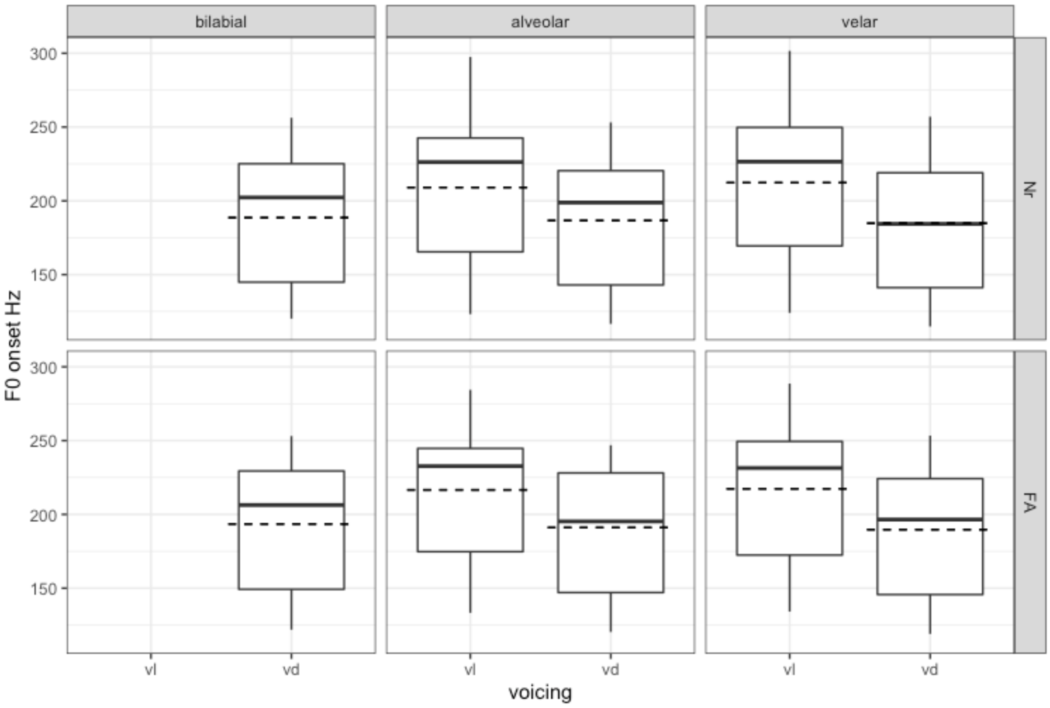


Figure 7.10. Boxplots of the fitted values of F0 onset for Voiceless and Voiced utterance-initial stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and place of articulation. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	189	41.3	-	-	193	39.7
Alveolar	209	45	187	41.4	217	42	191	40.1
Velar	212	46	185	41.8	217	44	190	40.2

Table 7.10. Mean and standard deviation of the fitted values of F0 onset for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate and place of articulation.

Figure and table 7.10 show that F0 onset for the Voiceless was significantly higher than for the Voiced stops by an average of 25.5 Hz across places of articulation in both normal and fast speech ($p<0.0001$). With regard to speech rate effect, fast speech resulted in significantly higher F0 onset than normal speech in the context of Voiceless alveolar stops by an average of 8 Hz ($p = 0.0148$) across. For Voiced alveolar context, however, F0 onset in fast speech was marginally higher than in normal speech by an average of 4 Hz ($p = 0.1641$). In terms of speech rate effect in the context of velar stops, F0 onset in fast speech was marginally higher than in normal speech by an average of 5 Hz following both Voiceless ($p = 0.1178$) and Voiced ($p = 0.1268$) stops. With respect to the impact of place of articulation on F0 onset for Voiceless stops, the pattern observed was /velar/ > /alveolar/ by an average of 3 Hz ($p = 0.3031$) in normal speech. With respect to the impact of place of articulation on F0 onset for

Voiced stops, the pattern observed was /bilabial/ > /alveolar/ > /velar/ in both normal and fast speech. The differences between each pair of comparison were found to be not significant (pairwise comparison: Appendix C Table C-100).

7.2.2 Utterance-medial stops CV:CVC (Trochaic)

F0 onset as a function of voicing and rate

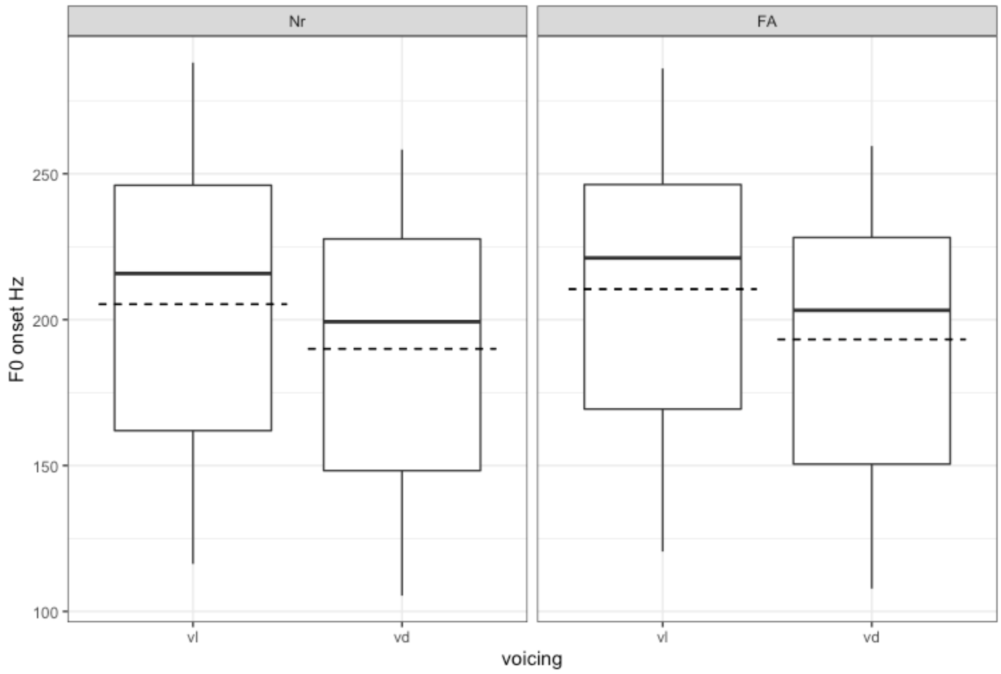


Figure 7.11. Boxplots of the fitted values of F0 onset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	205	44.1	190	41.9
Fast	210	43.1	193	40.9

Table 7.11. Mean and standard deviation of the fitted values of F0 onset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

Figure and table 7.11 display F0 onset values for Voiceless and Voiced stops as a function of voicing and rate. The results show that F0 onset values following Voiceless stops were significantly higher than following Voiced stops by an average of 16 Hz ($p<0.0001$) in both normal and fast speech rates. With regard to the impact of speech rate on F0 onset, F0 onset in fast speech was significantly higher than in normal speech by an average of 4 Hz in the context of both Voiceless ($p = 0.0013$) and Voiced ($p = 0.0099$) stops.

F0 onset as a function of voicing, rate, and place of articulation

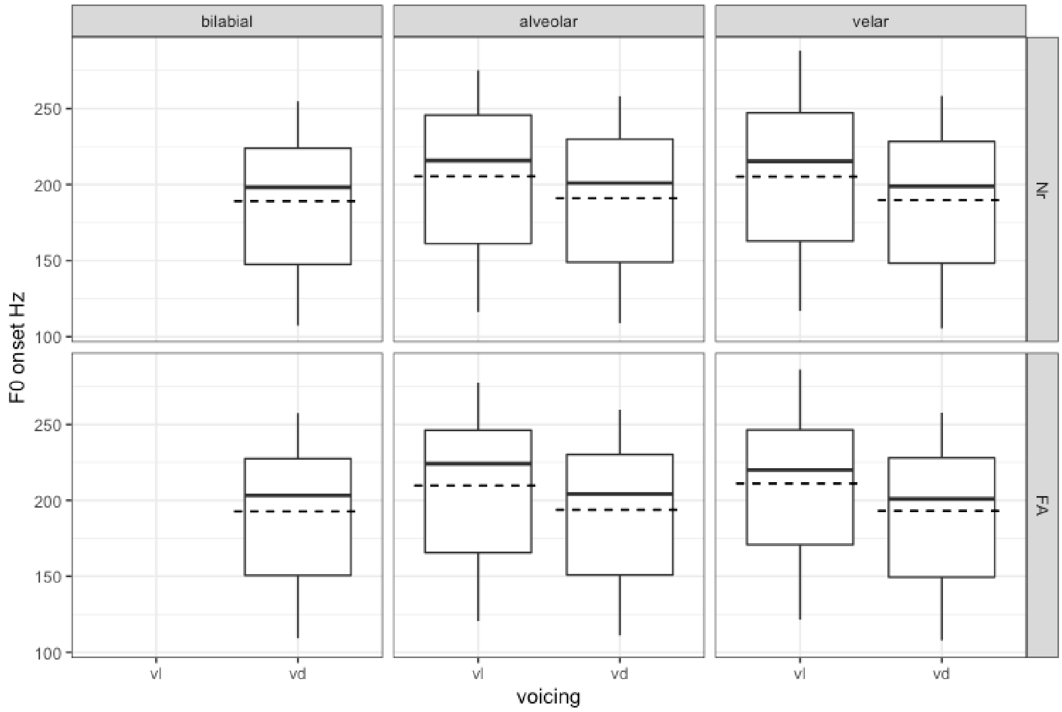


Figure 7.12. Boxplots of the fitted values of F0 onset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing speech rate, and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	189	40.9	-	-	193	39.9
Alveolar	205	43.8	191	41.8	210	43.3	194	41.2
Velar	205	44.5	190	43	211	43	193	41.9

Table 7.12. Mean and standard deviation of the fitted values of F0 onset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and place of articulation.

Figure and table 7.12 show that F0 onset for the Voiceless was significantly higher than for the Voiced stops by an average of 15.75 Hz across places of articulation in both normal and fast speech ($p<0.0001$). With regard to speech rate effect, fast speech resulted in marginally higher F0 onset in comparison to normal speech in the context of bilabial and alveolar stops across voicing categories (pairwise comparison: Appendix C Table C-105). For Voiceless velar stops, F0 onset in fast speech was significantly higher than in normal speech by an average of 6 Hz ($p = 0.014$). With respect to the impact of place of articulation on F0 onset, the differences between each pair of comparison were found to be not significant across voicing categories and speech rates (pairwise comparison: Appendix C Table C-105).

7.2.3 Utterance-medial stops CVCV:C (Iambic)

F0 onset as a function of voicing and rate

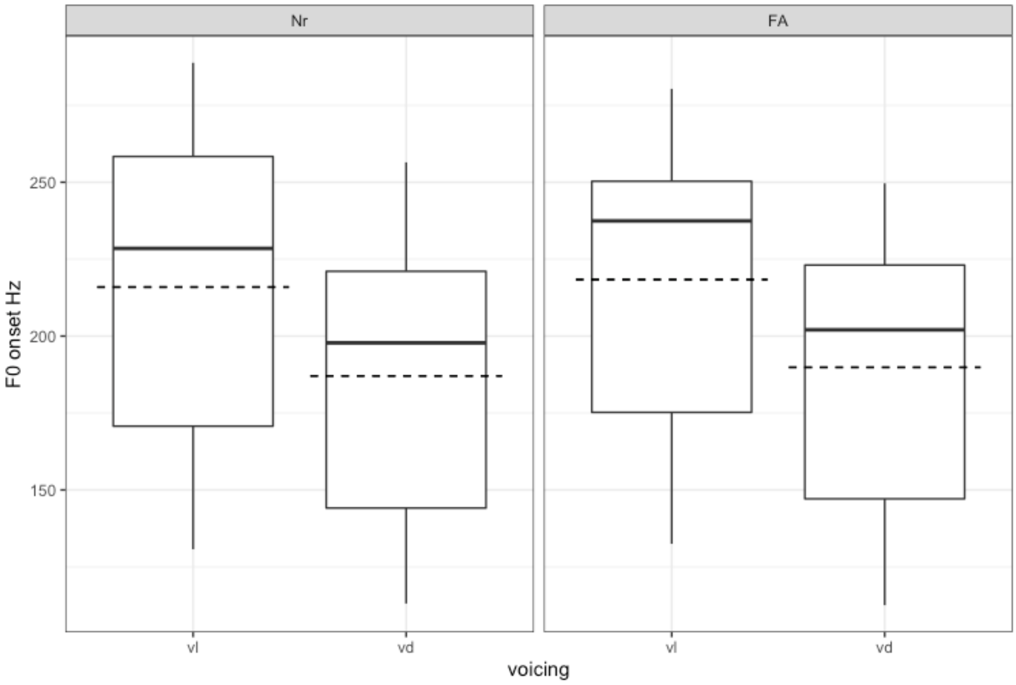


Figure 7.13. Boxplots of the fitted values of F0 onset for Voiceless and Voiced stops in utterance-medial stops (Iambic) grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	216	44.3	187	41.6
Fast	218	41.7	190	39.4

Table 7.13. Mean and standard deviation of the fitted values of F0 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

Figure and table 7.13 present F0 onset values for Voiceless and Voiced stops as a function of voicing and rate. The results show that F0 onset values following Voiceless stops were significantly higher than the ones following Voiced stops by an average of 28.5 Hz ($p<0.0001$) in both normal and fast speech rates. With regard to the impact of speech rate on F0 onset, F0 onset in fast speech tended to be higher than in normal speech by an average of 2.5 Hz in the context of both Voiceless ($p = 0.37$) and Voiced ($p = 0.22$) stops.

F0 onset as a function of voicing, rate, and place of articulation

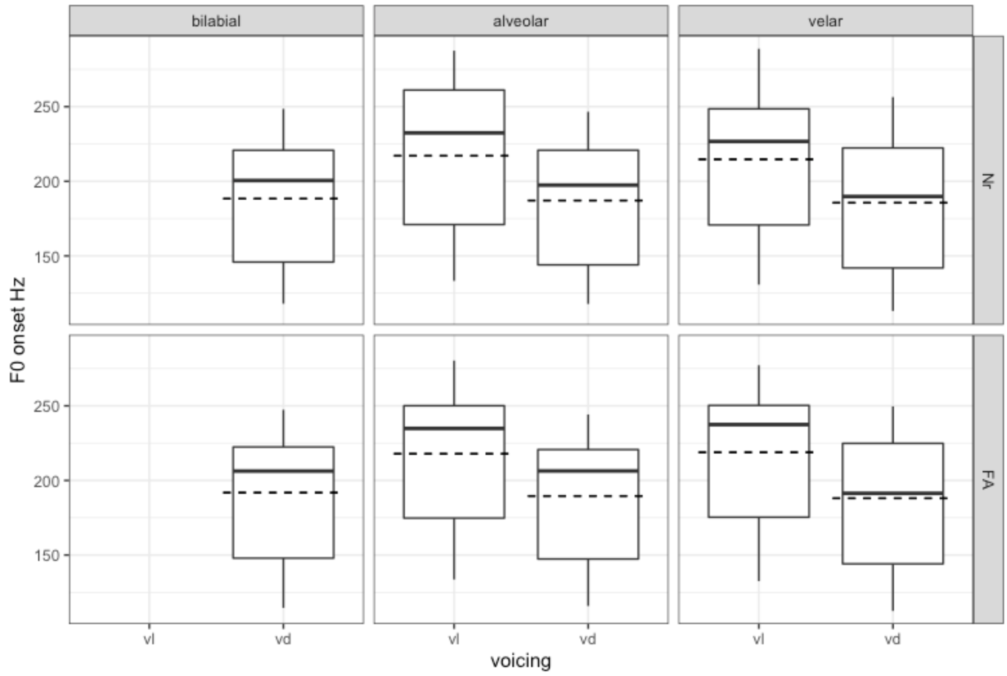


Figure 7.14. Boxplots of the fitted values of F0 onset for Voiceless and utterance-medial stops (Iambic) grouped by voicing speech rate, and place of articulation.

	Voiceless normal		Voiced normal		Voiceless fast		Voiced fast	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	188	41	-	-	192	38.9
Alveolar	217	44.7	187	40.7	218	42.6	190	38.8
Velar	215	43.9	186	43.1	219	40.8	188	40.6

Table 7.14. Mean and standard deviation of the fitted values of F0 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and place of articulation.

Figure and table 7.14 show that F0 onset for the Voiceless was significantly higher than for the Voiced stops by an average of 29.5 Hz across places of articulation in both normal and fast speech ($p<0.0001$). With regard to speech rate effect, fast speech resulted in marginally higher F0 onset in comparison to normal speech by an average of 2.8 Hz across voicing categories and places of articulation (pairwise comparison: Appendix C Table C-110). With respect to the impact of place of articulation on F0 onset, the pattern observed was /alveolar/ > /velar/ by an average of 2 Hz ($p = 0.78$) in normal speech. With respect to the impact of place of articulation on F0 onset for Voiced stops, the pattern observed was /bilabial/ > /alveolar/ > /velar/ in both normal and fast speech. The differences between each pair of comparison were found to be not significant across voicing categories and speech rates (pairwise comparison: Appendix C Table C-110).

7.2.4 Summary of results for F0 onset

The results established for F0 onset were in agreement with what was reported in the literature with regard to the raising of F0 onset following Voiceless stops in comparison to Voiced stops (House and Fairbanks, 1953; Lehiste and Peterson, 1961; Kingston and Diehl, 1995). This difference has been found in Najdi Arabic across positions, speech rates, places of articulation, vowel types, and genders. Such a finding gives an indication of the robustness of F0 onset in marking the distinction between Voiceless and Voiced stops in Najdi Arabic. F0 onset was proposed as correlate for [spread glottis] (or [tense]) in aspirating languages and as a correlate for [voice] in voicing languages.

It could be noticed that the difference in F0 onset as a function of voicing category is larger in utterance-initial and utterance-medial (Iambic) stops than in utterance-medial (Trochaic) stops. One explanation could be that the relation between the amount of aspiration in the stop and F0 onset leads to this positional variation. That is, the amount of aspiration in utterance-initial as well as utterance-medial (Iambic) stops is relatively more than in utterance-medial (Trochaic) due to the stress effect (Jessen, 2001; Ohala, 1983; Stevens, 2000). It has also been found that F0 onset is higher in fast speech than in normal speech.

7.3 F0 offset (Hz)

This section presents the results for F0 at the offset of the vowels preceding Voiced and Voiceless stops in Najdi Arabic. It has been reported that F0 offset is comparable to F0 onset considering that F0 offset is predicted to be lower preceding Voiced stops (Kingston and Diehl, 1995; Jessen, 1998). To check the effectiveness of F0 offset in the voicing contrast in Najdi Arabic, such a correlate was investigated in various contexts.

7.3.1 Utterance-medial stops CV:CVC (Trochaic)

F0 offset as a function of voicing and rate

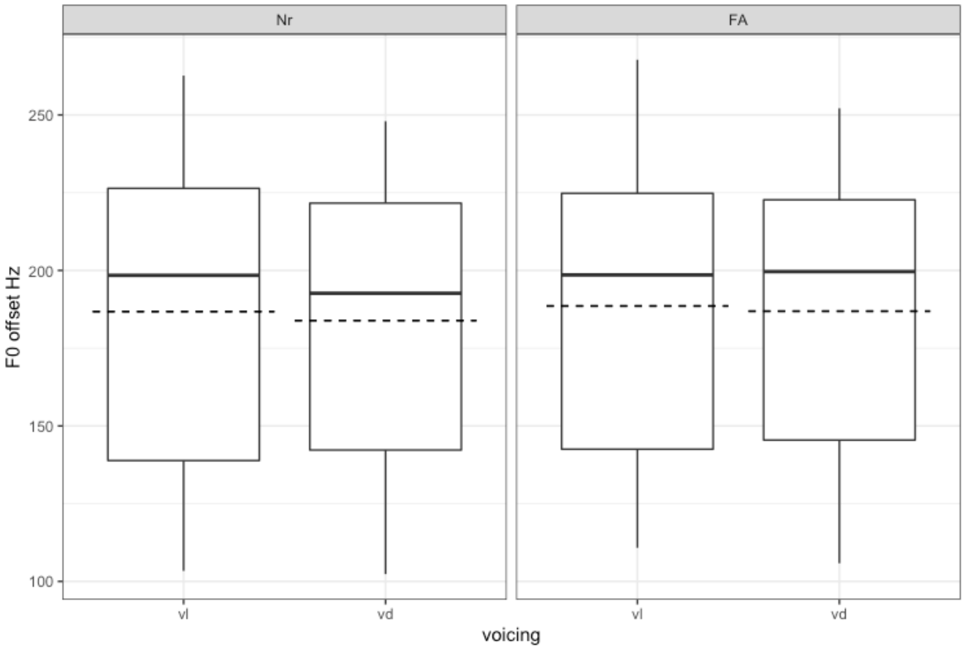


Figure 7.15. Boxplots of the fitted values of F0 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	187	45.7	184	42.3
Fast	189	43	187	40.2

Table 7.15. Mean and standard deviation of the fitted values of F0 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

Figure and table 7.15 show F0 offset values for Voiceless and Voiced stops as a function of voicing and rate. The results show that F0 offset values preceding Voiceless stops tended to be higher than the ones preceding Voiced stops by an average of 2.5 Hz in both normal ($p = 0.42$) and fast ($p = 0.59$) speech rates. With regard to the impact of speech rate on F0 offset, F0 offset in fast speech tended to be higher than in normal speech by an average of 2.5 Hz in the context of both Voiceless ($p = 0.59$) and Voiced ($p = 0.41$) stops.

F0 offset as a function of voicing, rate, and place of articulation

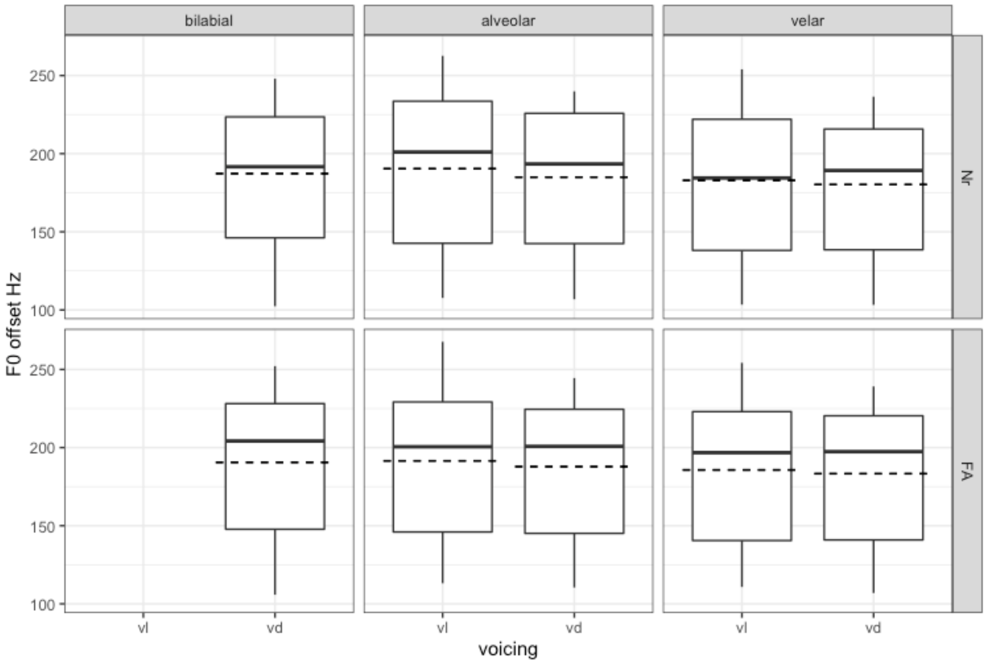


Figure 7.16. Boxplots of the fitted values of F0 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	187	42.9	-	-	191	41.2
Alveolar	190	46.1	185	42.3	192	43.7	188	39.9
Velar	183	45	180	41.7	186	42.3	183	39.4

Table 7.16. Mean and standard deviation of the fitted values of F0 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and place of articulation.

Figure and table 7.16 show that F0 offset for the Voiceless was marginally higher than for the Voiced stops by an average of 3.8 Hz across places of articulation in both normal and fast speech (pairwise comparison: Appendix C Table C-115). With regard to speech rate effect, F0 offset in fast speech tended to be higher than in normal speech by an average of 3 Hz across voicing categories and places of articulation (pairwise comparison: Appendix C Table C-115). With respect to the impact of place of articulation on F0 offset, the pattern observed was /alveolar/ > /velar/ by an average of 6.5 Hz preceding Voiceless stops in normal ($p = 0.168$) and fast ($p = 0.374$) speech. With respect to the impact of place of articulation on F0 offset for Voiced stops, the pattern observed was /bilabial/ > /alveolar/ > /velar/ in both normal and fast speech. The differences between each pair of comparison were found to be not significant across voicing categories and speech rates (pairwise comparison: Appendix C Table C-115).

7.3.2 Utterance-medial stops CVC̲V:C (Iambic)

F0 offset as a function of voicing and rate

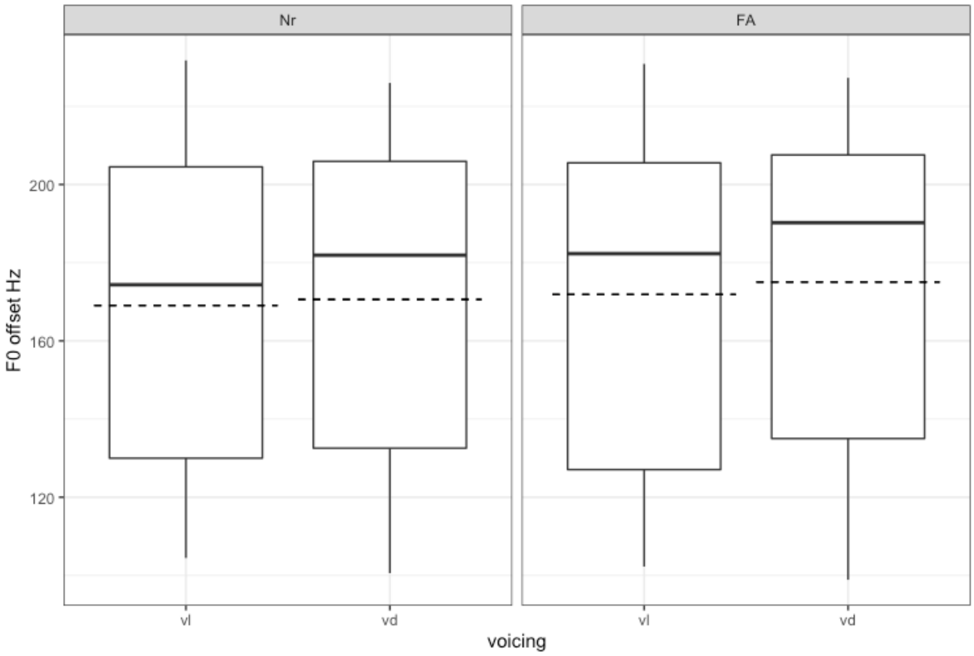


Figure 7.17. Boxplots of the fitted values of F0 offset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	169	40.4	171	38.7
Fast	172	39.2	175	37.9

Table 7.17. Mean and standard deviation of the fitted values of F0 offset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

Figure and table 7.17 display F0 offset values for Voiceless and Voiced stops as a function of voicing and rate. The results show that F0 offset values preceding Voiceless stops tended to be higher than the ones preceding Voiced stops by an average of 2.5 Hz in both normal ($p = 0.593$) and fast ($p = 0.373$) speech rates. With regard to the impact of speech rate on F0 offset, F0 offset in fast speech tended to be higher than in normal speech by an average of 3.5 Hz in the context of both Voiceless ($p = 0.433$) and Voiced ($p = 0.067$) stops.

F0 offset as a function of voicing, rate, and place of articulation

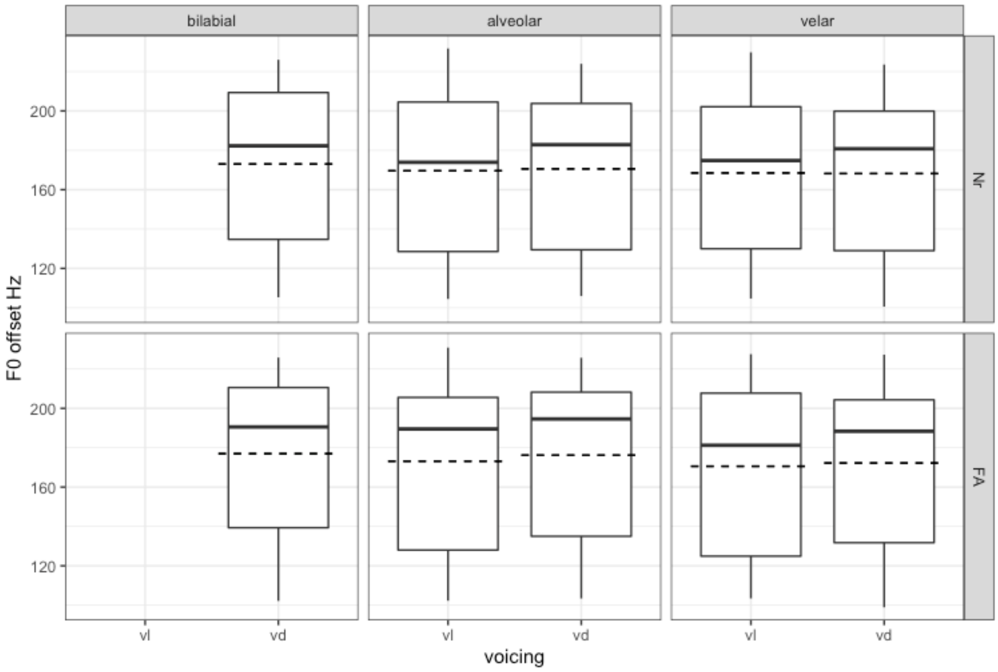


Figure 7.18. Boxplots of the fitted values of F0 offset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing speech rate, and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	173	38.4	-	-	177	37.6
Alveolar	170	40.3	171	38.9	173	38.5	176	38.5
Velar	168	40.6	168	38.7	171	40	172	37.7

Table 7.18. Mean and standard deviation of the fitted values of F0 offset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and place of articulation.

Figure and table 7.18 show that F0 offset for the Voiced was marginally higher than for the Voiceless stops by an average of 2 Hz across places of articulation in both normal and fast speech (pairwise comparison: Appendix C Table C-122). With regard to speech rate effect, F0 offset in fast speech tended to be higher than in normal speech by an average of 3.8 Hz across voicing categories and places of articulation (pairwise comparison: Appendix C Table C-122). With respect to the impact of place of articulation on F0 offset, the pattern observed was /alveolar/ > /velar/ by an average of 3 Hz across speech rates (pairwise comparison: Appendix C Table C-122). With respect to the impact of place of articulation on F0 offset for Voiced stops, the pattern observed was /bilabial/ > /alveolar/ > /velar/ in both normal and fast speech. The differences between each pair of comparison were found to be not significant across voicing categories and speech rates (pairwise comparison: Appendix C Table C-122).

7.3.3 Utterance-final stops CV:C

F0 offset as a function of voicing and rate

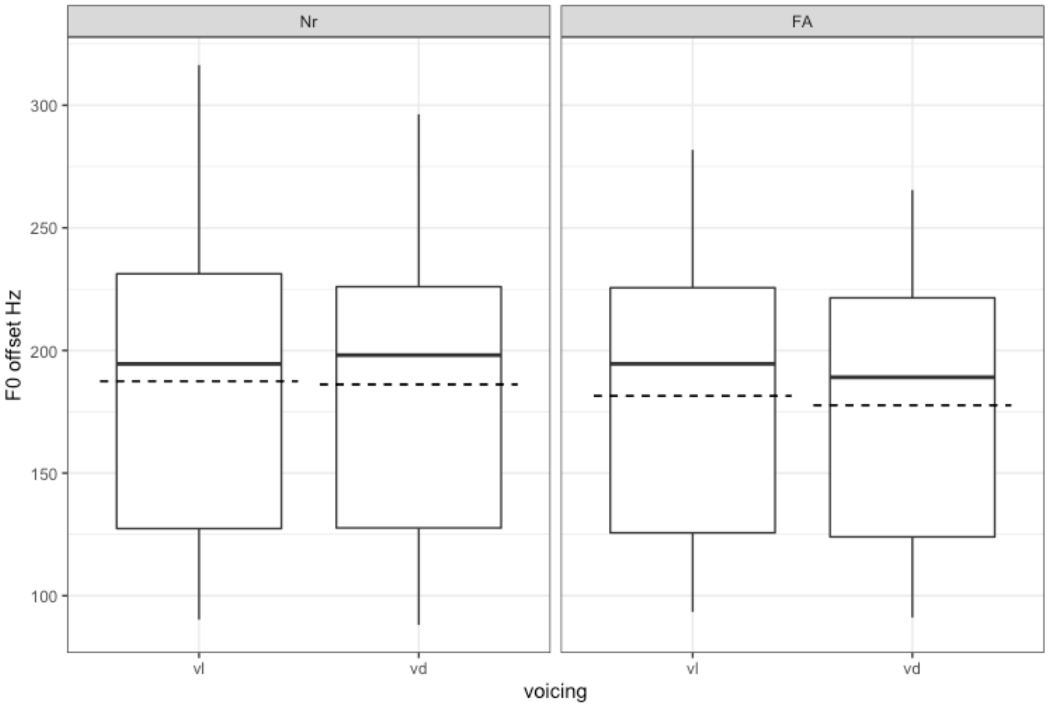


Figure 7.19. Boxplots of the fitted values of F0 offset for Voiceless and Voiced utterance-final stops grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	187	60.3	186	55.3
Fast	181	53.5	178	50.2

Table 7.19. Mean and standard deviation of the fitted values of F0 offset for Voiceless and Voiced utterance-final stops grouped by voicing and speech rate.

Figure and table 7.19 display F0 offset values for Voiceless and Voiced stops as a function of voicing and rate. The results show that F0 offset values for Voiceless stops tended to be higher than for Voiced stops by an average of 2 Hz in both normal ($p = 0.61$) and fast ($p = 0.17$) speech rates. As for the impact of speech rate on F0 offset, F0 offset in normal speech tended to be higher than in fast speech by an average of 6 Hz in the context of Voiceless stops (0.074). The results also showed that F0 offset in normal speech tended to be higher than in fast speech by an average of 8 Hz in the context of Voiced stops (0.0014).

F0 offset as a function of voicing, rate, and place of articulation

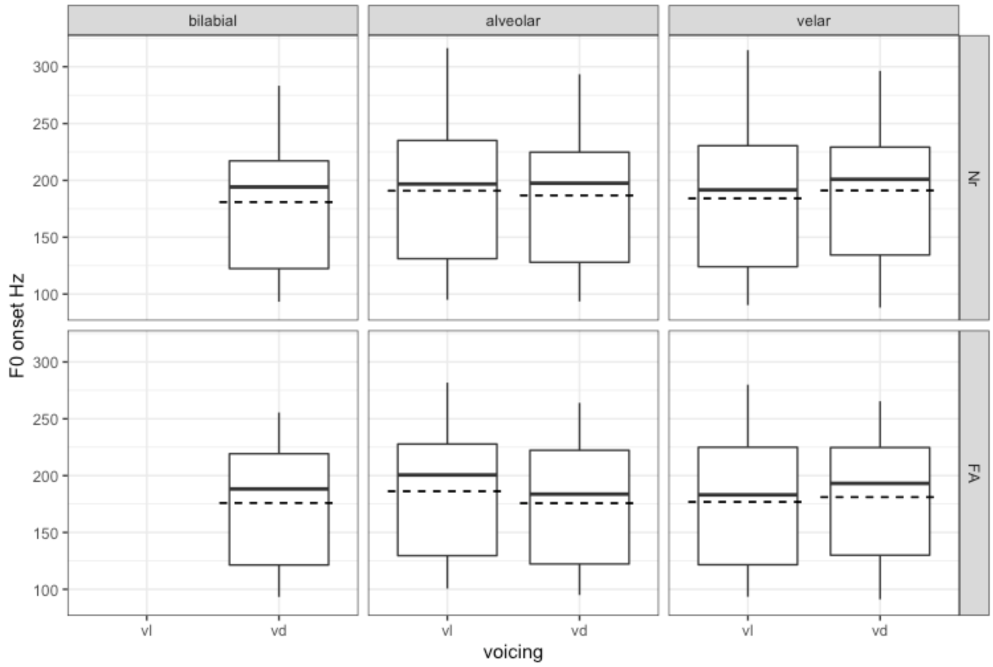


Figure 7.20. Boxplots of the fitted values of F0 offset for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	181	55.3	-	-	176	49.9
Alveolar	191	59.5	187	54.5	186	52.4	176	50.7
Velar	184	61.	191	55.7	177	54.3	181	50

Table 7.20. Mean and standard deviation of the fitted values of F0 offset for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and place of articulation.

Figure and table 7.20 show F0 offset values for Voiced and Voiceless utterance-final stops as a function of voicing category, rate, and place of articulation. With regard to voicing effect on F0 offset, F0 offset for alveolar stops tended to be higher in the Voiceless context than in the Voiced context by an average of 7 Hz across speech rate (pairwise comparison: Appendix C Table C-127). Velar stops on the other hand showed the opposite pattern; F0 offset for Voiced tended to be higher than for Voiceless across speech rates (pairwise comparison: Appendix C Table C-127). With regard to speech rate effect, F0 offset in normal speech tended to be higher than in fast speech by an average of 7.4 Hz across voicing categories and places of articulation (pairwise comparison: Appendix C Table C-127). With respect to the impact of place of articulation on F0 offset, the pattern observed was /alveolar/ > /velar/ by an average of 7 Hz across speech rates (pairwise comparison: Appendix C Table C-127). With respect to the impact of place of articulation on F0 offset for Voiced stops, the pattern observed was

/velar/ > /alveolar/ > /bilabial/ in normal and /velar/ > /alveolar/ =/bilabial/ in fast speech. The differences between each pair of comparison were found to be not significant across voicing categories and speech rates (pairwise comparison: Appendix C Table C-127).

7.3.4 Summary of results for F0 offset

It has been shown that there was inconsistency regarding **F0 offset** values preceding Voiceless and Voiced stops. F0 offset tended to be slightly higher in the context of Voiceless stops in utterance-medial (Trochaic) position only. This difference is not as robust as F0 onset which is in agreement with the expectation from the literature (Kingston and Diehl, 1995; Jansen, 2004). Utterance-final stops behaved differently showing inconsistency and a notable variation when considering the factors tested in the analysis. It could be suggested that this variation is caused by the devoicing of the utterance-final stops. That is, the lowering of F0 offset is a by-product of the voicing in the hold phase which induces an anticipatory effect causing a drop in F0 offset to initiate and sustain voicing (Kingston and Diehl, 1995; Kohler, 1984; Halle and Stevens, 1971). With regard to speech rate effect, there was a positional variation as well in which fast speech accompanies marginal increase in F0 offset in utterance-medial (trochaic) stops. The opposite pattern occurred in utterance-final stops showing higher F0 offset values for normal speech.

7.4 F1 onset (Hz)

This section presents the results for F1 frequency at the onset of the vowel following Voiceless and Voiced stops in Najdi Arabic with respect to various factors including voicing, position, rate, place of articulation, vowel type, and gender. Similar to F0 onset, F1 onset is another acoustic correlate that may mark the opposition between Voiceless and Voiced stops. Many studies reported higher F1 values after aspirated stops in comparison to unaspirated ones (Jessen, 2001; House and Fairbanks, 1953; Ohde, 1984). Such a claim is examined in the present analysis in Najdi Arabic.

7.4.1 Utterance-initial stops CV:C

F1 onset as a function of voicing and rate

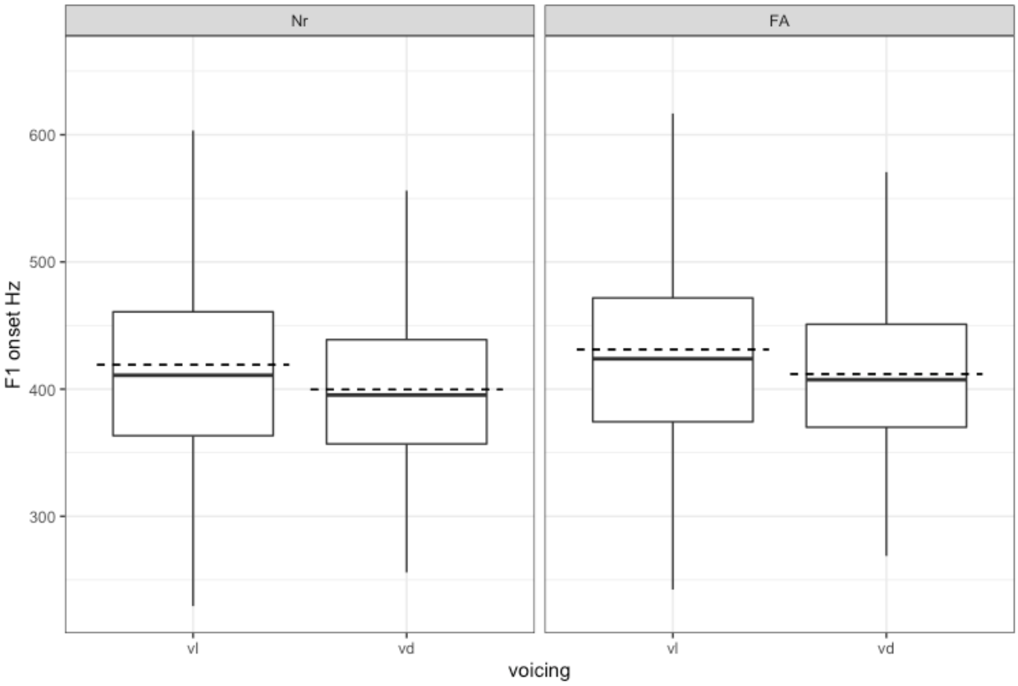


Figure 7.21. Boxplots of the fitted values of F1 onset for Voiceless and Voiced utterance-initial stops grouped by voicing and speech.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	419	85.2	400	66.7
Fast	431	85.1	418	66.5

Table 7.21. Mean and standard deviation of the fitted values of F1 onset for Voiceless and Voiced utterance-initial stops grouped by voicing and speech.

Figure and table 7.21 show the values of F1 onset for Voiceless and Voiced stops as a function of voicing and rate. F1 onset for the Voiceless was significantly higher than for the Voiced stops by an average of 16 Hz across speech rates ($p<0.0001$). For speech rate effect, the results show significant increase in fast speech by an average of 15 Hz across voicing categories ($p<0.0001$).

F1 onset as a function of voicing, rate, and place of articulation

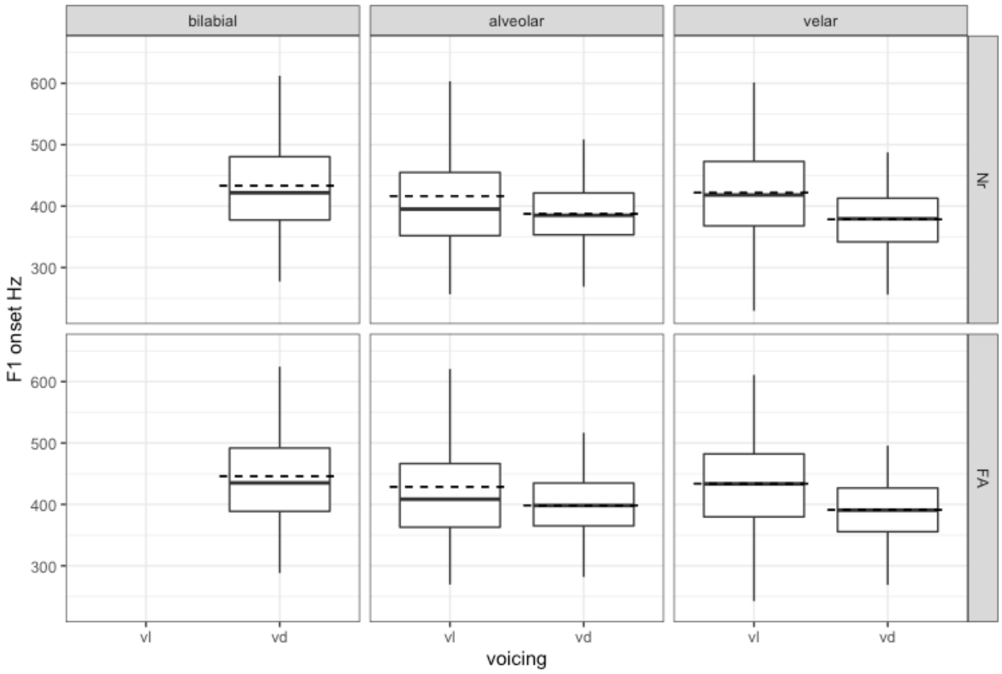


Figure 7.22. Boxplots of the fitted values of F1 onset for Voiceless and Voiced utterance-initial stops grouped voicing, speech rate, and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	433	78.8	-	-	446	78.3
Alveolar	416	90.2	387	53.3	429	90.4	398	53.2
Velar	422	80	378	50.6	434	79.6	391	50.5

Table 7.22. Mean and standard deviation of the fitted values of F1 onset for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate, and place of articulation.

Figure and table 7.22 show F1 onset values for Voiced and Voiceless utterance-initial stops as a function of voicing category, rate, and place of articulation. With regard to voicing effect on F1 onset, F1 onset for alveolar stops was significantly higher in the Voiceless context than in the Voiced by an average of 30 Hz across speech rate ($p<0.0001$). F1 onset for velar stops on the other hand showed that F1 onset for Voiceless was significantly higher than for Voiced stops by an average of 43.5 Hz across speech rates ($p<0.0001$). With regard to speech rate effect, F1 onset in fast speech was significantly higher than in normal speech by an average of 12.4 Hz across voicing categories and places of articulation (pairwise comparison: Appendix C Table C-132). With respect to the impact of place of articulation on F1 onset, the pattern observed was /velar/ >/alveolar/ by an average of 6.5 Hz in the Voiceless context across speech rates (pairwise comparison: Appendix C Table C-132). With respect to the impact of place of articulation on F1 onset for Voiced stops, the pattern observed was /bilabial/ >

/alveolar/ > /velar/ across speech rates. Within bilabial vs alveolar, the former showed higher F1 onset values by an average of 47 Hz (p<0.0001) across speech rates. For alveolar vs velar, the difference was found to be not significant across speech rate (pairwise comparison: Appendix C Table B-132).

7.4.2 Utterance-medial stops CVCV:C (Iambic)

F1 onset as a function of voicing and rate

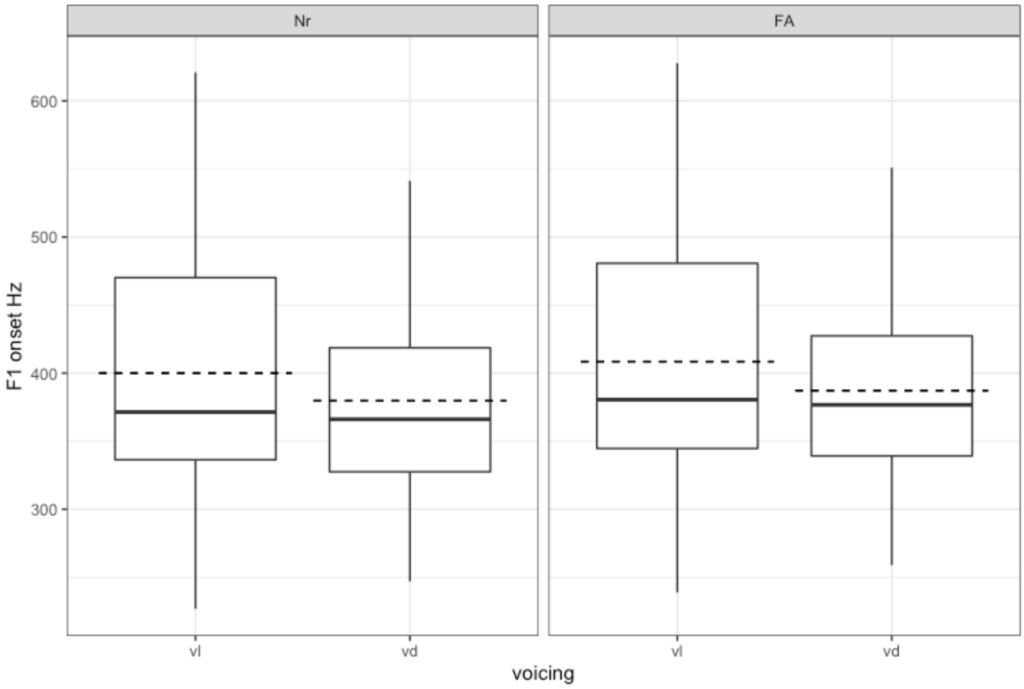


Figure 7.23. Boxplots of the fitted values of F1 onset for Voiceless Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	400	88.1	380	67.3
Fast	408	87.2	387	66.3

Table 7.23. Mean and standard deviation of the fitted values of F1 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

Figure and table 7.23 show the values of F1 onset for Voiceless and Voiced utterance-medial stops (Iambic) as a function of voicing and rate. F1 onset for the Voiceless was significantly higher than for the Voiced stops by an average of 20.5 Hz across speech rates (p<0.0001). For speech rate effect, the results show a marginal increase in fast speech by an average of 7.5 Hz in the context of Voiceless (p = 0.066) and Voiced (p = 0.055) stops.

F1 onset as a function of voicing, rate, and place of articulation

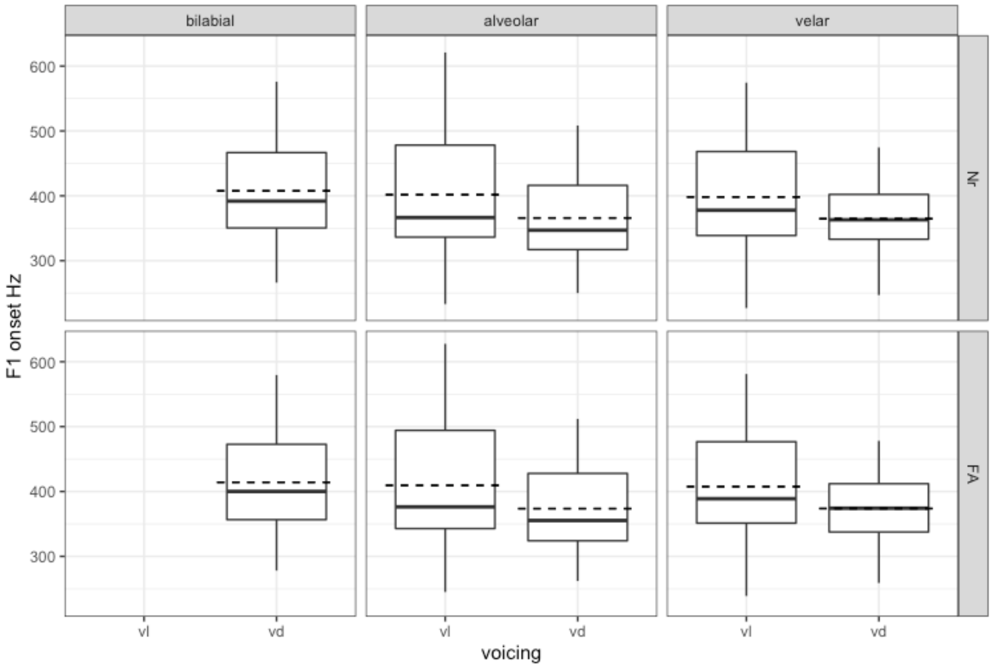


Figure 7.24. Boxplots of the fitted values of F1 onset for Voiceless Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	408	76.9	-	-	414	76.2
Alveolar	402	93.4	366	64	409	93.2	374	62.9
Velar	398	82.5	364	48.5	408	80.8	374	48.5

Table 7.24. Mean and standard deviation of the fitted values of F1 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and place of articulation.

Figure and table 7.24 show F1 onset values for Voiced and Voiceless utterance-medial stops (Iambic) as a function of voicing category, rate, and place of articulation. With regard to voicing effect on F1 onset, F1 onset was significantly higher in the Voiceless context than in the Voiced by an average of 35 Hz across speech rates and places of articulation ($p<0.0001$). With regard to speech rate effect, F1 onset in fast speech was marginally higher than in normal speech by an average of 8.6 Hz across voicing categories and places of articulation (pairwise comparison: Appendix C Table C-137). With respect to the impact of place of articulation on F1 onset, the pattern observed was /alveolar/ >/alveolar/ by an average of 2.5 Hz in the Voiceless context across speech rates (pairwise comparison: Appendix C Table C-137). With respect to the impact of place of articulation on F1 onset for Voiced stops, the pattern observed was /bilabial/ > /alveolar/ > /velar/ across speech rates. Within bilabial vs alveolar, the former showed higher F1 onset values by an average of 41.5 Hz ($p<0.0001$)

across speech rates. For alveolar vs velar, the difference was found to be not significant across speech rate (pairwise comparison: Appendix C Table C-137).

7.4.3 Summary of results for F1 onset

The results of **F1 onset** showed a significant increase in the context of Voiceless stops across positions, speech rates, places of articulation, vowel type, and genders. It has been established that F1 onset is a robust acoustic correlate that differentiates Voiceless and Voiced stops in Najdi Arabic. This pattern was found in many languages and confirmed for aspirating and voicing languages (House and Fairbanks, 1953; Ohde, 1984; Kingston and Diehl, 1994; Jessen, 1998; Al-Tamimi and Khattab, 2018). The positional variations in F1 onset between utterance-initial and utterance-medial (Iambic) stops could be explained by the amount of aspiration. That is, utterance-initial stops with longer aspiration induce more raising of F1 onset than that of utterance-medial stops. Interestingly, the opposite pattern took place with regard F1 onset in the context of Voiced stops. To illustrate, Voiced utterance-medial stops trigger more lowering of F1 onset than Voiced utterance-initial stops. That is, voicing in utterance-medial stops is stronger than that in utterance-initial stops considering that voicing is favourable in intervocalic contexts. The active voicing leads to F1 onset lowering as expected due the articulatory adjustments for maintaining voicing duration during closure (Jessen, 2001).

7.5 F1 offset (Hz)

This section displays the results for F1 at the offset of the vowels preceding Voiced and Voiceless stops in Najdi Arabic. F1 offset is expected to be lower preceding Voiced stops (Kingston and Diehl, 1994; Jessen, 1998) and consequently higher preceding Voiceless stops. The results are divided into two subsections: 1) utterance-medial-stops (Trochaic), and 2) utterance-final stop.

7.5.1 Utterance-medial stops (Trochaic) CV:CVC

F1 offset as a function of voicing and rate

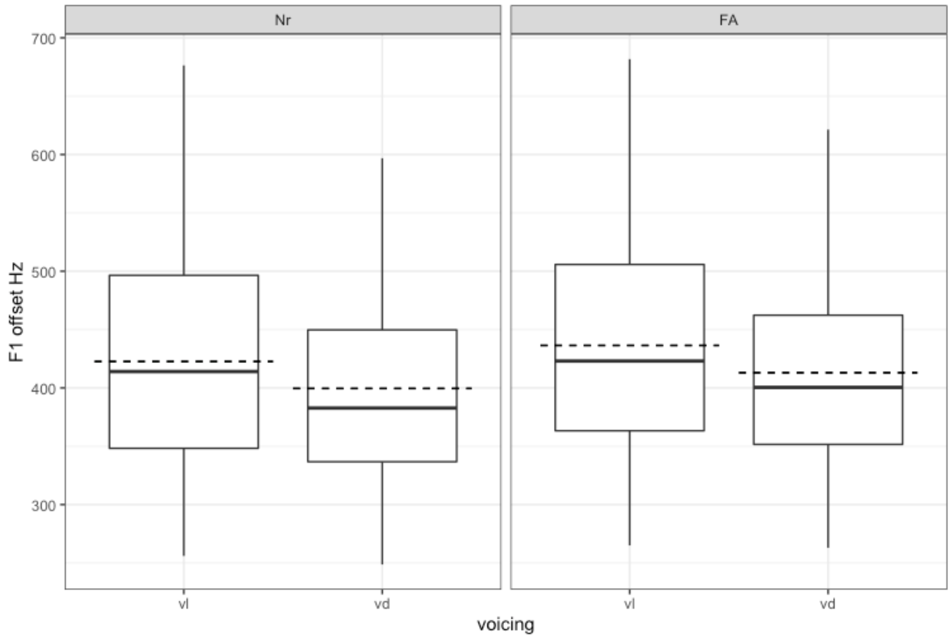


Figure 7.25. Boxplots of the fitted values of F1 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	423	88.8	400	81.7
Fast	436	85.7	413	79.6

Table 7.25. Mean and standard deviation of the fitted values of F1 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

Figure and table 7.25 show the values of F1 offset for Voiceless and Voiced utterance-medial stops (Iambic) as a function of voicing and rate. The results show that F1 offset for the Voiceless was significantly higher than for the Voiced stops by an average of 23 Hz across speech rates (p<0.0001). With regard to speech rate effect, the results show a significant increase of F1 offset in fast speech by an average of 13 Hz in the context of Voiceless (p = 0.007).

F1 offset as a function of voicing, rate, and place of articulation

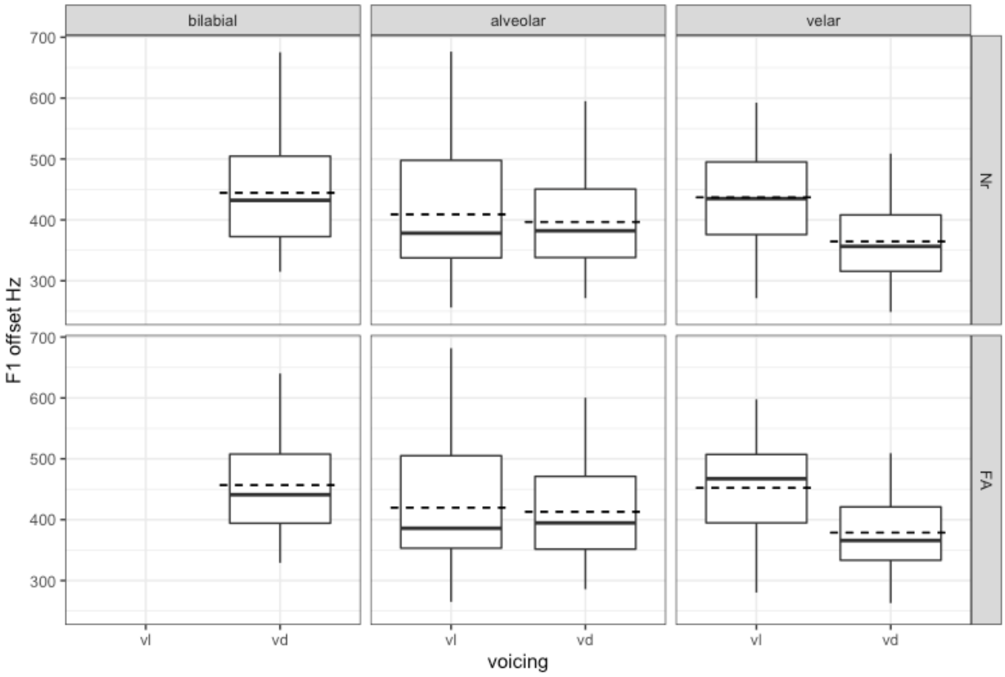


Figure 7.26. Boxplots of the fitted values of F1 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and place of articulation.

Table 7.26. Mean and standard deviation of the fitted values of F1 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	444	83.8	-	-	457	81.4
Alveolar	409	96.2	397	78.7	420	92.1	413	77.8
Velar	437	77.9	364	62.2	452	75.9	379	60.7

Figure and table 7.26 show F1 offset for Voiced and Voiceless utterance-medial stops (Iambic) as a function of voicing, rate, and place of articulation. With regard to voicing effect on F1 offset, F1 offset for alveolar stops was marginally higher in the Voiceless context than in the Voiced by an average of 9.5 Hz across speech rates (pairwise comparison: Appendix C Table C-144). F1 offset for velar stops was significantly higher in the Voiceless context than in the Voiced by an average of 73 Hz across speech rates ($p<0.0001$). With regard to speech rate effect, F1 offset for alveolar stops was marginally higher in the fast speech than in normal by an average of 13.5 Hz across voicing categories (pairwise comparison: Appendix C Table C-144). F1 offset for velar stops was significantly higher in fast speech than in normal speech by an average of 15 Hz across voicing categories (pairwise comparison: Appendix C Table C-144). With respect to the impact of place of articulation on F1 offset, the pattern observed was

/velar/ >/alveolar/ by an average of 30 Hz in the Voiceless context across speech rates ($p<0.0001$). With respect to the impact of place of articulation on F1 offset for Voiced stops, the pattern observed was /bilabial/ > /alveolar/ > /velar/ across speech rates. Within bilabial vs alveolar, the former showed a higher F1 offset values by an average of 54.5 Hz ($p<0.0001$) across speech rates. For alveolar vs velar, the difference was found to be significant by an average of 33.5 Hz across speech rate ($p<0.0001$).

7.5.2 Utterance-final stops CV:C

F1 offset as a function of voicing and rate

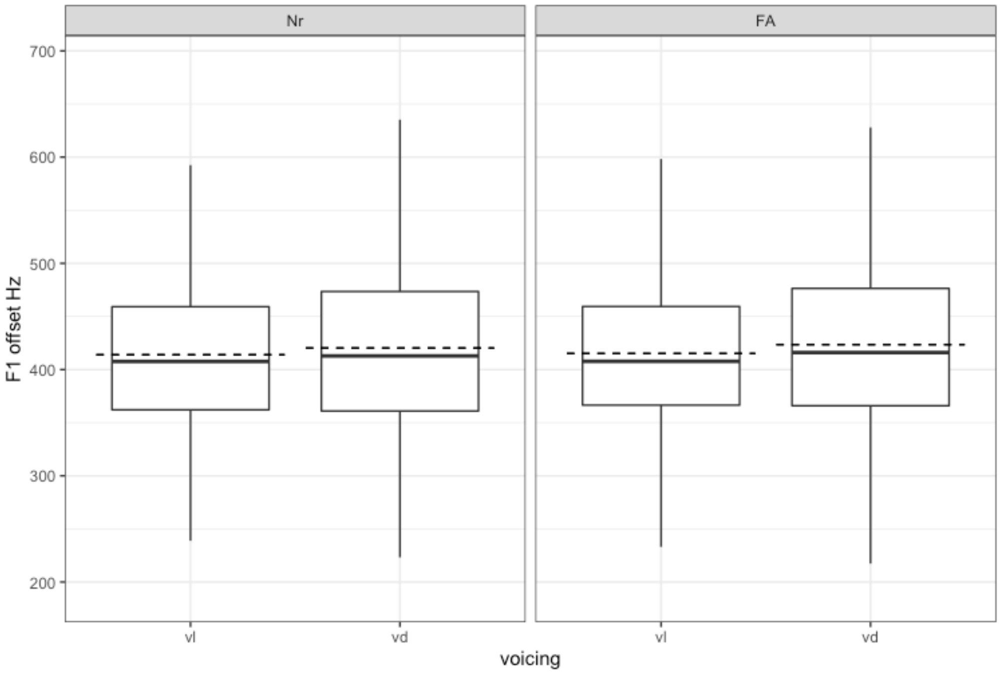


Figure 7.27. Boxplots of the fitted values of F1 offset for Voiceless and Voiced utterance-final stops grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	414	77.5	420	82.1
Fast	415	75	424	81.1

Table 7.27. Mean and standard deviation of the fitted values of F1 offset for Voiceless and Voiced utterance-final stops grouped by voicing and speech rate.

Figure and table 7.27 present F1 offset for Voiceless and Voiced utterance-final stops as a function of voicing and rate. The results show that F1 offset for the Voiced was marginally higher than for the Voiceless stops by an average of 6 Hz in normal speech ($p = 0.124$). F1 offset for the Voiced was significantly higher than for the Voiceless stops by an average of 9

Hz in fast speech (p = 0.03). The results show a marginal increase of F1 offset in fast speech in the context of Voiceless (p = 0.74) and Voiced (p = 0.33) stops.

F1 offset as a function of voicing, rate, and place of articulation

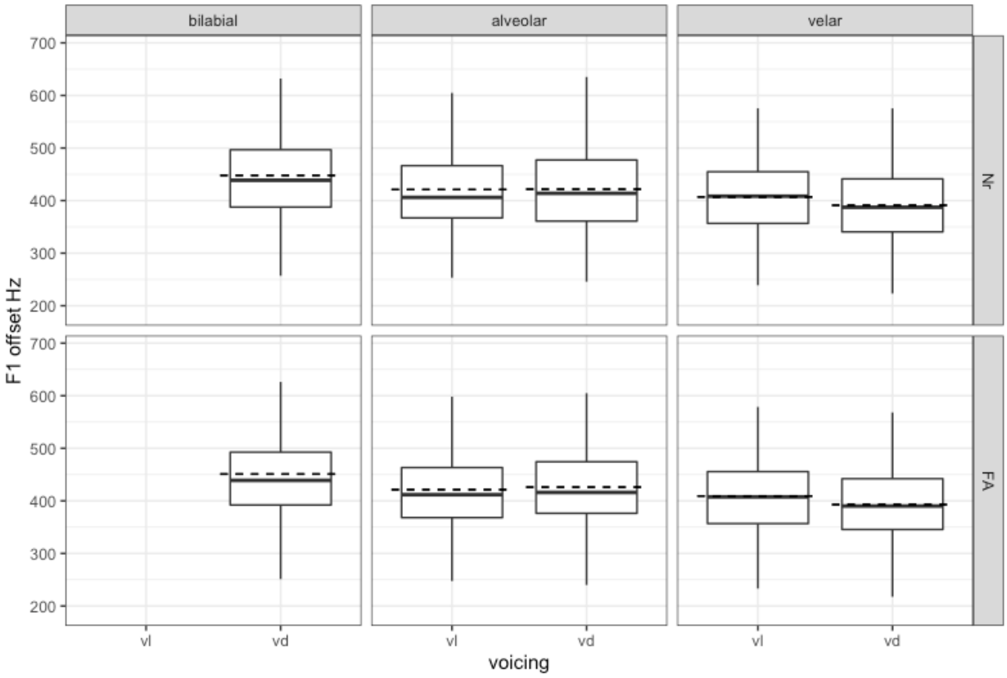


Figure 7.28. Boxplots of the fitted values of F1 offset for Voiceless and utterance-final stops grouped by voicing, speech rate, and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	448	88.2	-	-	451	86.5
Alveolar	421	81.1	422	78.7	421	78.8	426	77.8
Velar	407	72.9	391	68.1	409	70.7	393	67

Table 7.28. Mean and standard deviation of the fitted values of F1 offset for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and place of articulation.

Figure and table 7.28 show F1 offset for Voiced and Voiceless utterance-final stops as a function of voicing, rate, and place of articulation. With regard to voicing effect on F1 offset, F1 offset for alveolar stops was marginally higher in the Voiced context than in the Voiceless by an average of 3 Hz in normal (p = 0.92) and fast (p = 0.41). F1 offset for velar stops was significantly higher in the Voiceless context than in the Voiced by an average of 16 Hz in normal (p = 0.0029) and fast (p = 0.0026) speech rates. With regard to speech rate effect, F1 offset for was marginally higher in the fast speech than in normal by an average of 5.5 Hz across places of articulation and voicing categories (pairwise comparison: Appendix C Table

C-151). With respect to the impact of place of articulation on F1 offset for Voiceless stops, the pattern observed was /alveolar/ >/velar/ by an average of 13 Hz across speech rates (pairwise comparison: Appendix C Table C-151). With respect to the impact of place of articulation on F1 offset for Voiced stops, the pattern observed was /bilabial/ > /alveolar/ > /velar/ across speech rates. All the differences were found to be significant across speech rate (pairwise comparison: Appendix C Table C-151).

7.5.3 Summary of results for F1 offset

It has been established that **F1 offset** values tended to be slightly higher for Voiceless than for Voiced stops in utterance-medial position (Trochaic) with a considerable variation when taking into account the linguistic factors. These results are different from what was found in Lebanese Arabic which showed a consistent significant decrease in the context of Voiced stops (Al-Tamimi and Khattab, 2018). Utterance-final stops showed the opposite pattern in which F1 offset for Voiced was higher than for Voiceless. The devoicing of stops utterance-finally may explain the inconsistency in F1 offset. The results also showed no significant impact of speech rate on F1 offset in utterance-final stops.

7.6 H1-H2 onset (dB)

This section discusses the results of the difference between the amplitude of the first and second harmonics at the onset of the following vowel (H1-H2 onset) in the context of Voiceless and Voiced stops in Najdi Arabic. It has been reported in the literature that this parameter captures the voicing contrast in a number of aspirating languages (Stevens and Hanson, 1994; Klatt and Klatt, 1990; Chapin Ringo, 1988; Jessen, 1998). That is, aspirated stops induce higher values of H1-H2 at the onset of the following vowel compared to unaspirated stops. Some voicing languages showed the opposite pattern as in French and Italian (Ni Chasaide and Gobl, 1993). In Lebanese Arabic, H1-H2 onset’s results showed no significant difference between Voiceless and Voiced stops (Al-Tamimi and Khattab, 2018). It is worth looking at how this correlate is manifested in Najdi Arabic taking into consideration the presence of moderate aspiration in Voiceless stops (Al-Gamdi et al., 2019).

7.6.1 Utterance-initial stops CV:C

H1-H2 onset as a function of voicing and rate

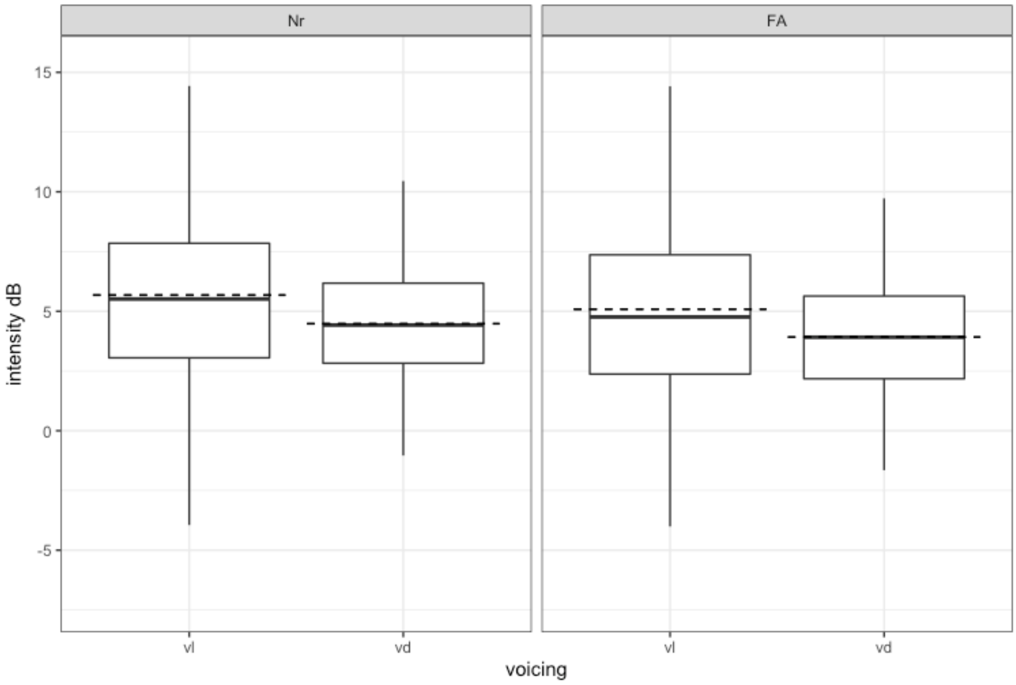


Figure 7.29. Boxplots of the fitted values of H1-H2 onset for Voiceless and Voiced utterance-initial stops grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	5.68	3.84	4.48	2.31
Fast	5.09	3.83	3.92	2.37

Table 7.29. Mean and standard deviation of the fitted values of H1-H2 onset for Voiceless and Voiced utterance-initial stops grouped by voicing and speech rate.

Figure and table 7.29 present the results of H1-H2 onset for Voiceless and Voiced utterance-initial stops as a function of voicing and rate. The results show that H1-H2 onset for the Voiceless was significantly higher than for the Voiced stops by an average of 1.2 dB across speech rates ($p<0.0001$). With regard to speech rate effect, the results show a significant increase of H1-H2 onset in normal speech by an average of 0.6 dB across voicing categories ($p<0.0001$).

H1-H2 onset as a function of voicing, rate, and place of articulation

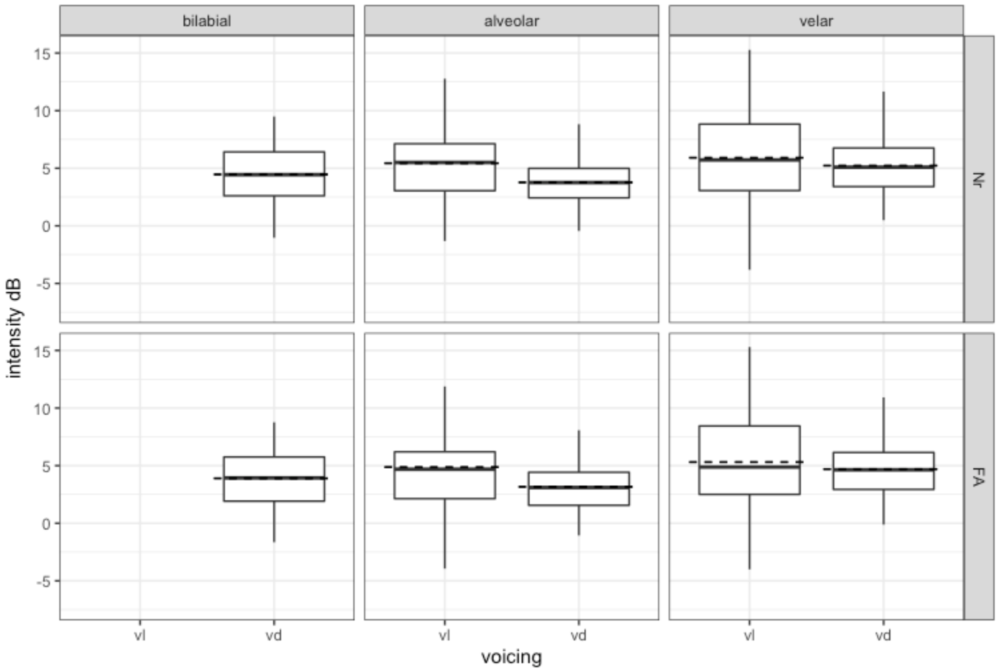


Figure 7.30. Boxplots of the fitted values of H1-H2 onset for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	4.46	2.27	-	-	3.90	2.39
Alveolar	5.44	3.46	3.77	2.08	4.87	3.47	3.16	2.15
Velar	5.91	4.16	5.22	2.35	5.31	4.15	4.70	2.31

Table 7.30. Mean and standard deviation of the fitted values of H1-H2 onset for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate and place of articulation.

Figure and table 7.30 show H1-H2 onset for Voiced and Voiceless utterance-initial stops as a function of voicing, rate, and place of articulation. With regard to voicing effect on H1-H2 onset, H1-H2 onset for alveolar stops was significantly higher in the Voiceless context than in the Voiced by an average of 1.7 dB across speech rates ($p<0.0001$). H1-H2 onset for velar stops was significantly higher in the Voiceless context than in the Voiced by an average of 0.7 dB in normal ($p = 0.00088$) and fast ($p = 0.0045$) speech rates. With regard to speech rate effect, H1-H2 onset was significantly higher in normal speech than in fast by an average of 0.8 dB across places of articulation and voicing categories (pairwise comparison: Appendix C Table C-156). With respect to the impact of place of articulation on H1-H2 onset for Voiceless stops, the pattern observed was /velar/ >/alveolar/ by an average of 0.6 dB across speech rates (pairwise comparison: Appendix C Table C-156). With respect to the impact of place of articulation on H1-H2 onset for Voiced stops, the pattern observed was /velar/ >

/bilabial/ > /alveolar/ across speech rates. All the differences were found to be significant across speech rate (pairwise comparison: Appendix C Table C-156).

7.6.2 Utterance-medial stops CVCV:C (Iambic)

H1-H2 onset as a function of voicing and rate

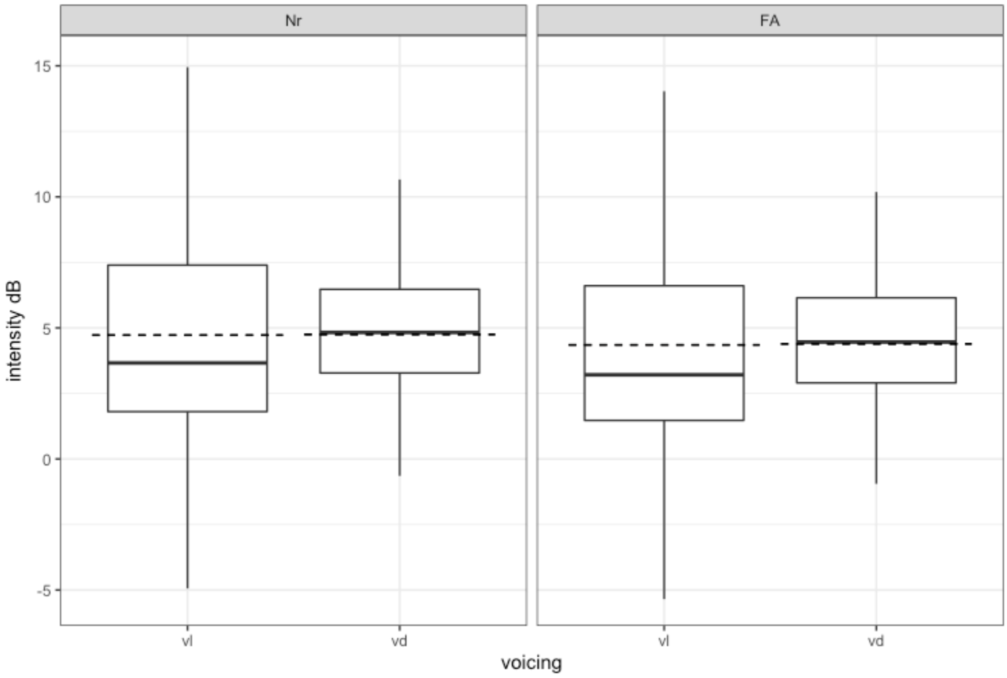


Figure 7.31. Boxplots of the fitted values of H1-H2 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	4.72	3.98	4.75	2.27
Fast	4.35	4.01	4.36	2.36

Table 7.31. Mean and standard deviation of the fitted values of H1-H2 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing and speech rate.

Figure and table 7.31 present H1-H2 onset values for Voiceless and Voiced utterance-medial (iambic) stops as a function of voicing and rate. The results show that H1-H2 onset for the Voiced was slightly higher than for the Voiceless stops by an average of 0.02 dB across speech rates ($p = 0.88$). With regard to speech rate effect, the results show a marginal increase of H1-H2 onset in normal speech by an average of 0.4 dB across voicing categories ($p = 0.095$).

H1-H2 onset as a function of voicing, rate, and place of articulation

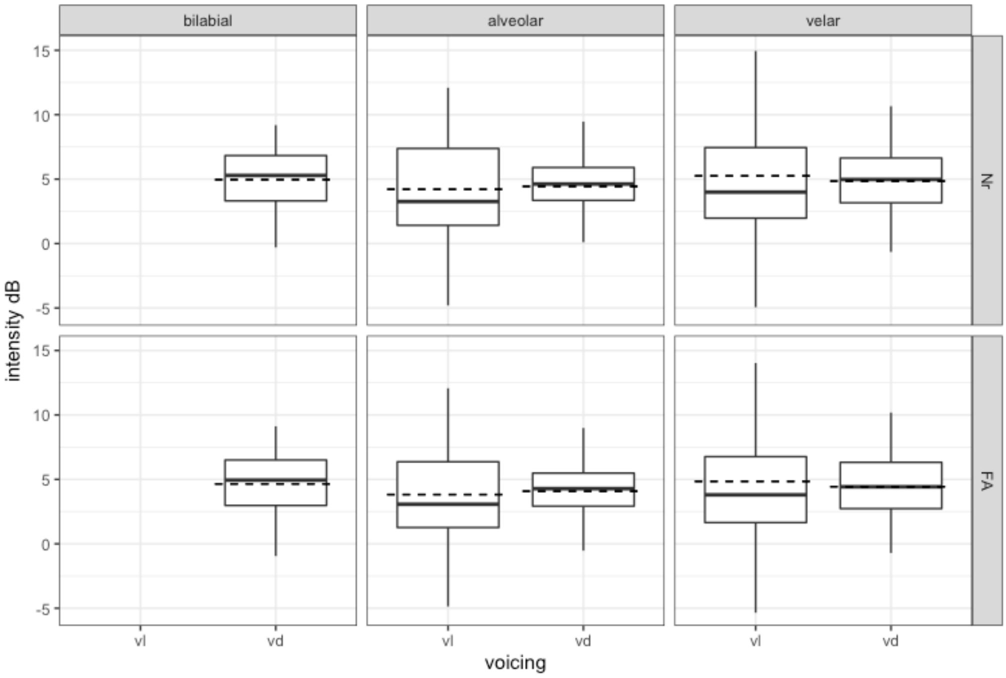


Figure 7.32. Boxplots of the fitted values of H1-H2 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	4.96	2.45	-	-	4.65	2.57
Alveolar	4.22	3.56	4.44	1.89	3.82	3.57	4.08	1.97
Velar	5.25	4.33	4.85	2.4	4.85	4.34	4.44	2.47

Table 7.32. Mean and standard deviation of the fitted values of H1-H2 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and place of articulation.

Figure and table 7.32 show H1-H2 onset for Voiced and Voiceless utterance-medial stops (Iambic) as a function of voicing, rate, and place of articulation. With regard to voicing effect on H1-H2 onset, H1-H2 onset for alveolar stops was marginally higher in the Voiced context than in the Voiceless by an average of 0.24 dB across speech rates (pairwise comparison: Appendix C Table C-159). H1-H2 onset for velar stops was marginally higher in the Voiceless context than in the Voiced by an average of 0.4 dB in normal and fast speech rates (pairwise comparison: Appendix C Table C-159). With regard to speech rate effect, H1-H2 onset was marginally higher in normal speech than in fast by an average of 0.4 dB across places of articulation and voicing categories (pairwise comparison: Appendix C Table C-159). With respect to the impact of place of articulation on H1-H2 onset for Voiceless stops, the pattern observed was /velar/ >/alveolar/ by an average of 1.03 dB across speech rates (pairwise comparison: Appendix C Table C-159). With respect to the impact of place of articulation on H1-H2 onset for Voiced stops, the pattern observed was /velar/ > /bilabial/ > /alveolar/ across speech rates. All the differences were found to be significant across speech rate (pairwise comparison: Appendix C Table C-159).

7.6.3 Summary of results for H1-H2 onset

It has been shown that there was a tendency for H1-H2 to be higher in the context of Voiceless stops compared to Voiced stops in Najdi Arabic. This difference was significant in utterance-initial stops with overlap between the two categories. There was ambiguity in H1-H2 onset pattern in utterance-medial stops showing variations in different vocalic contexts. The higher value of H1-H2 was proposed by Jessen (1998, 2001) to be a correlate for [tense] because of its association with the wider glottal opening that occur during the aspiration phase. The results for high vowels showed the opposite pattern, however. High vowels are produced with larger vocal tract that leads to a delay in the initiation of the voicing onset (Smith, 1976; Ohala, 1983). This articulatory gesture leads to longer aspiration. As it is expected, the aspiration phase is associated with glottal abduction which decreases with time until the end of the phase. Accordingly, the presence of breathiness that occurred at the onset of the vowel is low because the vocal cords are starting to close which consequently means more lower values of H1-H2.

7.7 H1-H2 offset (dB)

This section discusses the results of the difference between the amplitude of the first and second harmonics at the offset of the preceding vowel (H1-H2 offset) in the context of Voiceless and Voiced stops in Najdi Arabic.

7.7.1 Utterance-medial stops CV:CVC (Trochaic)

H1-H2 offset as a function of voicing and rate

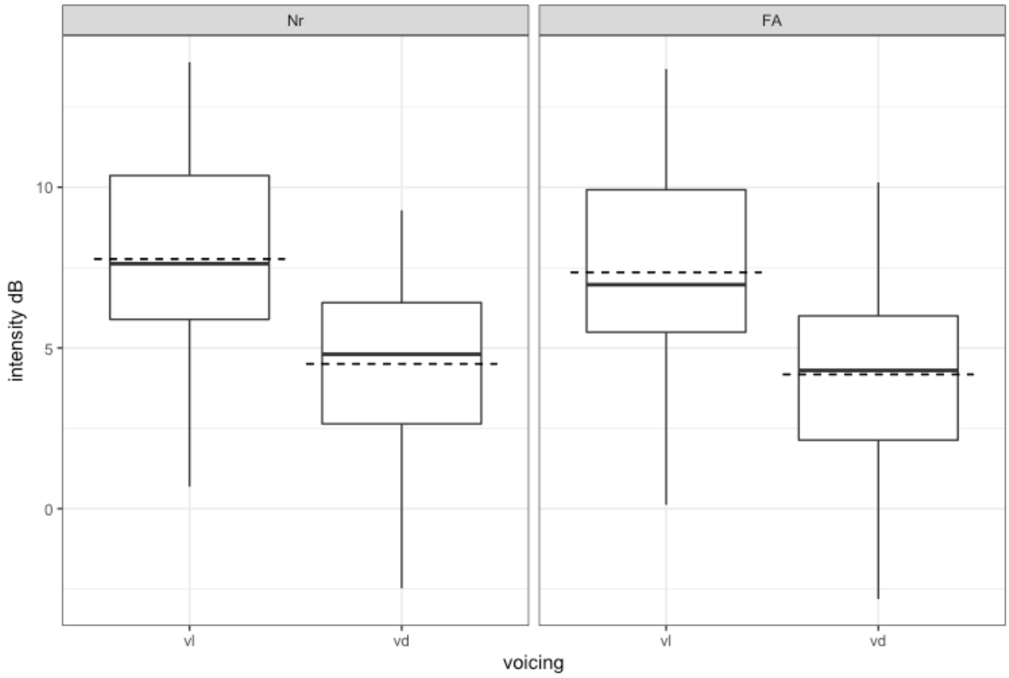


Figure 7.33. Boxplots of the fitted values of H1-H2 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	7.77	3.05	4.51	2.5
Fast	7.36	3.16	4.18	2.7

Table 7.33. Mean and standard deviation of the fitted values of H1-H2 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing and speech rate.

Figure and table 7.33 show H1-H2 offset values for Voiceless and Voiced utterance-medial (Trochaic) stops as a function of voicing and rate. The results show that H1-H2 offset for the Voiced was significantly higher than for the Voiceless stops by an average of 3.22 dB across speech rates ($p<0.0001$). With regard to speech rate effect, the results show a significant increase of H1-H2 offset in normal speech by an average of 0.4 dB across voicing categories ($p = 0.016$).

H1-H2 offset as a function of voicing, rate, and place of articulation

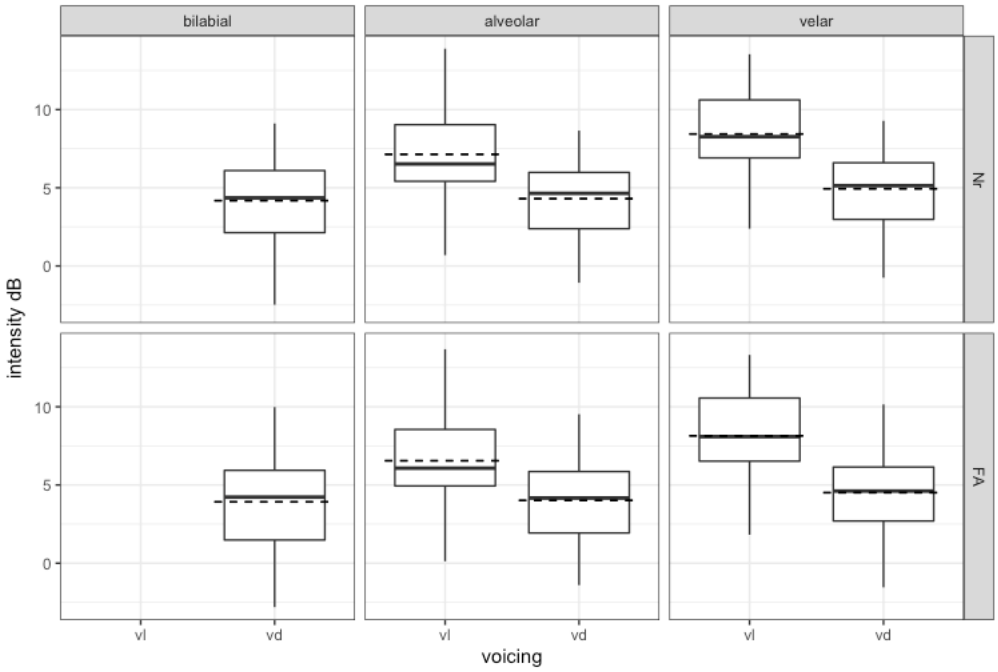


Figure 7.34. Boxplots of the fitted values of H1-H2 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	4.18	2.65	-	-	3.92	2.87
Alveolar	7.14	3.19	4.3	2.41	6.55	3.25	4.01	2.63
Velar	8.44	2.74	4.94	2.37	8.14	2.86	4.51	2.59

Table 7.34. Mean and standard deviation of the fitted values of H1-H2 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and place of articulation.

Figure and table 7.34 show H1-H2 offset for Voiced and Voiceless utterance-medial stops (Trochaic) as a function of voicing, rate, and place of articulation. With regard to voicing effect on H1-H2 offset, H1-H2 offset was significantly higher in the Voiceless context than in the Voiced by an average of 3.12 dB across speech rates (pairwise comparison: Appendix C Table C-164). With regard to speech rate effect, H1-H2 offset was marginally higher in normal speech than in fast speech across places of articulation and voicing categories (pairwise comparison: Appendix C Table C-164). With respect to the impact of place of articulation on H1-H2 offset for Voiceless stops, the pattern observed was /velar/ >/alveolar/ by an average of 1.4 dB across speech rates ($p<0.0001$). With respect to the impact of place of articulation on H1-H2 offset for Voiced stops, all the differences were found to be significant across speech rate categories (pairwise comparison: Appendix C Table C-164).

7.7.2 Utterance-final stops CV:C

H1-H2 offset as a function of voicing and rate

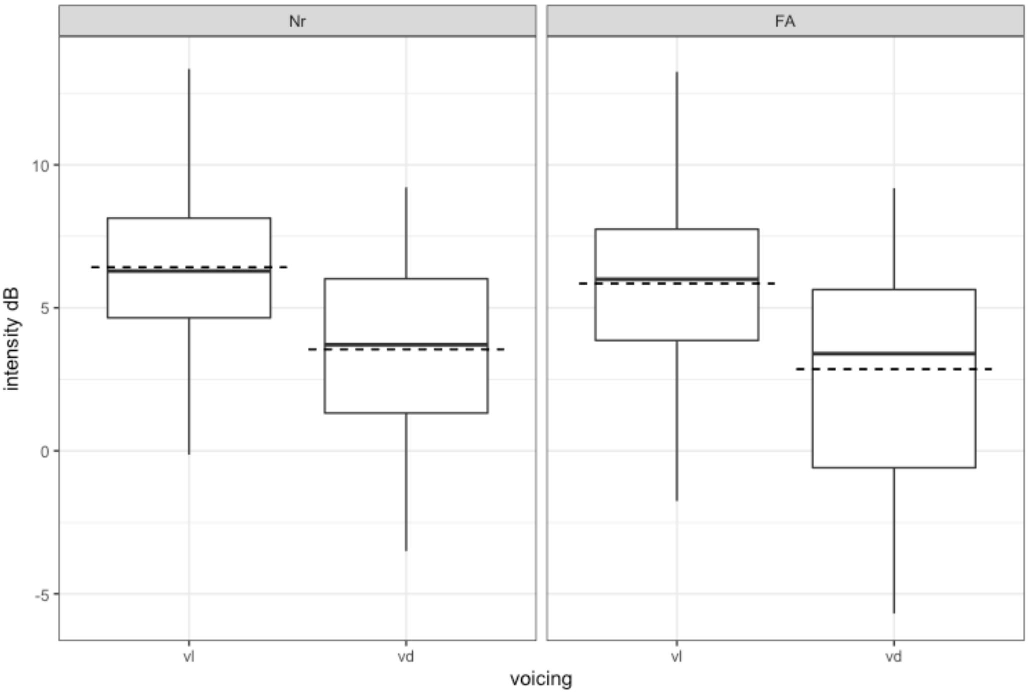


Figure 7.35. Boxplots of the fitted values of H1-H2 offset for Voiceless and Voiced utterance-final stops grouped by voicing and speech rate.

	Voiceless		**Voiced**	
	Mean	SD	Mean	SD
Normal	6.42	2.89	3.55	3
Fast	5.85	3.04	2.86	3.56

Table 7.35. Mean and standard deviation of the fitted values of H1-H2 offset for Voiceless and Voiced utterance-final stops grouped by voicing and speech rate.

Figure and table 7.35 show H1-H2 offset for Voiceless and Voiced utterance-final stops as a function of voicing and rate. The results show that H1-H2 offset for the Voiceless was significantly higher than for the Voiced stops by an average of 2.9 dB across speech rates ($p<0.0001$). With regard to speech rate effect, the results show a significant increase of H1-H2 offset in normal speech by an average of 0.6 dB for Voiceless ($p = 0.012$) and Voiced ($p = 0.0056$) stops.

H1-H2 offset as a function of voicing, rate, and place of articulation

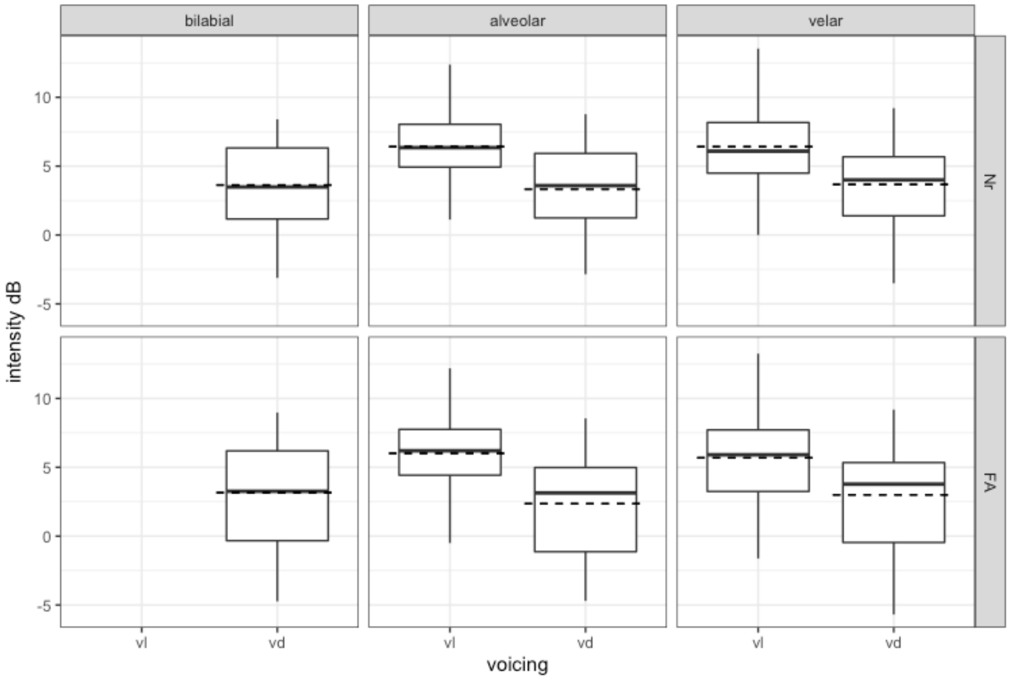


Figure 7.36. Boxplots of the fitted values of H1-H2 offset for Voiceless and Voiced utterance-final grouped by voicing, speech rate, and place of articulation.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Bilabial	-	-	3.63	3.05	-	-	3.16	3.54
Alveolar	6.42	2.69	3.33	2.97	6.01	2.86	2.36	3.55
Velar	6.42	3.08	3.67	2.97	5.69	3.2	3	3.56

Table 7.36. Mean and standard deviation of the fitted values of H1-H2 offset for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and place of articulation.

Figure and table 7.36 show H1-H2 offset for Voiced and Voiceless utterance-final stops as a function of voicing, rate, and place of articulation. With regard to voicing effect on H1-H2 offset, H1-H2 offset was significantly higher in the Voiceless context than in the Voiced by an average of 3 dB across speech rates (pairwise comparison: Appendix C Table C-169). With regard to speech rate effect, H1-H2 offset was marginally higher in normal speech than in fast speech across places of articulation and voicing categories (pairwise comparison: Appendix C Table C-169). With respect to the impact of place of articulation on H1-H2 offset, all the differences were found to be not significant across speech rate categories (pairwise comparison: Appendix C Table C-169).

7.7.3 Summary of results for H1-H2 offset

It has been presented that **H1-H2 offset** is a clear acoustic correlate that accompanies the distinction between Voiceless and Voiced stops in Najdi Arabic. This distinction is maintained across positions, rates, places of articulation, vowel types, and genders. The results of H1-H2 offset show a positional variation between utterance-medial and utterance-final stops. The former show relatively higher values than the latter. It could be noticed that H1-H2 offset in the context of high vowels is lower than the rest of the vowels across positions.

7.8 Conclusion

The following table summarises the acoustic measurements of the spectral correlates presented in this chapter under normal/fast speech rates and in every context. The correlates include burst intensity, F0 onset, F0 offset, F1 onset, F1 offset, H1-H2 onset, and H1-H2 offset.

		Utterance-initial			**Utterance-medial (trochaic)**			**Utterance-medial (iambic)**			**Utterance-final**		
Correlate	**Voicing**	**rate**	**M**	**SD**	**rate**	**M**	**SD**	**rate**	**M**	**SD**	**rate**	**M**	**SD**
Burst intensity	Voiceless	Nr	46.6	3.53	Nr	46.9	0.79	Nr	46.3	4.11	Nr	42.8	4.13
	Voiced	Nr	54.9	5	Nr	55.4	0.60	Nr	54.5	5.27	Nr	44.1	4.78
F0 onset	Voiceless	Nr	211	45.5	Nr	205	44.1	Nr	216	44.3	Nr	…..	…..
	Voiced	Nr	187	41.5	Nr	190	41.9	Nr	187	41.6	Nr	…..	…..
F0 offset	Voiceless	Nr	…..	…..	Nr	187	45.7	Nr	169	40.4	Nr	187	60.3
	Voiced	Nr	…..	…..	Nr	184	42.3	Nr	171	38.7	Nr	186	55.3
F1 onset	Voiceless	Nr	419	85.2	Nr	…..	…..	Nr	400	88.1	Nr	…..	…..
	Voiced	Nr	400	66.7	Nr	…..	…..	Nr	380	67.3	Nr	…..	…..
F1 offset	Voiceless	Nr	…..	…..	Nr	423	88.8	Nr	…..	…..	Nr	414	77.5
	Voiced	Nr	…..	…..	Nr	400	81.7	Nr	…..	…..	Nr	420	82.1
H1-H2 onset	Voiceless	Nr	5.68	3.84	Nr	…..	…..	Nr	4.72	3.98	Nr	…..	…..
	Voiced	Nr	4.48	2.31	Nr	…..	…..	Nr	4.75	2.27	Nr	…..	…..
H1-H2 offset	Voiceless	Nr	…..	…..	Nr	7.77	3.05	Nr	…..	…..	Nr	6.42	2.89
	Voiced	Nr	…..	…..	Nr	4.51	2.5	Nr	…..	…..	Nr	3.55	3
Burst intensity	Voiceless	FA	47.9	3.68	FA	48.3	1.07	FA	47.6	4.14	FA	44.1	3.69
	Voiced	FA	56.3	4.85	FA	56.4	0.86	FA	55.2	4.97	FA	45.5	4.56
F0 onset	Voiceless	FA	217	43	FA	210	43.1	FA	218	190	FA	…..	…..
	Voiced	FA	191	40	FA	193	40.9	FA	41.7	39.4	FA	…..	…..
F0 offset	Voiceless	FA	…..	…..	FA	189	43	FA	172	39.2	FA	181	53.5
	Voiced	FA	…..	…..	FA	187	40.2	FA	175	37.9	FA	178	50.2
F1 onset	Voiceless	FA	431	85.1	FA	…..	…..	FA	408	87.2	FA	…..	…..
	Voiced	FA	418	66.5	FA	…..	…..	FA	387	66.3	FA	…..	…..
F1 offset	Voiceless	FA	…..	…..	FA	436	85.7	FA	…..	…..	FA	415	75
	Voiced	FA	…..	…..	FA	413	79.6	FA	…..	…..	FA	424	81.1
H1-H2 onset	Voiceless	FA	5.09	3.83	FA	…..	…..	FA	4.35	4.01	FA	…..	…..
	Voiced	FA	3.92	2.37	FA	…..	…..	FA	4.36	2.36	FA	…..	…..
H1-H2 offset	Voiceless	FA	…..	…..	FA	7.36	3.16	FA	…..	…..	FA	5.85	3.04
	Voiced	FA	…..	…..	FA	4.18	2.7	FA	…..	…..	FA	2.86	3.56

Table 7.37. Overview of the acoustic measures (Mean and standard deviation) of the spectral correlates in Voiced and Voiceless stops under normal/fast speech rates.

The analysis for the spectral correlates shows crucial findings with regard to the distinction between Voiceless and Voiced stops in Najdi Arabic. F0 and F1 results appeared to be analogous to what has been found in both aspirating and voicing languages, i.e. the lowering of F0 and F1 in the context of Voiced stops and the raising of these correlates in the context of Voiceless stops. The distinction in terms of F0 onset is the most significant among the correlates in Voiceless and Voiced stops across all the factors. In terms of H1-H2, H1-H2 onset is a robust correlate that differentiates Voiceless and Voiced stops in utterance-initial position. H1-H2 offset, on the other hand, shows robust distinction across all the examined contexts.

The results for burst intensity show a pattern that might be unique to Arabic dialects in that burst intensity for Voiced is higher than for Voiceless stops. This is similar to what has been found in Lebanese Arabic (Al-Tamimi and Khattab, 2018).

The goal of investigating the spectral correlates is to investigate to what extent do these correlates strengthen the argument for the presence of [spread glottis] and [voice] in the voicing system for stops in Najdi Arabic. The findings form a crucial foundation for the phonological assumptions made with regard to the features that specify the voicing contrast in Najdi Arabic. The significant difference in F0 onset and H1-H2 onset/offset values indicates the presence of [spread glottis] for Voiceless stops and [voice] for Voiced stops.

Chapter 8. Regressive voicing assimilation in stop-stop cluster

The aim of this chapter is to closely examine the phonetic aspects of regressive voicing assimilation in stop-stop clusters across word boundaries in Najdi Arabic. In Chapter 2, it was established that voicing assimilation is a phonological process common to many aspirating and voicing languages. The traditional view of the nature of voicing assimilation posits that assimilated consonants (C1) are completely neutralised to the voicing status of the adjacent consonant (C2), thereby implying that the voicing targets in C1 are equivalent to those of C2 regardless of the underlying voicing of C1 (Chomsky and Halle, 1968). Many quantitative experimental studies, however, reveal results indicating incomplete neutralisation (Burton and Robblee, 1997; Barry and Teifour, 1999; Jansen, 2004; Kulikov, 2012). That is, some voicing targets of C1 preserve their underlying voicing. It has also been proposed that the ability of a stop to trigger voicing in the preceding stop relies on its phonological specification. In other words, actively (de)voiced stops (C2) are expected to trigger (de)voicing in the preceding stops (C1) (Jansen, 2004, 2007). Therefore, it is argued that Voiced stops in aspirating languages, as well as Voiceless stops in voicing languages, are not expected to trigger voicing in the preceding stops. In contrast, Voiceless stops in aspirating languages and Voiced stops in voicing languages are expected to trigger voicing in the preceding stop. On the other hand, the effect of C1 voicing on C2 has been described as a form of passive devoicing.

Based on the aforementioned theoretical assumptions, the following predictions can be made with regard to assimilation in stop-stop clusters in Najdi Arabic: 1) since both Voiceless and Voiced stops are proposed to be phonologically specified in Najdi Arabic, they should both trigger some form of (de)voicing in C1; 2) the assimilation process is expected to lead to incomplete neutralisation; 3) voiceless stops are expected to show less variation in their voicing targets than voiced stops for purely phonetic reasons.

The clusters included in the analysis are [bk], [kb], [dg], [tg], [gt]. It can be noticed that they differ in terms of voicing and place of articulation. They enable us to investigate the patterns required to decide whether Voiceless and Voiced stops trigger some (de)voicing in the preceding member of the cluster. The inclusion of [Voiced-Voiced] cluster [dg] is based on the assumption that C2 is expected to trigger some voicing and strengthen the voicing targets in C1 similar to what has been found in the previous literature (Barry and Teifour, 1999; (Jansen, 2004, 2007).

The results of analysis of stop-stop clusters are divided into three subsections: 1) general results: presenting examples of spectrograms and waveforms for each cluster with the percentages of voicing in C1 and C2 across tokens using the coding system showed earlier in table 5.1; 2) the acoustic features of C1: showing the statistical analysis (LMM) and pairwise comparisons for voicing targets in C1 as compared to C1 in the baseline environment in each cluster; 3) the acoustic features of C2: showing the statistical analysis (LMM) and pairwise comparisons for voicing targets in C2 compared to C2 in the baseline environment in each cluster.

8.1 General results

This section presents a set of examples of spectrograms and waveforms and the percentages of voicing in the first and second stops for each cluster. The coding system used in section 5.1 is also employed in this section to characterise the voicing patterns in all the tokens to conceptualize the differences between Voiced and Voiceless stops in terms of the degree of voicing in their closure phase.

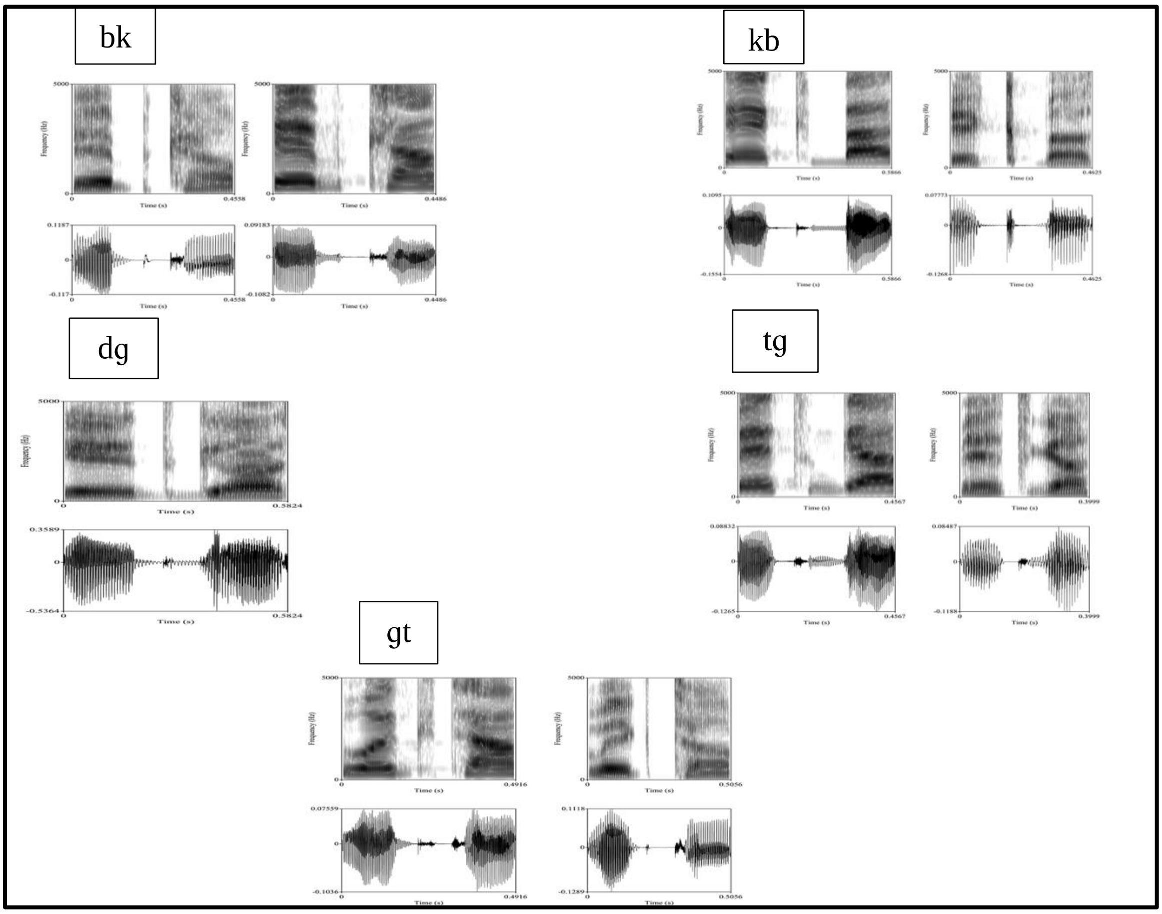


Figure 8.1. Examples of Spectrograms and waveforms for each of the clusters in the first and second stops.

	bk				kb				dg				tg				gt			
	C1		C2		C1		C2		C1		C2		C1		C2		C1		C2	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
U	11	11.8	93	100	87	100	15	17.2	1	1.2	7	8.5	87	98.8	26	29.5	9	10.9	82	100
UP	30	32.2	0	0	0	0	6	6.8	8	9.7	1	1.2	1	1.2	2	2.2	28	34.1	0	0
V	18	19.3	0	0	0	0	45	51.7	67	81.7	73	89	0	0	49	55.6	26	31.7	0	0
V80	3	3.2	0	0	0	0	8	9.1	0	0	0	0	0	0	2	2.2	4	4.8	0	0
VP	31	33.3	0	0	0	0	13	14.9	6	7.3	0	0	0	0	7	7.9	19	23.1	0	0
Total	93		93		87		87		82		82		88		88		82		82	

Table 8.1. The percentage of voicing in the closure for each cluster in the first and second stops. N refers to number of tokens.

Figure and table (8.1) mirror the voicing patterns of each cluster examined in the data. It can be noted that Voiceless stops in all clusters retained their voicelessness in both C1 and C2 positions. Voiced stops, on the other hand, showed variability in the amount of voicing in their closure. The results showed that Voiced stops in C1 tended to retain their voicing status within the voicing range (VP, V80, V) in 55.5% of the tokens of [bk] and in 59.2% of the tokens in [gt]. In [dg], the results showed that C1 were within the voicing range in 89% of the tokens. In C2 position, the results showed that Voiced stops tended to retain their voicing status within the voicing range in 75% of the tokens in [kb], in 89% in [dg], and 65% in [tg].

8.2 The acoustic features of C1 and C2 compared to their baseline environments

As established in the literature review, voicing assimilation in stop-stop clusters is expected to occur in both aspirating and voicing languages. The aim of this section is to present the acoustic features of the voicing targets of C1 and C2 in order to precisely detect any changes caused by the voicing assimilation process. To do so, the acoustic features of C1 in the cluster were compared to the acoustic features of C1 in the baseline context. Additionally, the acoustic features of C2 in the cluster were compared to the acoustic features of C2 in the baseline context. The baseline context for C1 is the same stop in utterance-final position whereas the baseline context for C2 is the same stop in utterance-initial position. Taking into consideration that some durational correlates of stops in utterance-final position are expected to be lengthened, only voicing duration, F0, F1, and burst intensity were used in the comparison. The LMM model predictions and pairwise comparisons are presented in the forthcoming sections.

8.2.1The acoustic features of C1 vs C1 baseline

Voicing duration C1 vs C1 baseline

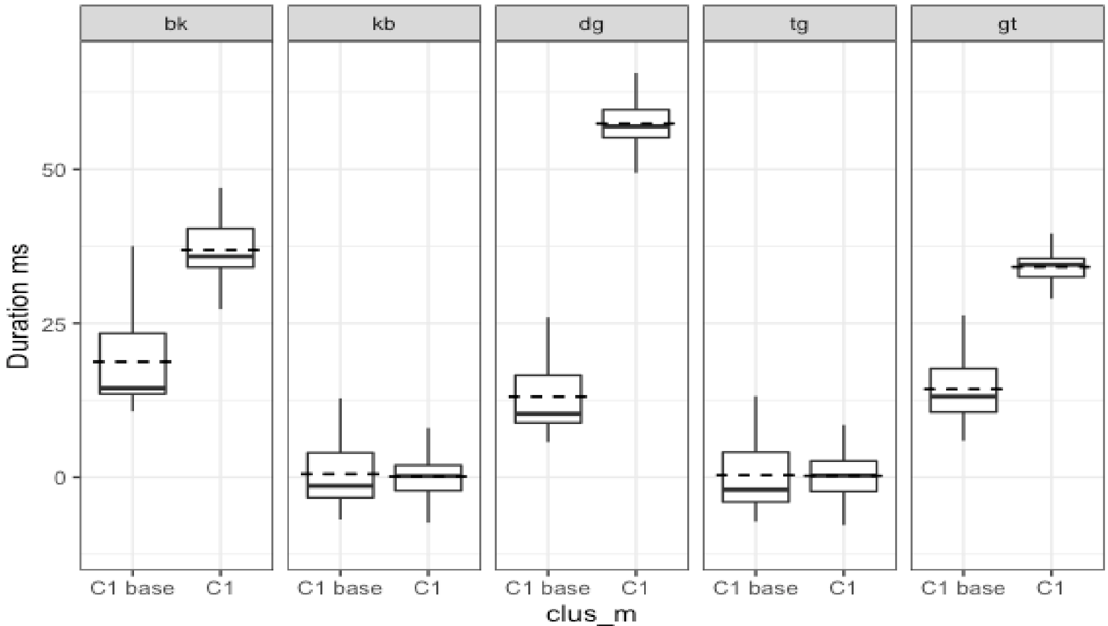


Figure 8.2. Boxplots of voicing duration for C1 and C1 baseline classified by cluster (C1: stop in cluster; C1 base: stop in baseline context). The dashed line represents the mean.

	C1		**C1 baseline**	
	Mean	SD	Mean	SD
[bk]	36.9	5.03	18.8	7.61
[kb]	0.119	4.69	0.53	6.27
[dg]	57.4	4.81	13.1	6.61
[tg]	0.24	4.75	0.35	6.64
[gt]	34.1	4.95	14.3	6.27

Table 8.2. Means and standard deviations of voicing duration for C1 and C1 baseline grouped by cluster.

Figure and table (8.2.) show voicing duration values for C1 and C1 baseline context for each cluster. With regard to [bk], voicing duration for [b] in the cluster was significantly longer than for [b] in the baseline context with an average of 18.1 ms ($p<0.0001$). For [kb], voicing duration for [k] in the cluster was marginally shorter than for [k] in the baseline context by an average of 0.42 ms ($p=0.71$). Moving to [dg], voicing duration for [d] in the cluster was significantly longer than in the baseline context by an average of 44.3 ms ($p<0.0001$). For [tg], voicing duration for [t] was marginally shorter in the cluster than in the baseline context by 0.11 ms ($p=0.9$). With regard to [gt], voicing duration for [g] was significantly longer in the cluster than in the baseline context by an average of 19.8 ms ($p<0.0001$).

F0 offset for C1 vs C1 baseline

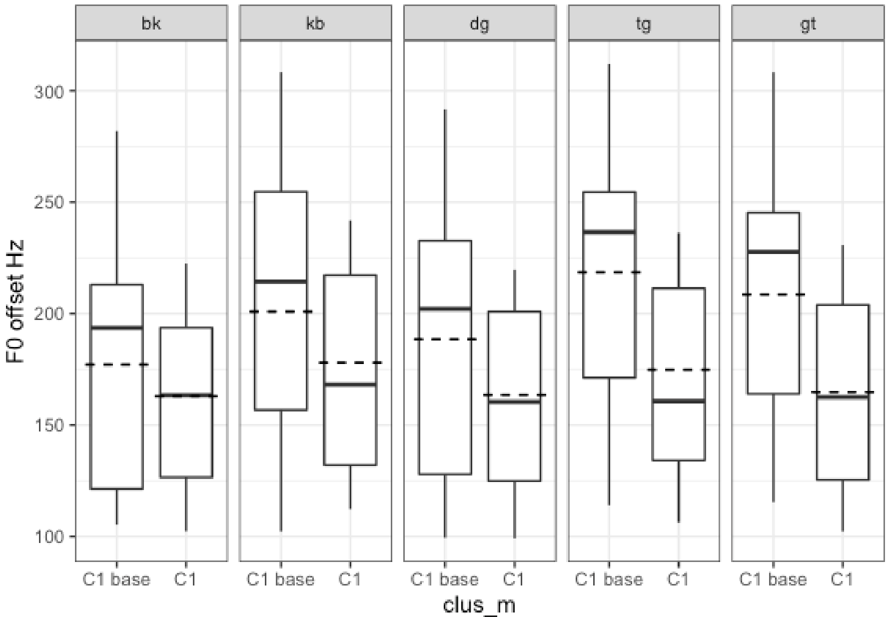


Figure 8.3. Boxplots of F0 offset for C1 and C1 baseline classified by cluster (C1: stop in cluster; C1 base: stop in baseline context). The dashed line represents the mean.

	C1		**C1 baseline**	
	Mean	SD	Mean	SD
[bk]	163	38.4	177	48.4
[kb]	178	40.9	201	65.4
[dg]	163	40.3	188	58.5
[tg]	175	43.2	218	55.2
[gt]	165	42.9	209	49.9

Table 8.3. Means and standard deviations of F0 offset for C1 and C1 baseline grouped by cluster.

Figure and table 8.3 show F0 offset values for C1 and C1 baseline context for each cluster. With regard to [bk], F0 offset for [b] in the cluster was marginally lower than for [b] in the baseline context with an average of 14 Hz (p=0.22). For [kb], F0 offset for [k] in the cluster was marginally lower than for [k] in the baseline context by an average of 23 Hz (p=0.06). Moving to [dg], F0 offset for [d] in the cluster was significantly lower than in the baseline context by an average of 25 Hz (p=0.02). For [tg], F0 offset for [t] was significantly lower in the cluster than in the baseline context by 43 Hz (p<0.0001). With regard to [gt], F0 offset for [g] was significantly lower in the cluster than in the baseline context by an average of 44 Hz (p<0.0001).

F1 offset for C1 vs C1 baseline

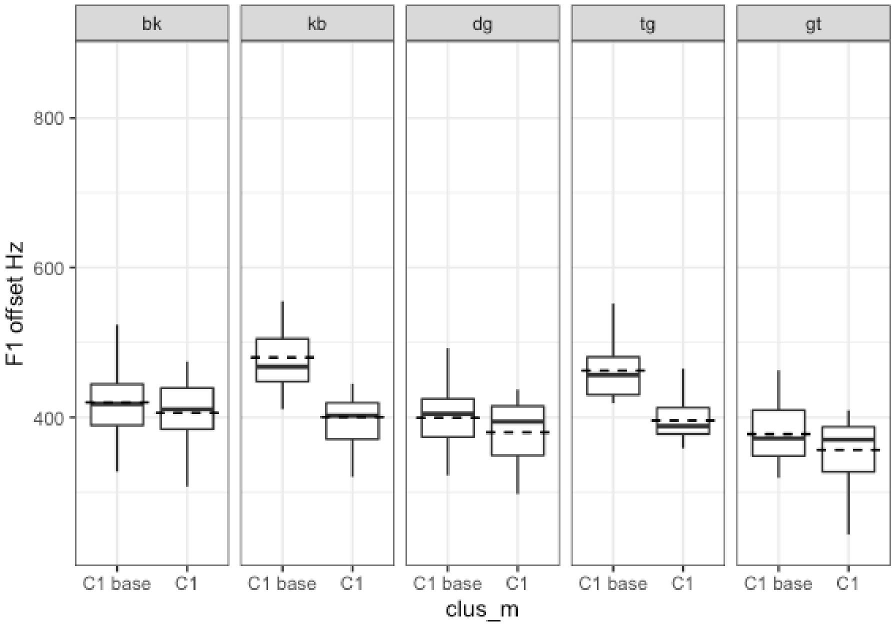


Figure 8.4. Boxplots of F1 offset for C1 and C1 baseline classified by voicing on the right side (vl = voiceless, vd = voiced) and cluster (C1: stop in cluster; C1 base: stop in baseline context). The dashed line represents the mean.

	C1		**C1 baseline**	
	Mean	SD	Mean	SD
[bk]	405	51.6	420	42.5
[kb]	401	61.8	480	84.6
[dg]	380	42.9	399	35.1
[tg]	396	25.1	462	39
[gt]	356	41.4	377	35.1

Table 8.4. Means and standard deviations of F1 offset for C1 and C1 baseline grouped by cluster.

Figure and table 8.4 present F1 offset values for C1 and C1 baseline context for each cluster. With regard to [bk], F1 offset for [b] in the cluster tended to be lower than for [b] in the baseline context with an average of 15 Hz (p=0.075). For [kb], F1 offset for [k] in the cluster was significantly lower than for [k] in the baseline context by an average of 79 Hz (p<0.0001). Moving to [dg], F1 offset for [d] in the cluster was significantly lower than in the baseline context by an average of 19 Hz (p=0.011). For [tg], F1 offset for [t] was significantly lower in the cluster than in the baseline context by 66 Hz (p<0.0001). With regard to [gt], F1 offset for [g] was significantly lower in the cluster than in the baseline context by an average of 21 Hz (p=0.0055).

Burst intensity for C1 vs C1 baseline

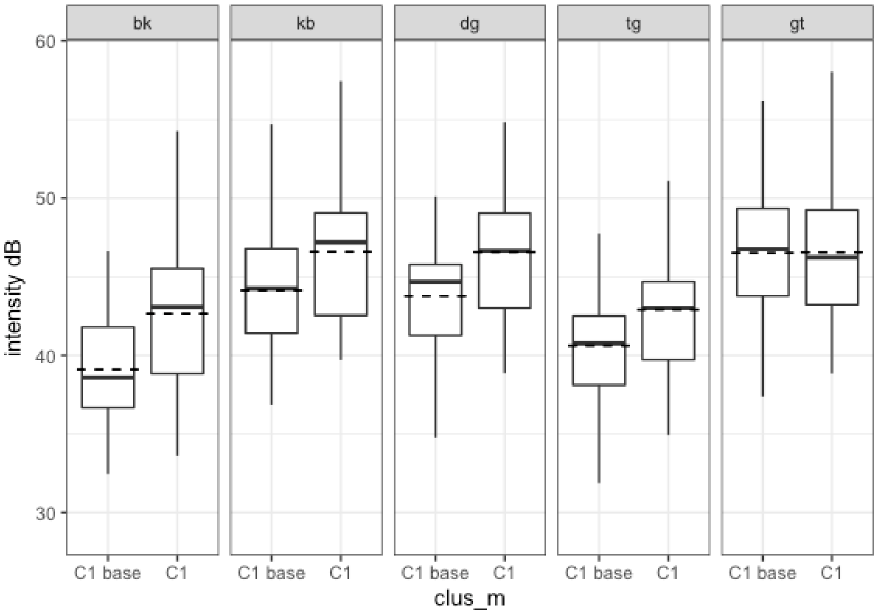


Figure 8.5. Boxplots of the fitted values of Burst intensity for C1 and C1 baseline classified by voicing on the right side (vl = voiceless, vd = voiced) and cluster (C1: stop in cluster; C1 base: stop in baseline context). The dashed line represents the mean.

	C1		**C1 baseline**	
	Mean	SD	Mean	SD
[bk]	42.6	5.43	39.1	4.16
[kb]	46.6	4.16	44.1	3.88
[dg]	46.6	4.89	43.8	3.67
[tg]	42.9	4.51	40.6	3.69
[gt]	46.6	4.32	46.5	3.67

Table 8.5. Mean and standard deviation of the fitted values of burst intensity for C1 and C1 baseline grouped by voicing and cluster.

Figure and table 8.5 show burst intensity values for C1 and C1 baseline context for each cluster. In terms of [bk], the results showed a marginal increase of burst intensity in the cluster context by an average of 3.5 dB ($p=0.073$). For [kb], burst intensity for [k] in the cluster was significantly higher than for [k] in the baseline context by an average of 2.5 dB ($p<0.0001$). Moving to [dg], burst intensity for [d] in the cluster was significantly higher than in the baseline context by an average of 2.8 dB ($p<0.0001$). For [tg], burst intensity for [t] was significantly higher in the cluster than in the baseline context by 2.3 dB ($p<0.0001$). With regard to [gt], burst intensity for [g] was marginally higher in the cluster than in the baseline context by an average of 0.1 dB ($p=0.33$).

8.2.2 The acoustic features of C2 vs C2 baseline

Voicing duration C2 vs C2 baseline

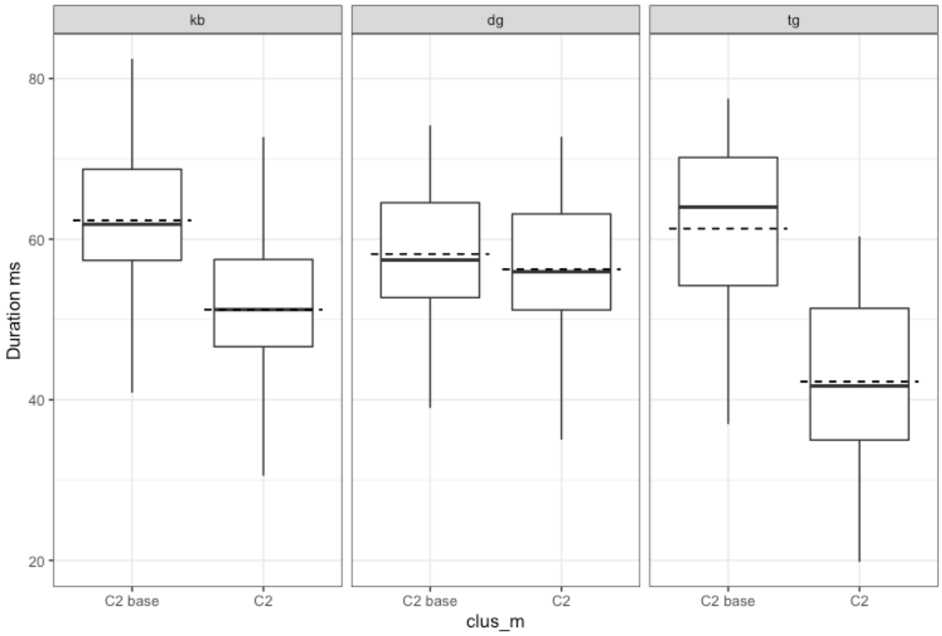


Figure 8.6. Boxplots of the fitted values of voicing duration for C2 and C2 baseline classified by cluster (C2: stop in cluster; C2 base: stop in baseline context). The dashed line represents the mean.

	C2		**C2 baseline**	
	Mean	SD	Mean	SD
[kb]	51.2	9.31	62.3	8.76
[dg]	56.2	8.7	58.2	8.23
[tg]	42.3	10.4	61.3	9.93

Table 8.6. Mean and standard deviation of the fitted values of voicing duration for C2 and C2 baseline grouped by cluster.

Figure and table 8.6 show voicing duration values for C2 and C2 baseline context for each cluster that has C2 as voiced stops. Clusters that have Voiceless C2 were excluded because there was no occurrence of voicing in any token in either the cluster or baseline contexts. With regard to [kb], voicing duration for [b] in the cluster was marginally shorter than for [b] in the baseline context with an average of 11.1 ms (p=0.22). For [dg], voicing duration for [g] in the cluster was marginally shorter than for [g] in the baseline context by an average of 2 ms (p=0.06). With regard to [tg], voicing duration for [g] was significantly shorter in the cluster than in the baseline context by an average of 19 ms (p<0.0001).

F0 onset for C2 vs C2 baseline

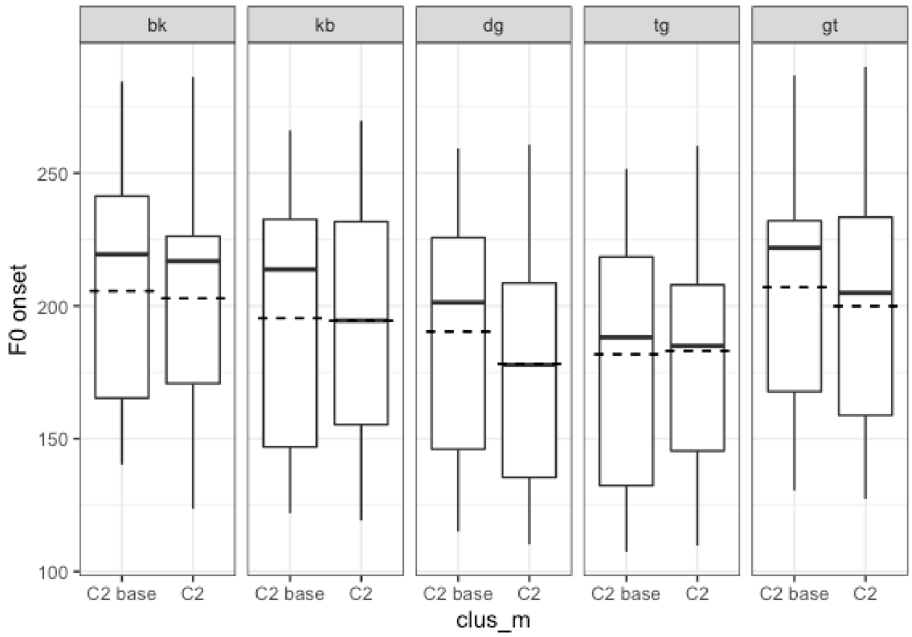


Figure 8.7. Boxplots of the fitted values of F0 onset for C2 and C2 baseline classified by cluster (C2: stop in cluster; C2 base: stop in baseline context). The dashed line represents the mean.

	C2		**C2 baseline**	
	Mean	SD	Mean	SD
[bk]	203	40.4	206	39.1
[kb]	194	42.7	195	43.9
[dg]	178	41.8	190	41.9
[tg]	183	40.2	182	43.4
[gt]	200	43.7	207	39.8

Table 8.7. Mean and standard deviation of the fitted values of F0 onset for C2 and C2 baseline grouped by cluster.

Figure and table 8.7 show F0 onset values following C2 and C2 baseline context for each cluster. In terms of [bk], the results showed a marginal decrease of F0 onset for [k] in the cluster context by an average of 3 Hz (p=0.8). For [kb], F0 onset for [b] was marginally lower in the cluster than in the baseline context by an average of 1 Hz (p=.89). Moving to [dg], F0 onset for [g] in the cluster was significantly lower than in the baseline context by an average of 12 Hz (p=0.26). For [tg], F0 onset for [g] was marginally higher in the cluster than in the baseline context by 1 Hz (p=0.89). With regard to [gt], F0 onset for [t] was marginally lower in the cluster than in the baseline context by an average of 7 Hz (p=0.54).

F1 onset for C2 vs C2 baseline

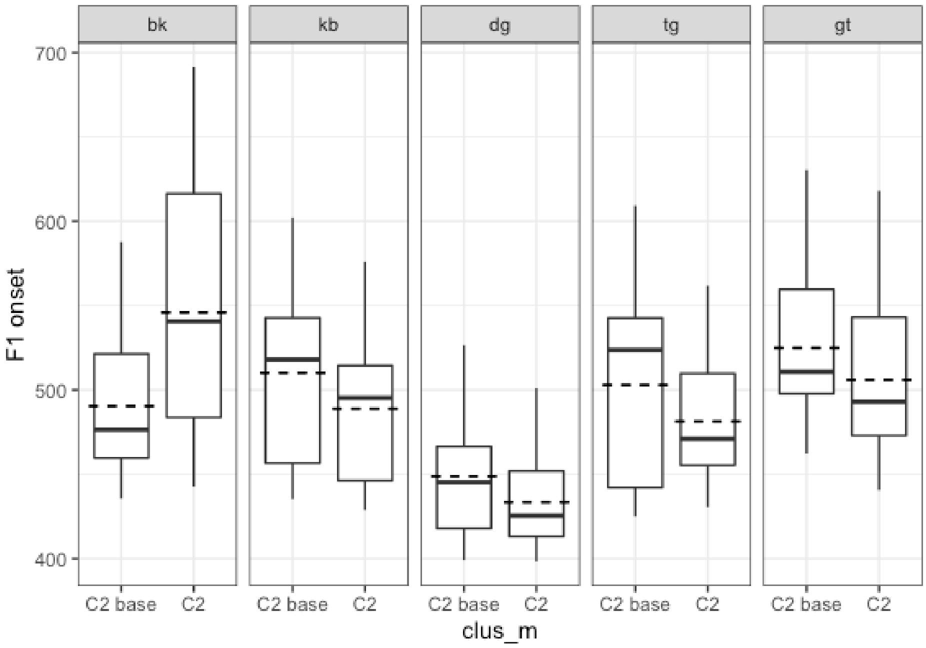


Figure 8.8. Boxplots of the fitted values of F1 onset for C2 and C2 baseline classified by cluster (C2: stop in cluster; C2 base: stop in baseline context). The dashed line represents the mean.

	C2		**C2 baseline**	
	Mean	SD	Mean	SD
[bk]	546	71.2	490	40.1
[kb]	489	42.1	510	50.5
[dg]	433	30.8	499	35.7
[tg]	481	34.5	503	56.3
[gt]	506	45.6	525	43.1

Table 8.8. Mean and standard deviation of the fitted values of F1 onset for C2 and C2 baseline grouped by cluster.

Figure and table 8.8 present F1 onset values following C2 and C2 baseline context for each cluster. In terms of [bk], F1 onset for [k] was significantly higher in the cluster context than in the baseline context by an average of 56 Hz (p<0.0001). For [kb], F1 onset for [b] in the cluster was significantly lower than in the baseline context by an average of 21 Hz (p=0.012). Moving to [dg], F1 onset for [g] in the cluster was significantly lower than in the baseline context by an average of 66 Hz (p<0.0001). For [tg], F1 onset for [g] was significantly lower in the cluster than in the baseline context by 22 Hz (p=0.006). With regard to [gt], F1 onset for [t] was significantly lower in the cluster than in the baseline context by an average of 19 Hz (p=0.01).

Burst intensity for C2 vs C2 baseline

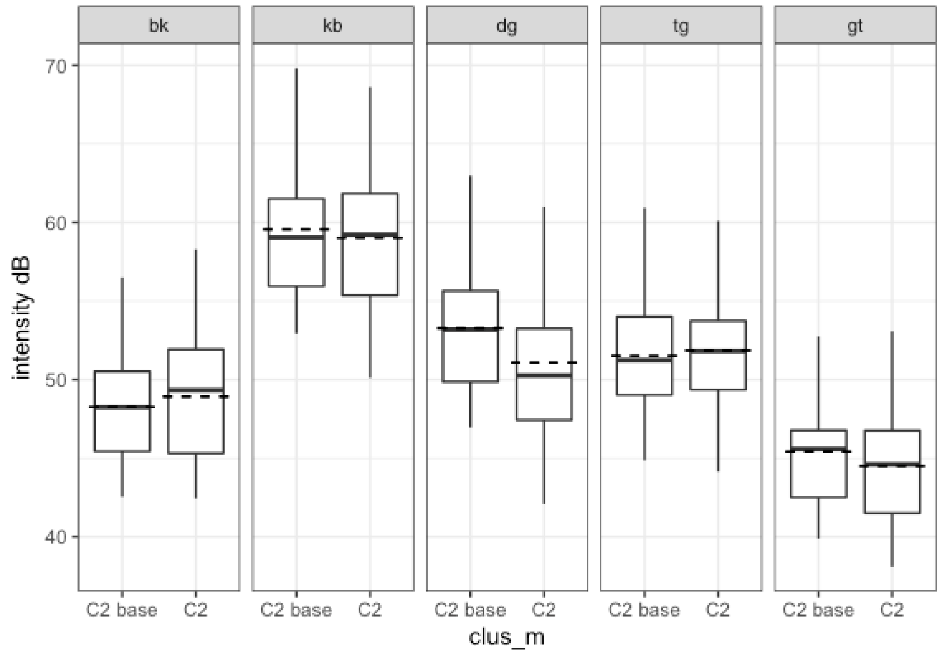


Figure 8.9. Boxplots of the fitted values of burst intensity for C2 and C2 baseline classified by cluster (C2: stop in cluster; C2 base: stop in baseline context). The dashed line represents the mean.

	C1		**C1 baseline**	
	Mean	SD	Mean	SD
[bk]	48.9	3.94	48.2	3.73
[kb]	59	5.23	59.6	4.34
[dg]	51.1	5.21	53.3	4
[tg]	51.9	4.04	51.5	3.93
[gt]	44.5	3.71	45.4	3.39

Table 8.9. Mean and standard deviation of the fitted values of burst intensity for C2 and C2 baseline grouped by cluster.

Figure and table 8.9 present burst intensity values for C2 and C2 baseline context for each cluster. In terms of [bk], burst intensity for [k] was marginally higher in the cluster context than in the baseline context by an average of 0.7 dB (p=0.28). For [kb], burst intensity for [b] in the cluster was marginally lower than in the baseline context by an average of 0.6 dB (p=0.42). Moving to [dg], burst intensity for [g] in the cluster was significantly lower than in the baseline context by an average of 2.2 dB (p=0.001). For [tg], burst intensity for [g] was marginally higher in the cluster than in the baseline context by 0.4 dB (p=0.6). Regarding [gt], burst intensity for [t] was marginally lower in the cluster than in the baseline context by an average of 0.9 dB (p=0.16).

8.3 Conclusion

Chapter 8 focuses on the acoustic correlates of C1 and C2 in stop-stop clusters across word boundaries in Najdi Arabic. The acoustic features of C1 were compared to C1 in the baseline context whereas the acoustic features of C2 were compared to C2 in the baseline context. The aim was to detect any changes in these features as a result of voicing assimilation. The acoustic features that were examined include voicing duration in the hold phase, F0 offset, F0 onset, F1 offset, F1 onset, and burst intensity.

Starting with the acoustic correlates of C1, Voiced stops in [voiced-voiceless] and [voiced-voiced] patterns showed a significant increase in voicing duration in the cluster context. On the other hand, Voiceless stops in [voiceless-voiced] pattern showed no difference between the cluster and the baseline contexts in their voicing duration. In terms of F0/F1, Voiced stops in [voiced-voiceless] and [voiced-voiced] patterns showed a significant lowering of F0/F1 offset in the cluster context. Voiceless stops in [voiceless-voiced] pattern, however, also showed a lowering of F0/F1 offset in the cluster context. In terms of burst intensity, Voiced stops in [voiced-voiceless] and [voiced-voiced] patterns showed a significant increase in burst intensity in the cluster context. Similarly, Voiceless stops in [voiceless-voiced] pattern showed a significant increase in the cluster context.

Moving to the acoustic correlates of C2, Voiced stops in [voiceless-voiced] and [voiced-voiced] patterns showed a tendency for shorter voicing duration in the cluster context. Voiceless stops showed no voicing in the hold phase in both the baseline and the cluster contexts. With regard to F0 onset, Voiced stops in [voiceless-voiced] pattern showed a tendency for lower F0 onset in the cluster context, but a significant decrease in [voiced-voiced] pattern. Similarly, Voiceless stops in [voiced-voiceless] pattern showed a tendency for lower F0 onset in the cluster context. For F1onset, Voiced stops in [voiceless-voiced] and [voiced-voiced] patterns showed a significant decrease in F1 onset in the cluster context. Voiceless stops in [voiced-voiceless] pattern, however, showed ambiguous results in which F1 onset for [k] showed higher values in the cluster context whereas [t] showed the opposite in the same context. Burst intensity results showed no significant differences in both Voiced and Voiceless stops.

The results showed that voicing assimilation in Najdi Arabic was incomplete in which the assimilated stops in C1 preserved some of their voicing targets. This is in agreement with what was found in the studies that implemented a quantitative approach in their investigation of voicing assimilation in both aspirating and voicing languages (Burton and Robblee, 1997; Barry and Teifour, 1999; Jansen, 2004; Kulikov, 2012). In terms of the participation of Voiceless and Voiced stops in the assimilation process in Najdi Arabic, the results showed that both Voiceless and Voiced stops triggered some degree of (de)voicing in the preceding stop. This finding provides a support for Jansen's claim (2004) that specified stops are expected to trigger (de)voicing in the preceding stop. The resistance to change in Voiceless and Voiced stops in C1 was manifested differently; Voiceless stops retained voiceless closure while Voiced stops retained the lowering of F0/F1 and raising of burst intensity.

The results demonstrate that Voiceless and Voiced stops in Najdi Arabic differ in their participation in the assimilation process. That is, Voiceless stops in C2 clearly trigger devoicing to the hold phase of the Voiced stops in C1. On the contrary, Voiced stops in C2 do not trigger voicing in the hold phase of the preceding Voiceless stops in C1. Yet, if we look closely to the behaviour of Voiced stops in the assimilation process, it can be noticed that Voiced stops trigger regressive voicing in the preceding stops in [Voiced-Voiced] clusters by increasing the voicing in the closure. They also trigger voicing to the preceding stops in [Voiceless-Voiced] clusters in F0, F1, and burst intensity but not in the voicing in the closure. These findings support the main arguments of this study which emphasise that both Voiceless and Voiced stops are specified and both trigger (de)voicing in the preceding stops. The change in F0/F1 and burst intensity as a result of voicing assimilation is important since they are considered in numerus studies among the acoustic correlates that actively differentiate between Voiceless and Voiced stops in both aspirating and voicing languages (Jansen 2004, 2008). Also, they are products of active articulatory gestures that might cause an anticipatory effect in the voicing targets in C1. Jansen (2004) argues that the correlates that are expected to spread from C2 to C1 are the ones that are produced with an active articulatory gesture.

Chapter 9. Discussion

The main goal of the present study, to establish a comprehensive understanding of the phonetic and phonological aspects of voicing contrast in Najdi Arabic, was achieved by investigating the acoustic correlates of Voiced and Voiceless stops in different phonetic contexts and drawing phonological conclusions related to the specification of stops and the activeness of distinctive features. This chapter will discuss the results in light of the theoretical frameworks and synthesize the similarities and differences observed between Najdi Arabic and the existing studies that describe voicing contrast in voicing and aspirating languages. This chapter is structured as follows: Section 9.1 focuses on the interaction between prevoicing and aspiration and linguistic factors, including place of articulation, vowel quality, phonetic context. Section 9.2 focuses on the phonological implications proposed in the realm of laryngeal realism and examines to what extent the predictions of the laryngeal realism account for the voicing contrast in Najdi Arabic, in addition to dissecting the parallelism between the types of evidence employed in laryngeal realism. Moreover, the categorization of languages that exhibit the features of both voicing and aspirating languages, in light of the acoustic features of voicing and aspirating languages, is discussed. Section 9.3 presents the limitations of the present study and suggestions for future research. Section 9.4 concluded the present work and gives some final remarks with respect to the laryngeal system of Najdi Arabic.

9.1 The interaction between prevoicing and aspiration and linguistic factors.

Starting with **place of articulation**, the results showed a significant effect of place of articulation on aspiration in utterance-initial and utterance-medial (iambic) stops in which the pattern appeared in the order velar > alveolar > bilabial. This effect was found in both Voiced and Voiceless stops (velar > alveolar for voiceless stops) and across speech rates, and this pattern aligns with the results found in aspirating languages, such as English (Caramazza et al., 1973; Klatt, 1975; Suomi, 1980; Docherty, 1992; Nearey and Rochet, 1994; Yao, 2009), and German (Jessen, 1998). It is also found in voicing languages such as Spanish (Rosner et al., 2000) and Portuguese (Lousada et al., 2010). Additionally, this phenomenon is found in Swedish which shows the features of both voicing and aspirating languages (Helgason and Ringen, 2008). These results are consistent with aerodynamic and articulatory explanations, i.e., that velars are produced with bigger size in the front cavity which leads to longer aspiration and more delay for the following vowel (Cho and Ladefoged, 1999). Regarding prevoicing, it is expected that prevoicing will be decreased as the place of articulation moves back in the mouth because the size of oral cavity in bilabial stops enhances the maintenance of air pressure difference required for voicing to be retained (Ohala, 1983; Keating, 1984). Prevoicing in Voiced stops in Najdi Arabic showed the same universal pattern across positions and speech rates in which that bilabial stops have longer prevoicing than alveolar and velar stops.

Moving on to **the quality of the following vowel**, the results of Najdi Arabic showed that aspiration in Voiceless stops is longer preceding high vowels /i:/ and /u:/ than the rest of the vowels in utterance-initial and utterance-medial (iambic) stops. These results are consistent with what has been reported in several studies (Docherty, 1992; Smith, 1987; Klatt, 1975; Jessen, 1998). With regard to the quality of the preceding vowel, the results showed variation in aspiration in utterance-medial (trochaic) and utterance-final stops. In terms of prevoicing, the results showed that prevoicing is longer before /e:/ and /o:/ in utterance-initial stops. In utterance-medial and utterance-final stops, prevoicing was not significantly affected by the quality of the preceding vowel. The results for aspiration seem consistent with the expectation that high vowels are associated with bigger size of oral cavity which causes a delay in the preparation for initiating the voicing onset of the following vowel and consequently means more time for aspiration (Smith, 1978; Ohala, 1983).

Regarding **syllable structure**, it has been proposed that prevoicing and aspiration are longer in monosyllabic words than in disyllabic words (Lisker and Abramson, 1967; Klatt, 1975; Flege, et al., 1998). However, this assumption cannot be tested in the present study because utterance-initial stops and utterance-final stops were tested in monosyllabic words while utterance-medial stops were tested in word-medial intervocalic disyllabic structure.

Moving on to **stress**, the results of aspiration and prevoicing agree with what has been expected, namely, that aspiration and prevoicing are longer in stressed syllables than in unstressed ones (Iverson and Salmons, 1995; Jacques, 1987; Kahn, 1976; Lavoie, 2001).

9.2 The phonological implication for the acoustic correlates of voicing.

Thus far we have investigated the acoustic correlates that account for the phonological opposition between Voiced and Voiceless stops in Najdi Arabic in utterance-initial, utterance-medial, and utterance-final positions while considering various linguistic factors. The acoustic results presented in the previous section prompt further thought about their phonological implications and demand a determination of which of these acoustic results are phonologically or phonetically motivated. As highlighted earlier, given the considerable number of experimental studies that have examined voicing contrast across languages, it has been proposed that Voiced and Voiceless stops in aspirating and voicing languages behave differently in each phonetic context. These studies have proven informative in terms of understanding the nature of the phonological representations that specify the voicing distinction and providing typological implications of their phonetic manifestations among the world's languages. The acoustic analysis of voicing contrast in Najdi Arabic presented in this study will contribute to accounting for the distinctive features that signal the contrast between Voiced and Voiceless stops. The phonological implications of the acoustic details will be presented in each context in the coming paragraphs considering several phonological approaches with a special focus on laryngeal realism.

9.2.1 VOT and feature specification.

In general, the results reveal a potential correlation between the phonological features that specify voicing contrast in Najdi Arabic and two possibilities with regard to the nature of the distinctive features: binary representational system and privative representational system. As mentioned earlier, some studies have proposed that voicing contrast is specified by the binary feature [±voice] in both aspirating and voicing languages (Keating, 1984; Kingston and Diehl, 1995; Wetzels and Mascaro, 2001) or [±spread glottis] (or [±tense]) for aspirating languages and [±voice] for voicing languages (Keating, 1990; Jessen, 2001). Other studies, adopting privative models of features, have demonstrated that [voice] specifies voicing in voicing languages and [spread glottis] specifies voicing in aspirating languages (Iverson and Salmons, 1995; Honeybone, 2005; Beckman et al., 2013; Jessen and Ringen, 2002).

Starting with **utterance-initial stops**, the results showed that Najdi Arabic contrasts prevoiced with aspirated stops. At normal speech rate, Voiced stops were prevoiced in 90.5% of the tokens by an average of -65.13 ms for the prevoicing duration. On the other hand, Voiceless stops were aspirated in 90% of the tokens by an average of 47.6 ms for the aspiration duration. In fast speech, Voiced stops were prevoiced in 86% of the tokens by an average of -40 ms for the prevoicing duration. On the other hand, Voiceless stops were aspirated in 56.4% of the tokens by an average of 41.2 ms for the aspiration duration.

Based on these findings, the binary system would conclude that Voiced stops in Najdi Arabic are specified with [+voice] while Voiceless stops are specified with [-voice] considering the traditional view based on SPE (Chomsky and Halle, 1968). Although the binary phonological specification seems straightforward and simple, it is not transparent in mapping from the distinctive features to the phonetic properties and consequently does not account for the phonetic features of voicing contrast in Najdi Arabic. Based on the definition of distinctive features adopted in the present work, a distinctive feature is associated with the phonetic property that carries the distinctiveness of the segment, which is aspiration in aspirating languages and active voicing in voicing languages (Jessen, 1998; Kingston and Diehl, 1995). The same issue arises when considering the studies that proposed [±spread glottis] to be the feature for aspirating languages and [±voice] for voicing languages and not accounting for languages with two-way contrast that show active voicing and aspiration in their laryngeal system. That is, the presence of prevoicing in utterance-initial position indicates an active articulatory and aerodynamic gesture in the production of voicing in an unfavourable context (Jessen, 2001; Jansen, 2004). Similarly, the presence of aspiration is a significant marker of an active glottal opening in the production of Voiceless stops. Accordingly, neither [±voice] or [±spread glottis] alone can accurately account for the voicing contrast in Najdi Arabic utterance-initial stops. Some studies have postulated that long aspiration in Voiceless stops in some voicing languages is nothing but a phonetic enhancement once used to strengthen the distinction between the two series of stops in the event voiced stops showed weak prevoicing (Cho and Keating, 1984; Fougeron, 2001). However, this is not the case in the results of the present study based on the aforementioned findings. This is because the enhancement is not needed in Najdi Arabic since the active voicing is expected to be more than enough to signal the opposition.

Moving on to the privative model in laryngeal realism, utterance-initial stops in Najdi Arabic posit a challenge to laryngeal realism also in that both prevoicing and aspiration are present in the voicing system. Laryngeal realism predicts a strong, transparent connection between the distinctive features and the acoustic details. Therefore, based on the assumption of laryngeal realism, voicing contrast in Najdi Arabic is overspecified with two distinctive features: [voice] and [spread glottis]. To test to what extent [voice] and [spread glottis] are active in the phonology of Najdi Arabic, both aspiration and prevoicing should be actively produced by the speaker (Beckman et al., 2013). Speech rate effects on prevoicing and aspiration are typically used in the realm of laryngeal realism to mirror the activeness of the features. That is, speech rate is expected to affect the specified stops but not the unspecified ones. This finding was confirmed in many studies that examine both voicing and aspirating languages. The results of speech rate effect on prevoicing and aspiration in Najdi Arabic showed that normal speech rate resulted in significantly longer prevoicing and aspiration in both Voiced and Voiceless utterance-initial stops, respectively. Taking into account the articulatory justification for speech rate effect on the specified stops, it can be proposed that when the speaker speaks slowly to make speech clearer, the acoustic outputs which are produced with active gestures will have a greater chance to increase their amount in comparison to spontaneous acoustic signals (Beckman et al., 2011). Therefore, it can be safely proposed that prevoicing and aspiration in Najdi Arabic are products of active gestures intended by the speaker.

Moving on to **utterance-medial stops (iambic)**, the results revealed that Najdi Arabic contrasts prevoiced with aspirated stops. At normal speech rate, Voiced stops were produced with voicing covering all the closure in 98% of the tokens with an average of 51.3 ms prevoicing duration. Voiceless stops, on the other hand, were aspirated in 70.8% of the tokens with an average of 42.1 ms aspiration duration. Voiceless stops showed no voicing in the closure in 98% of tokens. In fast speech, Voiced stops were produced with voicing covering all the closure in 98% of the tokens with an average of 42.4 ms prevoicing duration. Conversely, Voiceless stops were aspirated in 42.4% of the tokens with an average of 36.8 ms for aspiration duration and showed no voicing in the closure in 94% of the tokens. For **utterance-medial stops (trochaic)**, at normal speech rate, Voiced stops were produced with voicing covering all the closure in 98% of the tokens with an average of 49.5 ms for prevoicing duration. Voiceless stops, on the other hand, were aspirated 42.9% of the tokens with an average of 32.4 ms for aspiration duration and showed no voicing in the closure in 93% of the tokens. In fast speech, Voiced stops were produced with voicing covering all the closure in 97% of the tokens with an average of 44.2 ms for prevoicing duration. In contrast, Voiceless stops were aspirated in 15.9% of the tokens by an average of 28.9 ms for aspiration duration. Voiceless stops showed no voicing in the closure in 92% of the tokens.

Looking at the phonological implications for utterance-medial stops in Najdi Arabic, the voicing pattern for voiced stops was like that of voicing languages, regardless of speech rate or stress variations. Moreover, Voiceless stops were produced with no voicing in the closure in the majority of tokens which is a common pattern in aspirating languages, no matter the speech rate or stress variations. Aspiration was present, as well, in voiceless stops in iambic structure in the majority of tokens in normal speech but not in the majority of tokens in fast speech and definitely not in the majority of tokens in Voiceless stops in trochaic structures. These findings provide a robust indication that both Voiced and Voiceless stops in Najdi Arabic are produced with an active articulatory gesture. The presence of strong voicing that extends throughout the closure in Voiced stops requires active vocal folds vibration whereas the maintenance of full closure voicelessness and the production of long lag aspiration require active glottal abduction (Westbury, 1983; Ohala, 1997; Westbury and Keating, 1986). It is worth mentioning that the voicelessness of closure in specified Voiceless stops in intervocalic position is more important than the amount of aspiration as aspiration is expected to be shorter in non-initial stops (Beckman et al., 2013; Iverson and Salmon, 1995).

A single feature [spread glottis] or [voice] (privative or binary) cannot account for the phonetic manifestation of voicing contrast in Najdi Arabic in utterance-initial and utterance-medial intervocalic stops, as both of them are required. The pattern in Voiced stops in Najdi Arabic is similar to that of voicing languages such as Russian (Ringen and Kulikov, 2012) and Hungarian (Gósy and Ringen, 2009). On the contrary, the pattern of Voiceless stops in Najdi Arabic is similar to that of aspirating languages such as German (Jeesen, 1998; Jessen and Ringen, 2002). Although the duration of aspiration in Najdi Arabic might indicate that it occupies a middle position between voicing and aspirating languages, it is evident that the patterns of prevoicing, closure voicelessness, and aspiration, irrespective of duration, show the characteristics of both aspirating and voicing languages.

Another possible issue relates to the degree of speech rate effect on prevoicing and aspiration because the results for Najdi Arabic showed that prevoicing is more affected by speech rate than aspiration. However, it is hard to draw phonological conclusions based on this finding for several reasons. First, this issue has not received much attention in the literature of languages that show the features of both voicing and aspiration. Second, the degree of speech rate effect is not clear-cut due to the possibility of methodological variations between studies that investigated speech rate effects on duration (there might be a difference between normal and slow speech rate). Third, it might be challenging to consider the degree of speech rate effect as phonologically motivated because voicing and aspiration differ in their acoustic nature (voicing: periodic wave, aspiration: aperiodic wave). This is not, of course, to reject speech rate effect as a tool generally used for drawing phonological implications. Rather it questions the legitimacy of comparing the prevoicing and aspiration of specified stops as an analogy for the phonological representation based on the different susceptibility of prevoicing and aspiration to speech rate effect. However, Beckman et al. (2011) argued that the degree of speech rate effect can be used as the basis for a new categorisation within languages that show features of both voicing and aspiration; specifically, dialects deriving from originally voicing languages show more effect on prevoicing as opposed to dialects of aspirating languages which show more effect on aspiration. This assumption might fit the case of Najdi Arabic which is a dialect of a voicing language. However, more studies are needed to confirm or reject such a finding in similar dialects.

Moving on to **utterance-final stops**, the results revealed that Najdi Arabic showed considerable variation within the voicing targets for Voiced and Voiceless stops. At normal speech rate, Voiced stops were produced with complete voiceless closure in 59.2% of the tokens whereas the remaining tokens showed various degrees of voicing with an average of 13.7 ms. On the other hand, Voiceless stops were produced with heavy aspiration in 98.3% of the tokens with an average of 80.7 ms. Voiceless stops showed no voicing in the closure in 99% of the tokens. In fast speech, Voiced stops were produced with complete voiceless closure in 63.2% of the tokens. The remaining tokens showed various degrees of voicing with an average of 12.6 ms. Voiceless stops were produced with heavy aspiration in 59% of the tokens with an average of 56.3 ms and showed no voicing in the closure in 97% of the tokens.

Starting with the *complete laryngeal neutralisation* view proposed in the traditional approach in the formal linguistic frameworks, the results showed that Voiced and Voiceless stops preserved some degree of distinction in utterance-final position, a finding that supports the *incomplete laryngeal neutralisation* explanation proposed in laryngeal realism. Of note is that the majority of Voiced stops in utterance-final stops in Najdi Arabic were devoiced in both normal and fast speech rates, but a considerable percentage of them showed some voicing in the closure (40% in normal speech, 37% in fast speech). The devoicing process in Najdi Arabic can be explained as a case of passive devoicing, which is expected because utterance-final stops are a preferred position for devoicing, for aerodynamic and articulatory reasons. It is quite hard to maintain the air pressure difference required to enable vocal fold vibration in such a context (Belvins, 2004; Ohala, 1983, 1997, Iverson and Salmons, 2007). Considering the autosegmental approach proposed in Iverson and Salmons (1995, 2007), Najdi Arabic would be an interesting case to consider as both series of stops are proposed to be specified. Accordingly, it is problematic to assume the insertion (fortition) or delinking (loss) of [spread glottis] or [voice] in utterance-final stops.

The high incidence of final laryngeal neutralisation in both voicing and aspirating languages, with a lot of variability, makes it very hard to accurately predict the behaviour of stops in final position based on the voicing/aspirating distinction (Jansen, 2004). However, it could be assumed that voicing and aspirating languages might use different cues to maintain the distinction in stops in final position. It has been found that a voicing language, like French, increases the release burst properties of Voiceless stops while an aspirating language, like English, uses the preceding vowel duration (Flege and Hillenbrand, 1987). Najdi Arabic seems to behave like French in employing release burst properties rather than preceding vowel duration. The weight of the acoustic cues could be determined through a perception study to check what the listener relies on to detect the distinction, a subject beyond the scope of the present study.

Many previous studies that investigated speech rate effect on prevoicing and aspiration focused on utterance-initial stops. The present study, to my knowledge, might be the first to examine speech rate effect on utterance-final stops. The results showed that aspiration but not prevoicing showed a significant increase in slower speech. Based on this assumption, it could be assumed that Voiceless stops in Najdi Arabic are specified. Although the properties of utterance-final stops were increased, the distinction between aspiration for Voiced and Voiceless stops was preserved at both normal and fast speech rates which supports the argument for the specification of Voiceless stops in Najdi Arabic. Regarding Voiced stops, it can be proposed that since Voiced stops are specified and robustly prevoiced in utterance-initial and utterance-medial intervocalic positions, this assumption could extend to utterance-final position. That is, Voiced stops in utterance-final position are phonologically specified but passively devoiced.

9.2.2 The hierarchy of voicing correlates and phonologization

After examining the phonological implications for prevoicing and aspiration in the previous section, we now consider the contribution that other acoustic cues make to marking the phonological opposition between Voiced and Voiceless stops in Najdi Arabic. As made explicit in the present work, the integration model perspective proposed in the work of Jakobson and colleagues (1952, 1979, 1987), Jessen (1998, 2001), along with laryngeal realism, formed the foundation of the phonological and phonetic analysis that was implemented on voicing contrast in Najdi Arabic. That is, the phonetic reality is crucial in understanding the phonological representation. The previous two sections discussed the phonological implications of prevoicing/voicing duration and aspiration that were robustly realised in the majority of contexts of Voiced and Voiceless stops, respectively. Based on the hierarchical categorization proposed by Jessen (1998, 2001), voicing duration is the basic correlate for [voice] whereas aspiration duration is the basic correlate for [spread glottis] ([tense] in Jessen's model). The contribution of other correlates was considered in the present analysis to examine possible links between each correlate and the distinctive features [voice] and [spread glottis].

Closure duration results in utterance-medial and utterance-final stops showed inconsistency in that closure duration signalled the distinction in utterance-medial but not in utterance-final stops. Utterance-medial showed significantly longer Voiceless stops. This pattern has been confirmed for aspirating and voicing languages by many studies (Lisker, 1957; Kohler, 1984; Jacques, 1980; Al-Tamimi and Khattab, 2018). The distinction based on closure duration is not present in utterance-final stops in Najdi Arabic. The results showed the opposite pattern in alveolar stops, however; Voiced stops had significantly longer closure than Voiceless stops. The pattern in velar stops showed no difference between the two voicing categories. This manifestation of closure duration showing longer closure in Voiced stops was found in German (Jessen 1998) and Danish (Hutters 1985). Jessen (1998) postulated that the closure duration in Voiceless stops might be suppressed when the stop is heavily aspirated to increase the perceptibility of the aspiration phase. The results for Najdi Arabic support this claim in that Voiceless stops showed longer aspiration utterance-finally in comparison to the rest of the contexts which might have affected the closure duration. These findings indicate that closure duration is a non-basic correlate for [spread glottis] and [voice] in Najdi Arabic because it marks the distinction between the two voicing categories in a limited context; namely utterance-medial intervocalic position.

The results showed that there was no significant distinction between Voiceless and Voiced stops in terms of **the duration of the preceding vowel** across the examined contexts. Taking into account what was proposed in the literature regarding the inverse relation between closure duration and preceding vowel duration (Kohler, 1984; Kluender et al., 1988), it could be assumed that the ambiguity of closure duration in Najdi Arabic led to the suppression of voicing impact on the preceding vowel duration. Some other Arabic dialects showed no impact of voicing on the preceding vowel duration (Mitleb, 1984; Munro, 1993; De Jong and Zawaydeh, 2002; Al-Gamdi, 2013). The results for preceding vowel duration in Najdi Arabic indicate that this correlate is not effective in signalling the distinction which consequently means that it is not a correlate for [spread glottis] or [voice] in Najdi Arabic. As for **the following vowel duration**, the results showed a significant distinction whereby vowels following Voiceless stops were longer than following Voiced stops across the examined contexts. This is in agreement with what was found in aspirating and voicing languages (Jessen, 2001; Allen and Miller, 1999; Fischer-Jorgensen, 1968; Alghamdi, 1990; Al-Tamimi and Khattab, 2018). In terms of production, it is evident that the following vowel duration is shortened because of the time that is occupied by aspiration in Voiceless stops (Jessen, 2001). As for perception, the concept of enhancement is essential for understanding this variation in that the shortness of the following vowel increases the saliency of aspiration (Jessen, 2001). The opposite is found also in the inverse relation between the long following vowel and absence of aspiration in the context of Voiced stops. Accordingly, following vowel duration is a correlate for [spread glottis] and [voice] in Najdi Arabic. It is a non-basic correlate, however, due to 1) its limited contextual stability (not available in some contexts such as final stops) and 2) its dependency on the basic correlate: aspiration duration.

The results for release **burst duration and burst intensity** showed that both correlates were a robust cue to the distinction between Voiceless and Voiced stops in Najdi Arabic. Starting with release burst duration, the results were compatible with the prediction in the literature in that burst duration is longer in Voiceless than in Voiced stops (van Alphen and Smits, 2004). This distinction was found across all sources of variability in the factors included in the study. This finding suggests that burst duration is a correlate for [spread glottis] and [voice] in Voiceless and Voiced stops in Najdi Arabic, respectively. The results for burst intensity showed the opposite pattern; burst intensity for Voiced is higher than for Voiceless stops. Similar results were found in word-medial intervocalic stops in Lebanese Arabic (Al-Tamimi and Khattab, 2018). This pattern is against what was proposed in the literature. It was observed that burst intensity is higher in Voiceless stops because of the high air pressure that built up behind the constriction whereas the prevoicing in Voiced stops leads to a drop in the burst intensity (Jessen, 1998). The case of Najdi Arabic as well as Lebanese Arabic is problematic and might indicate a language-specific feature in Arabic. Since the present work and the work of Al-Tamimi and Khattab are the only two studies that looked at this correlate in the dialects of Arabic, it is hard to draw any conclusions.

As highlighted earlier in Chapter 2, it has been found that F0 onset is higher after voiceless stops than after voiced stops (House and Fairbanks, 1953; Ohde, 1984). This difference has been explained in terms of articulation in that F0 raising is a product of stiffening the vocal folds while F0 lowering is a repercussion of slacking the vocal folds (Halle and Stevens, 1971). Other explanations have proposed aerodynamic views in that F0 raising is a result of high airflow that follows the burst in aspirated stops (Hombert et al. 1979). Kingston and Diehl (1994), in contrast, proposed that F0 lowering after voiced stops is audience designed in that it is intended by the speaker to signal the voicing distinction. Taking into account the predictions of laryngeal realism, the presence of [voice] implicates active F0 lowering whereas the presence of [spread glottis] implicates active F0 raising. The results of F0 onset in Najdi Arabic showed a significant effect of voicing on the two series of stops. The difference in F0 onset present in Najdi Arabic is notably more than what has been reported in voicing languages such as Lebanese Arabic (Al-Tamimi and Khattab, 2018) and Russian (Kolikuv, 2012). However, it is relatively similar to the difference reported in Qatari Arabic which showed the features of both voicing and aspirating languages (Kulikov 2020). Nevertheless, it might be challenging to compare between these studies without considering the differences in their methodologies.

A number of studies have challenged the predictions of laryngeal realism. It has been found that some voicing languages such as French and Italian showed raised F0 after Voiceless stops (Kirby and Ladd 2016, 2018). Kirby and Ladd concluded that F0 raising after Voiceless stops in voicing languages is against the laryngeal realism prediction which posits that unspecified stops are not supposed to raise F0 onset since they are not produced with active articulatory gestures. Similar results have been reported in various languages (Bang et al., 2018; Dmitrieva et al., 2015). Kirby and Ladd (2018) further argue that this variation should be explained as a different voicing mechanism applied differently and indicates a language-specific implementations of voicing. That is, the distinction between voicing and aspirating languages alongside the proposed laryngeal specification in laryngeal realism does not account for F0 effect variation across languages (Kingston, 2007).

To get more in-depth analysis of F0 onset manifestation in Najdi Arabic, a linear mixed effect model was built to investigate interaction between voicing, and F0 onset across speech rates, gender, position, place of articulation, and the degree of aspiration. Speaker and word were added as random effects (The model, the statistical output, and the pairwise comparisons are presented in the Appendix C: table C-9). The predictions of the model are presented in figure and table 9.1 below.

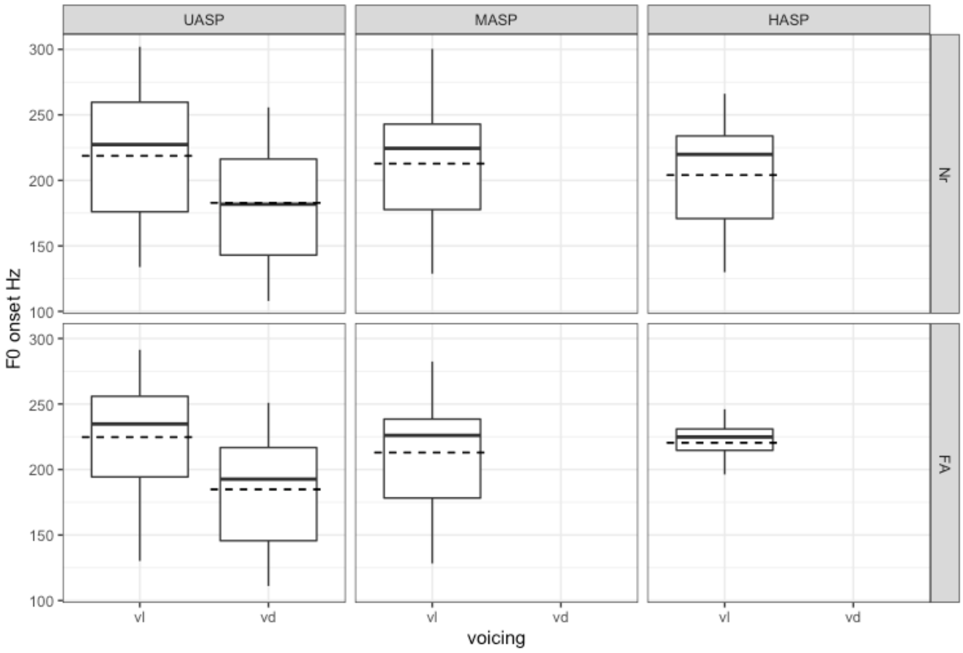


Figure 9.1. Boxplots of F0 onset values classified by voicing (vl = voiceless, vd = voiced), speech rate (FA = fast, Nr = Normal), and degree of aspiration (UASP = unaspirated, MASP = moderately aspirated, HASP = heavily aspirated see Table 5.1 Chapter 5). The dashed line represents the mean.

	Voiceless UASP		**Voiceless MASP**		**Voiceless HASP**		**Voiced UASP**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Normal	218	41.3	212	36.7	204	39.5	184	40.1
Fast	225	36.3	212	38.4	219	19.5	185	38.9

Table 9.1. Means and standard deviations of F0 onset values classified by voicing and speech rate.

The results presented in figure and Table 9.1 showed that F0 onset following Voiceless unaspirated stops was significantly higher than after moderately and heavily aspirated stops across speech rates. All the differences were found to be significant. The results also showed that F0 onset following Voiceless unaspirated stops was significantly higher than following Voiced stops by an average of 34 Hz in normal speech and 40 Hz in fast speech. It is important to note that aspiration in Voiced stops was measured regardless of the phonetic voicing in the closure. Aspiration as mentioned earlier (section 4.4.1) refers to the aperiodic noise that follows the burst.

These results reveal interesting discoveries about aspiration, voicing, and F0 onset. First of all, it can be noticed that F0 raising following voiceless stops indicates the presence of an active gesture and therefore implies the presence of [spread glottis] in Najdi Arabic. The implementation of F0 raising in Najdi Arabic is unlike that of aspirating languages which show F0 onset raising after aspirated stops rather than unaspirated stops such as in Cantonese

(Zee, 1980) and Danish (Petersen, 1983). The results also showed an inverse relationship between F0 onset and aspiration: the less aspiration the higher the F0 onset values (HASP in fast speech was produced by two speakers only S10, S29). It could be assumed that F0 onset raising is used to mark the voicing distinction in case of weaker aspiration (Jessen, 2001). However, this might not be the case in Najdi Arabic because of the presence of prevoicing and F0 lowering in Voiced stops which would be more than enough to mark the voicing distinction in case of short aspiration in Voiceless stops.

The results for **F0 offset** showed that this property was not robust in the distinction between Voiceless and Voiced stops in Najdi Arabic. F0 offset was slightly higher before Voiceless than before Voiced stops in utterance-medial position, but the difference was not significant. It was observed in the literature that F0 offset was not as effective as F0 onset in marking the voicing contrast in stops (Jansen, 2004; Kingston and Diehl, 1995; Kohler, 1984). Jessen (2001) did not include F0 offset as a correlate in his model indicating the ambiguity of this correlate. Taking into consideration that F0 offset received little attention in the studies that looked at voicing contrast in aspiration and voicing languages, more investigation of this acoustic property might result in more reliable findings.

Moving to **F1 onset**, the results suggest that F1 onset is a robust acoustic correlate that significantly marks the distinction between Voiceless and Voiced stops across the examined contexts. This distinction was reported in many aspirating and voicing languages (House and Fairbanks, 1953; Ohde, 1984; Kingston and Diehl, 1994; Jessen, 1998; Al-Tamimi and Khattab, 2018; Kulikov 2020). The lowering of F1 onset in the context of Voiced stops suggests an active articulatory adjustment such as larynx lowering that was reported in some voicing languages (Kingston and Diehl, 1994; Van Alphen and Smits, 2004; Al-Tamimi and Khattab, 2018). It was also reported for Qatari Arabic which is described as a dialect that has features of both aspirating and voicing languages (Kulikov, 2020). The raising of F1 onset after Voiceless stops in Najdi Arabic is compatible with the pattern found in aspirating languages (House and Fairbanks, 1953; Kingston and Diehl, 1994; Jessen, 1998). These results support the conclusion that F1 onset is a robust correlate for [spread glottis] and [voice] in Najdi Arabic. In terms of the results of **F1 offset**, similar to F0 offset, the results showed no significant distinction in Voiceless and Voiced stops. This again emphasizes that F0/F1 following the stop is a more effective cue than in the vowel preceding it (Kingston and Diehl, 1994; Jessen, 2001; Jansen, 2004).

The results of **H1-H2 onset** showed a distinction between Voiceless and Voiced stops in utterance-initial stops. As noted in the literature, this correlate is associated with glottal opening and indicates a breathiness in the onset of the vowel that is used as a cue for the distinction between Voiced and Voiceless stops in aspirating and voicing languages (Stevens and Hanson, 1994; Klatt and Klatt, 1990; Chapin Ringo, 1988; Ni Chasaide and Gobl, 1993). The difference in **H1-H2 offset**, however, showed more reliability in the differentiation between Voiceless and Voiced stops across the contexts. The results of **H1-H2 offset** in Najdi Arabic were more robust than the results reported in Lebanese Arabic (Al-Tamimi and Khattab, 2018) which emphasizes the necessity of looking at the manifestation of voicing contrast across the dialects of Arabic.

It can be noticed that there were some differences in the correlates that robustly mark the distinction between Voiceless and Voiced stops in Najdi Arabic based on the phonetic context. In utterance-initial stops, the most important correlates were aspiration, prevoicing, burst duration, burst intensity, F0 onset, F1 onset, H1-H2 onset, and FVD. In utterance-medial stops, the most important correlates were aspiration, voicing duration, burst duration, burst intensity, F0 onset, F1 onset, H1-H2 offset, and FVD. In utterance-final stops, the most important correlates were aspiration, burst duration, burst intensity, and H1-H2 offset. The acoustic correlates for Voiceless stops in Najdi Arabic showed a manifestation that is similar to what was reported for aspirated stops in aspirating languages. On the other hand, the acoustic correlates for Voiced stops showed a similar pattern to that was found in prevoiced stops in voicing languages. These findings form robust evidence for overspecification in the voicing system of Najdi Arabic in which both categories of voicing are specified; Voiceless stops are specified with [spread glottis] while Voiced stops are specified with [voice].

Najdi Arabic is a fruitful case to apply the hierarchical structure model of acoustic correlates proposed in the literature (Jessen, 2001). Which of the acoustic correlates in Najdi Arabic is “basic” for the voicing distinction and which of them is “non-basic” as in the terminology used by Jessen (2001)? There are two standards that were used by Jessen (2001): *contextual stability* and *perceptual saliency*. The former refers to the availability of the correlate in all contexts. For instance, release properties and voicing are available to be examined in all positions whereas closure duration is not (no closure in utterance-initial stops). The latter, however, refers to the perceptual sensitivity of the correlate that affects the listeners’ judgment. Many perceptual studies showed that prevoicing and aspiration are more salient than F0, F1, and H1-H2 (Diehl and Molis, 1995; Abramson and Lisker, 1985). Although determining the perceptual saliency of voicing correlates in Najdi Arabic is beyond the scope of the present study, it could be initially assumed that prevoicing and aspiration have more weight in marking the distinction than the rest of the correlates. It can also be initially assumed, based on the production results, that preceding and following vowel duration are less perceptually salient for Najdi speakers. This assumption was supported by several studies that examined voicing effect on preceding vowel duration (Mitleb, 1984; Al-Gamdi, 2013; Zawaydeh and De Jong, 2002).

According to the Jakobsonian approach and the integration model perspective proposed in the work of Jessen (1998, 2001) and laryngeal realism (Kohler, 1984; Keating, 1984; Kingston and Diehl, 1995; Jessen, 1998, 2001; Harris, 1994; Iverson and Salmons, 1995; Jessen and Ringen, 2002; Honeybone, 2005; Beckman et al., 2013; Al-Tamimi and Khattab, 2018; Al-Gamdi et al., 2019; Kulikov, 2020), there is a robust connection between phonetic concreteness and the distinctive features that specify voicing contrast. This connection is manifested in phonetic aspects including articulation, acoustics and perception. Following Jessen's model, the present study employed acoustic analysis to characterize the voicing contrast in Najdi Arabic and considered the principle of 'contextual stability' to evaluate the strength of the correlate in marking the opposition. Some of the acoustic correlates reflect an active articulatory event such as the robust prevoicing in Voiced stops in utterance-initial stops which indicates active vocal folds vibration (Jessen, 2001). The importance of perception in the distinction between the two categories of voicing stems from the fact that, despite a distinction in the production, the perceptibility or 'perceptual saliency' (Kingston and Diehl, 1995; Jessen, 2001) of the correlate is another tool that determine to what extent this correlate is crucial for distinction. The acoustic results of the present study are a starting point that will enrich perception analysis in future research that aims to depict the voicing contrast in Najdi Arabic or in any other dialect that shows features of both aspirating and voicing languages.

The results for the acoustic correlates that signal voicing distinction in Najdi Arabic, so far, provide a compelling justification for the proposal that the voicing contrast in Najdi Arabic is overspecified with two features in the phonology: [spread glottis] and [voice]. This proposal is based on the presence of active prevoicing in utterance-initial and utterance-medial voiced stops, as well as the presence of moderate aspiration in utterance-initial voiceless stops. Additionally, both prevoicing and aspiration were found to be affected by speech rate variations (slower rate yielded longer duration). Although overspecification was not expected in the laryngeal realism approach, the tools used to diagnose the distinctive features including active/passive (de)voicing and speech rate effect seemed indicative of the presence of two active features in the laryngeal system of Najdi Arabic. The results also showed that Najdi Arabic apparently has taken a middle position between voicing and aspirating languages. Beckman et al., (2013) proposal of numeric values for the phonetic distinctive features might be effective in the description of voicing contrast in Najdi Arabic (detailed description of this proposal is presented in section 2.2.1). A special section will look at the implementation of the numeric values in Najdi Arabic which will consider all the phonetic contexts including voicing assimilation in clusters.

An important question might be raised with regard to the middle position that Najdi Arabic takes between voicing and aspirating languages. That is, is Najdi Arabic closer to voicing or aspirating languages? One useful method of classification is to look at voicing assimilation in stop-stop clusters across word boundaries. It is an approach that may provide insights about the activeness and specification of voiced and voiceless stops in Najdi Arabic. This issue will be examined in the following section which focuses on the phonetic and phonological aspects of voicing assimilation in stop-stop clusters across word boundaries in Najdi Arabic. After presenting the characterization of voicing assimilation in Najdi Arabic, a clearer, sharper picture of voicing contrast in Najdi Arabic that covers almost all possible phonetic contexts informative for determining the status of the laryngeal contrast in Najdi Arabic will enhance our understanding of the aspects of voicing contrast in Najdi Arabic.

9.2.3 The phonological implications for the acoustic correlates of stop-stops clusters

This section aims to shed light on the phonological implications of the acoustic correlates of stop-stop clusters across word boundaries in Najdi Arabic. It has been established in many studies in the literature that voicing assimilation is a common process in many languages. Various phonological theories have attempted to characterise the manifestation of the phonological specification in terms of the occurrence of assimilation in consonant clusters across word boundaries. The following paragraphs summarise the basic predictions of each of these attempts and compare them with the results of voicing assimilation in Najdi Arabic in order to reach conclusions regarding the interactions between the phonological and phonetic aspects of such a process in Najdi Arabic.

The traditional view of phonological feature analysis assumes the complete phonetic neutralisation of C1 to C2 across word boundaries (Chomsky and Halle 1968). That is, the voicing correlates of C1 and C2 must be identical in the case when the process takes place. This assimilation between stops across word boundaries, in the traditional view, is categorical, meaning that C1 is fully assimilated or fully unassimilated to C2. The traditional phonological approach implicates that all acoustic correlates associated with [voice] are expected to participate in the assimilation process (Jansen 2004) which means both temporal and spectral acoustic correlates carry over from C2 to C1. The traditional view also assumes that the ability of Voiced or Voiceless stops to participate in the assimilation process is not determined by their phonetic manifestation. Moreover, the traditional view posits the presence of [-voice] in the phonology. Therefore, it is expected based on this assumption that [-voice] C2 would devoice the preceding stop in consonant clusters across word boundaries.

Looking at the results of stop-stop cluster assimilation in Najdi Arabic allows us to test the accuracy of the predictions of the feature phonological approach. Starting with the complete phonetic neutralisation prediction, the results showed that both Voiced and Voiceless stops in C1 position preserved some of their voicing properties in all cluster patterns [voiced-voiceless], [voiced-voiced], and [voiceless-voiced]. The distinction between Voiced and Voiceless stops in C1 were maintained in both the baseline and the cluster contexts. For instance, it has been found that the majority of Voiced stops in C1 position in [voiced-voiceless] clusters have voicing in their hold phase. On the other hand, Voiceless stops were found fully voiceless in all tokens in [voiceless-voiced] clusters. These results refute the basic assumption of the traditional approach. With regard to the cues that participate in the assimilation process, this prediction was not tested due to the possible lengthening of the durational correlates in the baseline context (utterance-final position). Accordingly, it is hard to assume that any change in the durational correlates is actually caused by assimilation and not by the phonetic context. Nevertheless, the nature of the durational correlates might be problematic in terms of the spread from the trigger consonant to the target one in which that time is not expected to participate in a coarticulation process (Jansen, 2004). Moving to the activeness of [-voice] in the assimilation process, the binary featural system would propose that Voiceless stops in Najdi Arabic are specified with [-voice]. The results of [voiced-voiceless] clusters in Najdi Arabic showed that voiced stops were partially devoiced in the majority of the tokens, thereby supporting the claim that [-voice] is active in the phonology.

Laryngeal realism approach posits a different view regarding voicing assimilation in stop-stop clusters across word boundaries. As discussed earlier, laryngeal realism postulates a transparent and strong connection between the phonetic details and the phonological representation which is attractive on the empirical grounds. In terms of voicing assimilation, laryngeal realism assumes that the process of assimilation is gradient and consequently incomplete. The ability of Voiced and Voiceless stops to trigger or undergo voicing assimilation is determined by their phonological specification status (Iverson and Salmons, 1995; Jansen, 2004; Honeybone, 2005). That is, specified stops are expected to trigger (de)voicing whereas unspecified stops are expected to undergo passive (de)voicing. Accordingly, it has been proposed that Voiceless stops in aspirating languages and Voiced stops in voicing languages are predicted to influence the voicing of the neighbouring sound. With regard to the voicing correlates predicted to spread from the trigger to the target in the assimilation process, the correlates subject to assimilation are predicted to be the ones that are associated with active articulatory movement (Jansen, 2004). On the other hand, passively (de)voiced stops produced without active gestures for voicing production are not predicted to spread any form of (de)voicing to the adjacent sound.

Voicing assimilation in Najdi Arabic posits an interesting case for the analysis under the laryngeal realism approach. Assuming Najdi Arabic has the features of both aspirating and voicing languages and employs both [voice] and [spread glottis] to specify the voicing distinction, laryngeal realism would predict Voiced and Voiceless stops are expected to trigger voicing assimilation in the adjacent sound. By examining the results presented in this chapter, it is important first to note that voicing assimilation in stop-stop clusters in Najdi Arabic is gradient and incomplete which can be observed in all the acoustic correlates examined in both C1 and C2 stops in all patterns. The distinction between Voiced and Voiceless stops in C1 and C2 positions were acoustically maintained in all the cues investigated in the analysis. By delving deeply into the acoustic details of Voiced and Voiceless stops in C1 position, it can be noticed that both Voiced and Voiceless stops were affected to a certain degree by the voicing status of C2. Starting with Voiceless stops, the acoustic analysis revealed that Voiceless stops were prone to voicing assimilation in F0/F1 and burst intensity but not in voicing in the closure. For Voiced stops, the results showed a degree of devoicing in their closure but not in F0/F1 and burst intensity. These results support two main assumptions: 1) both Voiced and Voiceless stops trigger some form of (de)voicing in C1, 2) both Voiced and Voiceless stops in C1 position showed some degree of resistance to the assimilation process in that voiceless stops retained their voicelessness during the hold phase while voiced stops retained the voicing features of their F0/F1(lowering) and burst intensity (raising).

These results entail questions about the behaviour of Voiced and Voiceless stops in Najdi Arabic not explicitly addressed in the laryngeal realism approach. First of all, it has been proposed in laryngeal realism that specified stops trigger (de)voicing and do not undergo passive (de)voicing (Jansen, 2004; Kulikov, 2012; Beckman et al., 2013). This conclusion builds from the assumption that [voice] is the privative feature specifying voicing contrast in voicing languages, and [spread glottis] is the privative feature specifying voicing contrast in aspirating languages. Voicing assimilation in Najdi Arabic is problematic because both the trigger (C2) and the target (C1) are predicted to be actively (de)voiced. The issue regarding this assumption is how to determine to what extent each of Voiced and Voiceless stops trigger (de)voicing (C1) or undergo (de)voicing (C2), in addition to which voicing properties are predicted to be spreading from the trigger, and what voicing properties are predicted to resist the change in the target. It is noticeable from the results that F0/F1 and burst intensity were spilled over and maintained in Voiced stops in C1 and C2, respectively. On the other hand, voicelessness during the hold phase was the correlate that spread and was maintained in Voiceless stops in C1 and C2, respectively. These results might accord with the prediction that voicing properties subject to spreading are those associated with an active articulatory event in the production of the stop. It could be proposed that F0/F1 and burst intensity in Voiced stops in Najdi Arabic are products of the articulatory mechanism that facilitate the production of voicing including larynx lowering and tongue root advancement (Jansen, 2004).

Interestingly, the effect of C1 on C2 (progressive (de)voicing) appeared to be almost in line with the same pattern in regressive voicing in Najdi Arabic. The results showed that Voiced stops were slightly devoiced when preceded by a Voiceless stop. Voicing duration of Voiced stops in C2 position tended to be shorter than in the baseline context in [voiceless-voiced] cluster. However, Voiced stops tended to maintain the lowering of F0/F1 with more significant impact on F1 than on F0, supporting the previous finding that Voiced stops retain their effect on F0/F1 while their hold phases were slightly devoiced. Voiceless stops, on the other hand, retained their voicelessness during the closure in all the tokens of [voiced-voiceless] clusters. In terms of F0/F1 for Voiceless stops, the results revealed some variations based on the place of articulation of C2. F0/F1 values in [gt] were significantly lower in the cluster context than in the baseline context. On the other hand, [bk] showed no significant effect on F0/F1follwoing C2.

The present study is not the first to investigate a language that has both prevoicing and aspiration: this feature has been investigated in Hebrew (Raphael et al., 1995), Najdi Arabic (Flege and Port, 1981; Al-Gamdi et al., 2019), Swedish (Helgason and Ringen, 2008; Beckman et al., 2011), Norwegian (Ringen and Dommelen, 2013), Turkish (Unal-Logacev et al., 2018; Ogut et al., 2006; Feizollahi, 2010), Saudi North dialect (Alanazi, 2018), and Qatari Arabic (Kulikov, 2020). However, none of these studies described the voicing assimilation in stop-stop clusters across word boundaries in terms of their phonological or phonetic aspects. Laryngeal realism does not attempt to account for such languages, making it difficult to precisely dissect the typological properties that these languages might share. One of the suggested solutions to characterise the phonetic and phonological aspects of voicing assimilation in Najdi Arabic is to employ the numeric values of the phonetic distinctive features proposed by Beckman et al. (2013). This issue will be discussed in detail in the next section.

The following section concentrates on the parallelism between the types of evidence used in laryngeal realism in the light of the results of analysis of voicing contrast in Najdi Arabic in the present study.

9.3 The compatibility between the types of evidence used in laryngeal realism

The predictions of the theory of laryngeal realism originate from the assumption of a robust connection between acoustic details and phonological features. Because of the complexity and variance in the acoustic signal, laryngeal realism posits an articulatorily oriented explanation that systematizes the evaluation of whether a specific acoustic feature is phonetically or phonologically motivated. Specifically, phonologically motivated correlates are produced by an active articulatory gesture whereas phonetically motivated correlates result from coarticulation with no active gesture. Accordingly, the types of evidence that are used in laryngeal realism include the phonetic cues of the segment, speech rate effects, and voicing assimilation. Among the phonetic cues of the segment, the types of evidence for actively specified Voiced stops include the presence of robust prevoicing in utterance-initial and intervocalic positions. In contrast, the types of evidence for voiceless stops include the presence of long lag aspiration in utterance-initial stops and active devoicing in intervocalic stops. As for the speech rate effect, it is expected that prevoicing in specified Voiced stops and aspiration in the specified Voiceless stops will be lengthened in slower rate speech in comparison to fast rate speech. For voicing assimilation, laryngeal realism posits that specified stops are expected to trigger some (de)voicing to the adjacent sound in stop-stops clusters.

The previous paragraph is a brief synopsis of the main types of evidence proposed in the literature of laryngeal realism. All these types have been tested in the present study of voicing contrast in Najdi Arabic. As shown in Chapter 5, 6,7, and 8, here are the main findings:

a) Voiced stops were realised with robust prevoicing, longer burst duration, longer following vowel, higher burst intensity, and F0/F1/H1-H2 onset lowering in utterance-initial position.
b) Voiced stops were realised with robust voicing in the closure that extends throughout the closure phase, longer closure, short burst duration, longer following vowel, high burst intensity, low H1-H2 offset, and low F0/F1 onset in utterance-medial positions.

c) Voiced stops were passively devoiced in the majority of tokens in utterance-final position. The distinction was maintained by release properties: high burst intensity, short burst duration, low H1-H2 offset, and short aspiration.
d) Voiceless stops were realised with moderate aspiration in utterance-initial and heavy aspiration in utterance-final positions and to a lesser extent in utterance-medial positions (more aspirated tokens in iambic than trochaic structures), long burst duration, low burst intensity, and high F0/F1 onset.
e) Voiceless stops were actively voiceless (no voicing in the closure) in utterance-medial and utterance-final positions.
f) Voiceless stops were realised with long burst duration, low burst intensity, and long aspiration.
g) Both prevoicing and aspiration were longer in normal speech rate in comparison to fast speech rate.
h) Voiced stops in C2 trigger some voicing in voiceless stops in C1 (F0/F1 offset lowering) and in voiced stops in C1 (more voicing in the closure, F0/F1 offset lowering).
i) Voiceless stops in C2 trigger some devoicing in voiced stops in C1 (shorter voicing)

It is notable that all the results accord with the predictions of laryngeal realism regarding the phonetic cues of voiced and voiceless stops; the response of aspiration and prevoicing to change in speech rate; and the triggering of (de)voicing in the preceding stops.

However, one of the issues in need of more attention is the degree of change in aspiration and prevoicing as a response to speech rate. The present study extended the use of speech rate effect to utterance-medial and utterance-final stops. The results revealed that prevoicing is more affected by change in speech rate in utterance-initial position whereas aspiration is more affected by change in speech rate in utterance-final position. Looking at the former, this pattern has been found in Swedish and Qatari Arabic where the effect is shown more in prevoicing than in aspiration (Beckman et al., 2011; Kulikov, 2020). Of note is that the response of aspiration to change in speech rate in Najdi Arabic in initial stops (prevoicing: 24 ms, aspiration: 6.4 ms) is relatively small when compared to that found in Swedish (prevoicing: 29.4, aspiration: 18.7 ms) or Qatari Arabic (prevoicing: 19 ms, aspiration: 12 ms). When comparing the results with voicing and aspirating languages, the difference between normal and fast speech in prevoicing in Najdi Arabic is like that found in voicing languages, such as Russian (Kulikov, 2012) and French (Kessinger and Blumstein, 1997).

However, the difference in aspiration appears to be the smallest among aspirating languages (Allen and Miller, 1999; Beckman et al., 2011). These differences might be caused by variations within the studies in their participants' ability to express the difference between slow (or normal) and fast speech. Another possibility is that the small effect of speech rate on aspiration is language-specific and related to Arabic phonology (Kulikov, 2020) which might be the case for both Qatari and Najdi Arabic which showed the least effect of speech rate on aspiration.

Comparing the results of Najdi and Qatari Arabic, in Qatari Arabic 77% of initial stop tokens were prevoiced (Kulikov, 2020). In Najdi Arabic, on the other hand, 90% of tokens were prevoiced. Kulikov (2020) argues that this devoicing in Qatari Arabic might be caused by the growing use of English among younger native speakers of Qatari Arabic. These differences lead us to consider the variation among dialects of Arabic in the phonetic manifestation of voicing contrast. The mean aspiration found in Qatari Arabic in Kulikov's study was 54 ms for /t/ and 62 ms for /k/. In the Flege and Port (1981) study of Najdi Arabic, mean aspiration was 37 ms for /t/ and 54 ms for /k/, while in the Alghamdi (1990) study of Saudi Ghamidi dialects, mean aspiration was 32 ms for /t/ and 42 ms for /k/. In Alanazi's (2018) examination of North Saudi Arabic, mean aspiration was 58 ms for /t/ and 72 ms for /k/. In the Al-Gamdi et al. (2019) study of Najdi Arabic, mean aspiration was 76 ms for /t/ and 82 ms for /k/. Table 7.2 below summarises the reported aspiration values across peninsular Arabic dialects. (Prevoicing is not reported due to the fact that all the studies looked at stops in non-utterance-initial position which made it possible that their voicing might be coming from the preceding consonant).

	Source	Mean aspiration	
Najdi	Flege and Port (1981)	/t/ = 37 ms	/k/ = 54 ms
Ghamidi	Alghamdi (1990)	/t/ = 32 ms	/k/ = 42 ms
North Saudi	Alanazi (2018)	/t/ = 58 ms	/k/ = 72 ms
Najdi	Al-Gamdi et al. (2019)	/t/ = 76 ms	/k/ = 82 ms
Qatari	Kulikov (2020)	/t/= 54 ms	/k/ = 62 ms
Najdi	Present study	/t/ = 45 ms	/k/ = 49 ms

Table 9.2. Mean aspiration for /t/ and /k/ in modern peninsular dialects of Arabic.

It can be noticed that the differences between means of aspiration among this subset of modern Arabic dialects are not great, ranging from 32 to 76 for /t/ and from 42 to 82 for /k/.

The variability in the phonetic cues of voiced and voiceless stops within across dialects and the languages that show the features of both voicing and aspirating languages raises questions about the features that specify voicing contrast in the phonology of these languages. It has been proposed in the previous section on phonological implications in Najdi Arabic that both [voice] and [spread glottis] are active in the phonology. To address the variability among the phonetic cues in voicing and aspirating languages, Beckman et al. (2013) proposed a framework using numeric values for phonetic distinctive features as a scale to address the strength of the presence of the features in the phonology. Based on this framework, voiceless stops in aspirating languages, such as German are specified with [9 spread glottis] while voiced stops are unspecified with [1 spread glottis], whereas voiced stops in voicing languages, such as Russian are specified with [9 voice], and voiceless stops are unspecified with [1 voice].

Applying the same framework to Najdi Arabic, based on the results of the acoustic correlates in Najdi Arabic, it could be proposed that voiced stops in Najdi Arabic are specified with [9 voice] while voiceless stops are specified with [8 spread glottis]. [9 voice] is evident in case of Najdi Arabic because of two important findings: a) the presence of robust voicing in the majority of the tokens in utterance-initial and utterance-medial intervocalic stops, b) voiced stops in C2 position trigger voicing in the preceding C1 in stops-stops clusters. [8 spread glottis] for voiceless stops in Najdi Arabic is proposed because of two findings: a) the presence of (moderate) aspiration in utterance-initial position, b) voiceless stops in C2 trigger some devoicing in C1 in stop-stop clusters, c) the voiceless stops do not undergo passive voicing in the majority of the tokens in utterance-medial intervocalic position. Considering that the types of evidence proposed in laryngeal realism fit the findings found in Najdi Arabic, it is recommended that a new category be added to the typology of Beckman et. al.'s framework presented in table 9.3 below. (For intervocalic stops).

Aspirating languages		
German type		
voiceless	[9sg] [Øvoice]	Passive voice cannot apply, glottal width too great
voiced	[1sg] [[Øvoice]	Passive voice applies, small glottal width, no numerical specification for [voice]
Icelandic type		
voiceless	[9sg] [Øvoice]	Passive voice cannot apply, glottal width too great
voiced	[5sg] [Øvoice]	Passive voice cannot apply, glottal width too great
Voicing languages		
Russian		
voiceless	[1voice] [Øsg]	Passive voice cannot apply; phonetic rules do not change numerical specifications
voiced	[9voice] [Øsg]	Active voice
Voicing Aspirating languages		
Najdi Arabic		
voiceless	[8sg] [[Øvoice]	Passive voice cannot apply, glottal width too great
voiced	[9voice] [[Øsg]	Active voice

Table 9.3 Summary of analysis in Beckman et al.'s framework (Najdi Arabic is added to the analysis)

One of the contributions of the present study is providing a clear justification for determining the numeric values for the phonetic distinctive features proposed in Beckman et al. (2013). Moreover, the results of the present study enrich the framework by adding a possible new category extending the predictions of laryngeal realism. It can be assumed that for a language to be classified using the numeric distinctive phonetic features, it is crucial to test Voiced and Voiceless stops across the phonetic contexts that might be informative for the specification of the stops such as utterance-initial and utterance-medial intervocalic, as well as in stops-stop clusters. In terms of Arabic dialects, it could be assumed that the features for Qatari Arabic are [8voice] and [8 spread glottis], where [8voice] results from the higher percentage of devoiced voiced stops than in Najdi Arabic. Nevertheless, it is hard to characterize the voicing system in Qatari Arabic without checking the phonetic manifestation in the remaining positions not tested in Kulikov (2020).

9.4 Limitations of the present study and suggestions for future research.

The present study examined voicing contrast considering various linguistic and non-linguistic factors by investigating various phonetic contexts exclusively for stops. It is important to look at the other classes of obstruents, including fricatives and affricates, for by doing so, a clearer picture of the laryngeal system of Najdi will emerge. Specifically, the specification of fricatives can be examined by checking if Voiced and Voiceless fricatives trigger (de)voicing in adjacent sounds. As the symmetry between the stops and fricatives has been proposed in some studies under the operation of laryngeal realism (Jansen, 2004; Beckman and Ringen, 2009), it could be suggested that investigating fricatives in Najdi Arabic might enrich the study of laryngeal phonology.

One of the main goals of the present study was to clarify which acoustic correlates were robust in marking the voicing distinction in Najdi Arabic. Although a considerable number of durational and spectral acoustic correlates have been found acoustically strong in signalling the voicing distinction, the relative contribution of each correlate and how the correlates mutually mark the distinction have not been examined in the present study. Such an approach has been employed in various studies using Random Forests statistical analysis which aims to specify the weight of each of the correlates in the voicing contrast (Al-Tamimi and Khattab, 2018). Another way could be by testing the perceptual saliency proposed by Jessen (2001) in which the importance of a correlate is determined by checking its role in listeners' judgement in perceiving the sound.

In a voicing assimilation experiment, the correlates tested were voicing duration, F0, F1, and burst intensity. None of the other durational correlates were tested because of the possibility of lengthening taking place in the baseline context in the experiment (utterance-final position). It could be suggested the baseline context should be in a neutral position not affected by lengthening as used in Jansen (2004).

An issue the present study has not discussed is a sound change that is possibly taking place in the modern dialects of Saudi Arabia. As highlighted earlier, the presence of long lag aspiration in Voiceless stops was found in Najdi Arabic, North Saudi Arabic, and Ghamidi Saudi Arabic (Flege and Port, 1981; Alanazi, 2018; Alghamdi, 1990). Therefore, additional studies are needed to decide the source and scale of the change in the voicing system of modern Saudi Arabic dialects.

9.5 Conclusion and final remarks

The purpose of the present study was to investigate the phonetic and phonological aspects of voicing contrast in stops in Najdi Arabic. The investigation was carried out considering various contexts within the word and the utterance. Furthermore, the study looked at some phonological processes including speech rate effect on voicing and aspiration, passive and active voicing, final laryngeal neutralisation, and regressive voicing assimilation in stop-stop clusters across word boundaries.

The robust connection between the phonetic reality and the distinctive features lies at the heart of the present study. As made evident throughout the stages of the analysis, the integration model perspective was adopted by employing what has been used in the literature of laryngeal realism as tools to examine the two-way interactions between the acoustic signal and the phonological features in Najdi Arabic which is a variety that shows features of both aspirating and voicing languages.

Chapter 1 started with an opening section that highlighted the purpose and the importance of the present study. It concentrated on the general consensus with respect to voicing contrast in aspirating and voicing languages and how the case of voicing contrast in Najdi Arabic, as well as other languages that show the features of both aspirating and voicing languages, is worthy of researchers' attention. It also presented a brief sketch of the basic principles of laryngeal realism that aimed to prepare the reader for the theoretical discussions in the remaining part of Chapter 1. The remaining part of Chapter 1 focused on three essential topics: distinctive features, the phonological and phonetic specification, and the phonetics-phonology interactions. These three topics formed the basic foundation that the arguments that the present study was built on. The first topic provided the context around the emergence of distinctive features and how the notion of *distinctivity* in the acoustic signal is linked to the phonological representation in the Jakobsonian framework. The second topic explored the definition of specification and how to differentiate the phonetic behaviour of specified and unspecified segments. The link between passive and active voicing and specification was discussed as well. The third topic described the theoretical discussions that characterise the interaction between phonetics and phonology and how this is implemented in the present analysis.

Chapter 2 presented the phonetic and phonological aspects of voicing contrast across languages. The theoretical models of voicing were discussed in the light of three main issues: 1) the nature of the distinctive features, 2) the connection between the phonetic reality and the distinctive features, and 3) the predictions for languages that show features of aspirating and voicing languages. A special section concentrated on the basic principles of the laryngeal realism approach and the types of evidence employed to examine the connection between the acoustic signal and the distinctive features. This chapter also focused on the production of voicing and aspiration and the required articulatory adjustments to initiate, maintain, or inhibit them. The acoustic correlates found to mark the distinction of voicing contrast in the literature were discussed with a focus on the aspirating/voicing categorization and what that means in the case of Najdi Arabic. A special section previewed the studies that looked at languages with two-way contrast that employ both aspiration and voicing to signal the voicing distinction.

Chapter 3 was more specific to studies that looked at voicing contrast in stops in the modern dialects of Arabic. This chapter was divided into two main subsections: 1) studies that focused on phonetic aspects only and 2) studies that adopted the laryngeal realism approach in the investigation of voicing contrast in dialects of Arabic. The basic features of Najdi Arabic were presented. This chapter ended by presenting the rationale of the study and why it was crucial to examine voicing contrast in Najdi Arabic. The research question and primary predictions were discussed.

Chapter 4 presented the methods of the analysis including participants, stimuli, procedures, acoustic analysis, and statistical analysis. To achieve the goals of the study and answer the research questions, the acoustic and statistical analysis were carefully constructed considering all variability in the data. The tokens produced by 32 native speakers of Najdi Arabic were examined using Praat software to capture the acoustic details for the correlates that signal the distinction between Voiceless and Voiced stops in Najdi Arabic. The extracted results were analysed employing a series of Linear Mixed Effects (LMM) statistical models that were built to investigate the voicing contrast in each context. A very detailed set of criteria was used to describe voicing, devoicing, aspiration in the examined stops in each context.

Chapter 5, 6, 7, and 8 presented the acoustic results for the correlates in Voiceless and Voiced stops in utterance-initial, utterance-medial, utterance-final, and stop-stop clusters. The section for each correlate was closed with a short summary that provided a brief discussion of the results and linked them with the previous studies. The results of the analysis provided the answers for the first research question: What are the acoustic correlates of stop-voicing contrast in Najdi Arabic and how are they implemented across the following phonetic contexts: utterance-initial, utterance-medial intervocalic, utterance-final, and across-word-boundary clusters.

Chapter 9 reviewed the results of the acoustic analysis provided in each context in Chapter 5, 6, 7, and 8 and discussed the phonological implications. There was adequate evidence that Voiceless stops are specified with [spread glottis] while Voiced stops are specified with [voice]. This was supported by the acoustic characteristics of Voiceless stops which behave similar to aspirating languages and Voiced stops which behave similar to voicing languages. The robust voicing found in Voiced stops in utterance-initial and utterance-final contexts indicates the presence of active voicing. The presence of aspiration and complete devoicing of closure in Voiceless stops across the contexts indicated the presence active devoicing. The results for the speech rate effect showed that both voicing and aspiration were longer in slow speech compared to fast speech. The final devoicing (laryngeal neutralisation) in stops in utterance-final stops indicated that Voiceless stops were heavily aspirated while Voiced stops were passively devoiced. The neutralisation was incomplete in that the distinction was preserved in the release properties. The results for regressive voicing assimilation in stop-stop clusters showed that the two categories of voicing triggered some (de)voicing in the preceding stop and they both showed resistance to change in the process. All these findings presented in Chapter 9 provided the answer for the research question: Employing the laryngeal realism approach, how does the voicing system of Najdi Arabic behave in terms of the following processes: speech-rate effect on the acoustic correlates of stops across the examined phonetic contexts, the acoustic activeness of voicing/devoicing of stops across the examined phonetic contexts, and regressive voicing assimilation in across-word-boundary clusters.

To answer the third question which stated that: is Najdi Arabic a voicing or an aspirating language? What does that mean in terms of the phonological representation/specification? The present work argues that Najdi Arabic takes a middle position between voicing and aspirating languages and used the numeric distinctive features proposed by Beckman et al. (2013) to describe the overspecification in voicing contrast in Najdi Arabic. Based on the acoustic results, it was concluded that voicing contrast in Najdi Arabic is specified with [8spread glottis] and [9 voice].

This study was an attempt to draw researchers' attention to the phonetic and phonological aspects of languages that show features of both aspirating and voicing languages. It aimed to provide insightful analysis that will enrich the theories that focus on the connection between phonetics and phonology. The present work might inspire more investigations of the modern dialects of Arabic which showed variation in their laryngeal systems. On the basis of the current study, I argue for more in-depth analysis of the phonetic and phonological behaviour of Voiceless and Voiced stops in languages with two-way contrast that contrast prevoiced and aspirated stops; it might be insufficient to limit the investigation to reporting the duration values for moderate VOT.

References

Abboud, P.F. (1964) *The syntax of Najdi Arabic*. PhD thesis. The University of Texas at Austin.

Abboud, P.F. (1979) 'The verb in northern Najdi Arabic1', *Bulletin of the School of Oriental and African studies*, 42(3), pp. 467-499.

Abdelli-Beruh, N.B. (2004) 'The stop voicing contrast in french sentences: Contextual sensitivity of vowel duration, closure duration, voice onset time, stop release and closure voicing', *Phonetica*, 61(4), pp. 201-219.

Abdelli-Beruh, N.B. (2009) 'Influence of place of articulation on some acoustic correlates of the stop voicing contrast in Parisian French', *Journal of Phonetics*, 37(1), pp. 66-78.

Abramson, A.S. and Lisker, L. (1970) *Proceedings of the sixth international congress of phonetic sciences*. Prague.

Abramson, A.S. and Lisker, L. (1985) 'Relative power of cues: F0 shift versus voice timing', *Phonetic linguistics: Essays in honor of Peter Ladefoged*, pp. 25-33.

Abramson, A.S. and Whalen, D.H. (2017) 'Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions', *Journal of phonetics*, 63, pp. 75-86.

Alanazi, S. (2018) *The acquisition of English stops by Saudi L2 learners*. PhD thesis. University of Essex.

Alezetes, E.D. (2007) *A Markedness Approach to Epenthesis in Arabic Speakers' L2 English*. PhD thesis. The university of Montana.

Al-Gamdi, N., Al-Tamimi, J. and Khattab, G. (2019). 'The acoustic properties of laryngeal contrast in Najdi Arabic initial stops', *In Proceedings of the 19th International Congress of Phonetic Sciences*. Melbourne.

Al-Gamdi, N.A. (2013). *Saudi ESL learners' awareness of voicing effect on the preceding vowel duration in CVC English words*. MA thesis. Southern Illinois University.

Al-Masri, M. and Jongman, A. (2004). 'Acoustic correlates of emphasis in Jordanian Arabic: Preliminary results', *In Proceedings of the 2003 Texas Linguistics Society Conference*. Somerville, MA: Cascadilla Proceedings Project, pp. 96-106.

Al-Sweel, A.A.I. (1987) 'Verbal and nominal forms of Najdi Arabic', *Anthropological linguistics*, pp. 71-90.

Al-Tamimi, J. (2017). 'Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations'. *Laboratory Phonology*, 8(1), pp. 1-40

Al-Tamimi, J. and Khattab, G. (2018) 'Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops', *Journal of Phonetics*, 71, pp. 306-325.

Alghamdi, M. (1990) *Analysis, synthesis and perception of voicing in Arabic*. PhD thesis. University of Reading.

Alghmaiz, B.A. (2013) *Word-initial consonant cluster patterns in the Arabic Najdi dialect*. MA thesis. Southern Illinois University at Carbondale.

Allen, J.S. and Miller, J.L. (1999) 'Effects of syllable-initial voicing and speaking rate on the temporal characteristics of monosyllabic words', *Journal of the Acoustical Society of America*, 106(4), pp. 2031-2039.

Alsudairi, M. and Abusharaf, R.M. (2015) 'Migration in pre-oil Qatar: a sketch', *Studies in Ethnicity and Nationalism*, 15(3), pp. 511-521.

Ashby, M. and Maidment, J. (2005) *Introducing phonetic science*. Cambridge University Press.

Baer, T. (1975) *Investigation of phonation using excised larynxes*. Massachusetts Institute of Technology.

Bang, H.-Y., Sonderegger, M., Kang, Y., Clayards, M. and Yoon, T.-J. (2018) 'The emergence, progress, and impact of sound change in progress in Seoul Korean: Implications for mechanisms of tonogenesis', *Journal of Phonetics*, 66, pp. 120-144.

Barr, D.J., Levy, R., Scheepers, C. and Tily, H.J. (2013) 'Random effects structure for confirmatory hypothesis testing: Keep it maximal', *Journal of memory and language*, 68(3), pp. 255-278.

Barry, M. and Teifour, R. (1999) *14th International Congress of Phonetic Sciences*, 3, p. 2429-2432

Bates, D., Mächler, M., Bolker, B. and Walker, S. (2014) 'Fitting linear mixed-effects models using lme4', *arXiv preprint arXiv:1406.5823*.

Beckman, J., Helgason, P., McMurray, B. and Ringen, C. (2011) 'Rate effects on Swedish VOT: Evidence for phonological overspecification', *Journal of Phonetics*, 39(1), pp. 39-49.

Beckman, J., Jessen, M. and Ringen, C. (2009) 'German fricatives: coda devoicing or positional faithfulness', *Phonology*, 26(2), pp. 231-268.

Beckman, J., Jessen, M. and Ringen, C. (2013) 'Empirical evidence for laryngeal features: Aspirating vs. true voice languages', *Journal of Linguistics*, 49(2), pp. 259-284.

Bellem, A. (2014). 'Triads, emphatics and interdentals in Arabic sound system typology'. *Journal of Semitic studies*, 34, pp. 9-41.

Best, C.T. and Halle, P.A. (2010) 'Perception of initial obstruent voicing is influenced by gestural organization', *Journal of Phonetics*, 38(1), pp. 109-126.

Blevins, J. (2004) *Evolutionary phonology: The emergence of sound patterns*. Cambridge University Press.

Boersma, P. and Weenink, D. (2016). Praat: Doing Phonetics by Computer. Version 5.3.84. http://www.praat.org/

Brown, J. (2016) 'Laryngeal assimilation, markedness and typology', *Phonology*, 33(3), pp. 393-423.

Burton, M.W. and Robblee, K.E. (1997) 'A phonetic analysis of voicing assimilation in Russian', *Journal of phonetics*, 25(2), pp. 97-114.

Caramazza, A., Yeni-Komshian, G.H., Zurif, E.B. and Carbone, E. (1973) 'The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals', *The Journal of the Acoustical Society of America*, 54(2), pp. 421-428.

Charles-Luce, J. (1985) 'Word-final devoicing in German: Effects of phonetic and sentential contexts', *Journal of Phonetics*, 13(3), pp. 309-324.

Chasaide, A.N. and Gobl, C. (1993) 'Contextual variation of the vowel voice source as a function of adjacent consonants', *Language and Speech*, 36(2-3), pp. 303-330.

Chen, M. (1970) 'Vowel length variation as a function of the voicing of the consonant environment', *Phonetica*, 22(3), pp. 129-159.

Chen, Y. (2011) 'How does phonology guide phonetics in segment–f0 interaction?', *Journal of Phonetics*, 39(4), pp. 612-625.

Cho, T. and Ladefoged, P. (1999) 'Variation and universals in VOT: evidence from 18 languages', *Journal of phonetics*, 27(2), pp. 207-229.

Chodroff, E. and Wilson, C. (2014) 'Burst spectrum as a cue for the stop voicing contrast in American English', *Journal of the Acoustical Society of America*, 136(5), pp. 2762-2772.

Chomsky, N. and Halle, M. (1968) 'The sound pattern of English'. New York: Harper & Row.

Coetzee, A.W., Beddor, P.S., Shedden, K., Styler, W. and Wissing, D. (2018) 'Plosive voicing in Afrikaans: Differential cue weighting and tonogenesis', *Journal of Phonetics*, 66, pp. 185-216.

Cohn, A.C. (1990) *Phonetic and phonological rules of nasalization*. PhD thesis. University of California, Los Angeles.

Davidson, L. (2016a) 'Acoustic effects of phonetic conditions and laryngeal specification on phonation in English voiceless obstruents', *The Journal of the Acoustical Society of America*, 140(4), pp. 3223-3224.

Davidson, L. (2016b) 'Variability in the implementation of voicing in American English obstruents', *Journal of Phonetics*, 54, pp. 35-50.

De Jong, K. and Zawaydeh, B. (2002) 'Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration', *Journal of Phonetics*, 30(1), pp. 53-75.

Deterding, D. and Nolan, F. (2007) Aspiration and voicing of Chinese and English plosives. In *Proceedings of the 16th international congress of phonetic Sciences*, p. 385-388. Citeseer.

Diehl, R.L. and Molis, M.R. (1995) 'Effect of Fundamental Frequency on Medial [+ Voice]/[– Voice] Judgments', *Phonetica*, 52(3), pp. 188-195.

Dmitrieva, O., Llanos, F., Shultz, A.A. and Francis, A.L. (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, pp. 77-95.

Docherty, G.J. (1992) *An experimental phonetic study of the timing of voicing in English obstruents*. Berlin, New York: Foris Publications.

Esposito, A. (2002) 'On vowel height and consonantal voicing effects: data from Italian', *Phonetica*, 59(4), pp. 197-231.

Feizollahi, Z. (2010) *Two case studies in the phonetics-phonology interface: evidence from Turkish voicing and Norwegian coalescence*. PhD thesis. Georgetown University.

Fischer-Jørgensen, E. (1968) 'Perceptual dimensions of vowels', *STUF-Language Typology and Universals*, 21(1-6), pp. 94-98.

Flege, J.E. (1982) 'Laryngeal timing and phonation onset in utterance-initial English stops', *Journal of Phonetics*, 10(2), pp. 177-192.

Flege, J.E. and Hillenbrand, J. (1987) 'A differential effect of release bursts on the stop voicing judgments of native French and English listeners', *Journal of Phonetics*, 15(2), pp. 203-208.

Flege, J.E. and Port, R. (1981) 'Cross-language phonetic interference: Arabic to English', *Language and speech*, 24(2), pp. 125-146.

Gósy, M. (2001) 'The VOT of the Hungarian voiceless plosives in words and in spontaneous speech', *International Journal of Speech Technology*, 4(1), pp. 75-85.

Greisbach, R. (2001) Experimentelle Testmethodik in Phonetik und Phonologie: Untersuchungen zu segmentalen grenzphänomenen im deutschen. Frankfurt am Main: Lang.

Haggard, M., Ambler, S. and Callow, M. (1970) 'Pitch as a voicing cue', *The Journal of the Acoustical Society of America*, 47(2B), pp. 613-617.

Hall, T.A. (2001) *Distinctive feature theory*. Berlin: Walter de Gruyter.

Halle, M., Hughes, G.W. and Radley, J.P. (1957) 'Acoustic properties of stop consonants', *The Journal of the Acoustical Society of America*, 29(1), pp. 107-116.

Halle, M. and Stevens, K.N. (1971) 'A Note on Laryngeal Features', *Phonology and Phonetics*, p. 45.

Harris, J. (1994) *English sound structure*. Blackwell Oxford.

Hayward, K. (2014) *Experimental phonetics: An introduction*. Routledge.

Helgason, P. and Ringen, C. (2008) 'Voicing and aspiration in Swedish stops', *Journal of phonetics*, 36(4), pp. 607-628.

Henderson, E.J. (1965) 'The topography of certain phonetic and morphological characteristics of South East Asian languages', *Lingua*, 15, pp. 400-434.

Heselwood, B. and Maghrabi, R. (2015). 'An Instrumental-Phonetic Justification for Sībawayh's Classification of ṭā', qāf and hamza as majhūr Consonants'. *Journal of Semitic studies*, 60(1), pp.131-175.

Hombert, J.-M., Ohala, J.J. and Ewan, W.G. (1979) 'Phonetic explanations for the development of tones', *Language*, pp. 37-58.

Honeybone, P. (2005) 'Diachronic evidence in segmental phonology: the case of obstruent laryngeal specifications', *The Internal Organization of Phonological Segments*, 77, p. 317.

Honeybone, P.G. (2002) *Germanic obstruent lenition: Some mutual implications of theoretical and historical phonology*. Newcastle University.

House, A.S. and Fairbanks, G. (1953) 'The influence of consonant environment upon the secondary acoustical characteristics of vowels', *The Journal of the Acoustical Society of America*, 25(1), pp. 105-113.

Hutters, B. (1985) 'Vocal fold adjustments in aspirated and unaspirated stops in Danish', *Phonetica*, 42(1), pp. 1-24.

Ingham, B. (1994) *Najdi Arabic : central Arabian*. Amsterdam ; Philadelphia: J. Benjamins Pub. Co.

Iverson, G.K. and Salmons, J.C. (1995) 'Aspiration and laryngeal representation in Germanic', *Phonology*, 12(3), pp. 369-396.

Iverson, G.K. and Salmons, J.C. (1999) 'Glottal spreading bias in Germanic', *Linguistische Berichte*, pp. 135-151.

Iverson, G.K. and Salmons, J.C. (2003) 'Legacy specification in the laryngeal phonology of Dutch', *Journal of Germanic linguistics*, 15(1), pp. 1-26.

Iverson, G.K. and Salmons, J.C. (2006) 'On the typology of final laryngeal neutralization: Evolutionary Phonology and laryngeal realism'.

Iverson, G.K. and Salmons, J.C. (2007) 'Domains and directionality in the evolution of German final fortition', *Phonology*, 24(1), pp. 121-145.

Iverson, G.K. and Salmons, J.C. (2011) 'Final devoicing and final laryngeal neutralization', *The Blackwell companion to phonology*, pp. 1-22.

Iwata, R. and Hirose, H. (1976) 'Fiberoptic acoustic studies of Mandarin stops and affricates', *Annual Bulletin Research Institute of Logopedics and Phoniatrics*, 10, pp. 47-60.

Jacques, B. (1984) *Étude acoustique des consonnes du français parlé à Montréal: analyse spectrale, intensité, durée*.

Jakobson, R. (1939) 'Zero sign', *Russian and Slavic grammar studies 1931–1981*, pp. 151-60.

Jakobson, R., Fant, C.G. and Halle, M. (1952) 'Preliminaries to speech analysis: The distinctive features and their correlates'. Cambridge: MIT Press.

Jakobson, R. and Waugh, L. (1979) 'An Instance of Interconnection between the Distinctive Features', in *Contributions to Comparative Mythology*. De Gruyter Mouton, pp. 59-61.

Jakobson, R., Waugh, L.R. and Taylor, M. (1987) *La forma sonora de la lengua*. Fondo de cultura económica México.

Jansen, W. (2004) *Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch*. Citeseer.

Jansen, W. (2007) 'Phonological ‘voicing’, phonetic voicing, and assimilation in English', *Language Sciences*, 29(2-3), pp. 270-293.

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

Jessen, M. (1998) *Phonetics and Phonology of Tense and Lax Obstruents in German*. John Benjamins Publishing.

Jessen, M. (2001) 'Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants', in *Distinctive feature theory*. De Gruyter Mouton, pp. 237-294.

Jessen, M. and Ringen, C. (2002) 'Laryngeal features in German', *Phonology*, 19(2), pp. 189-218.

Jessen, M. and Roux, J.C. (2002) 'Voice quality differences associated with stops and clicks in Xhosa', *Journal of Phonetics*, 30(1), pp. 1-52.

Johnstone, T.M. (1967) *Eastern Arabian dialect studies*. Oxford UP.

Kabrah, R. (2011) 'Regressive voicing assimilation in Cairene Arabic', in Broselow, E. and Ouali, H. (eds.) Perspectives on Arabic linguistics: papers from the annual symposia on Arabic linguistics Amsterdam: John Benjamins, pp. 21-33.

Kahn, D. (1976) Syllable-based Generalizations in English Phonology, PhD thesis. Indiana University Linguistics Club.

Kahn, D. (2015) *Syllable-based generalizations in English phonology*. Routledge.

Katz, D. (1987) *Grammar of the Yiddish language*. Bloomsbury Academic.

Keating, P.A. (1984) 'Phonetic and phonological representation of stop consonant voicing', *Language*, pp. 286-319.

Keating, P.A. (1988a) *A survey of phonological features*. Indiana University Linguistics Club.

Keating, P.A. (1988b) 'Underspecification in phonetics', *Phonology*, 5, pp. 275-292.

Keating, P.A. (1990) 'Phonetic representations in a generative grammar', *Journal of phonetics*, 18(3), pp. 321-334.

Kessinger, R.H. and Blumstein, S.E. (1997) 'Effects of speaking rate on voice-onset time in Thai, French, and English', *Journal of phonetics*, 25(2), pp. 143-168.

Khattab, G. (2002) 'VOT production in English and Arabic bilingual and monolingual children', *AMSTERDAM STUDIES IN THE THEORY AND HISTORY OF LINGUISTIC SCIENCE SERIES 4*, pp. 1-38.

Khattab, G., Al-Tamimi, F. and Heselwood, B. (2006). 'Acoustic and auditory differences in the/t/-/T/opposition in male and female speakers of Jordanian Arabic'. In Perspectives on Arabic Linguistics XVI: Papers from the sixteenth annual symposium on Arabic linguistics, pp. 131-160, John Benjamins Cambridge, UK.

Kim, C.-W. (1970) 'A theory of aspiration', *Phonetica*, 21(2), pp. 107-116.

Kingston, J. and Diehl, R.L. (1994) 'Phonetic knowledge', *Language*, 70(3), pp. 419-454.

Kingston, J. and Diehl, R.L. (1995) 'Intermediate properties in the perception of distinctive feature values', *Papers in laboratory phonology*, 4, pp. 7-27.

Kingston, J., Diehl, R.L., Kirk, C.J. and Castleman, W.A. (2008) 'On the internal perceptual structure of distinctive features: The [voice] contrast', *J Phon*, 36(1), pp. 28-54.

Kiparsky, P. (2003) 'Syllables and moras in Arabic', *The syllable in optimality theory*, pp. 147-182.

Kirby, J.P. (2010) *Cue selection and category restructuring in sound change*. The University of Chicago.

Kirby, J.P. (2018) 'Onset pitch perturbations and the cross-linguistic implementation of voicing: Evidence from tonal and non-tonal languages', *Journal of Phonetics*, 71, pp. 326-354.

Kirby, J.P. and Ladd, D.R. (2016) 'Effects of obstruent voicing on vowel F 0: Evidence from “true voicing” languages', *The Journal of the Acoustical Society of America*, 140(4), pp. 2400-2411.

Klatt, D.H. (1970) 'Synthesis of stop consonants in initial position', *The Journal of the Acoustical Society of America*, 47(1A), pp. 93-94.

Klatt, D.H. (1975) 'Voice onset time, frication, and aspiration in word-initial consonant clusters', *Journal of speech and hearing research*, 18(4), pp. 686-706.

Klatt, D.H. and Klatt, L.C. (1990) 'Analysis, synthesis, and perception of voice quality variations among female and male talkers', *the Journal of the Acoustical Society of America*, 87(2), pp. 820-857.

Kluender, K.R., Diehl, R.L. and Wright, B.A. (1988) 'Vowel-length differences before voiced and voiceless consonants: An auditory explanation', *Journal of phonetics*, 16(2), pp. 153-169.

Kohler, K.J. (1984) 'Phonetic explanation in phonology: the feature fortis/lenis', *Phonetica*, 41(3), pp. 150-174.

Kuhn, M. and Johnson, K. (2013) *Applied predictive modeling*. Springer.

Kulikov, V. (2012) *Voicing and voice assimilation in Russian stops*. The University of Iowa.

Kulikov, V. (2019) 'Laryngeal Contrast in Qatari Arabic: Effect of Speaking Rate on Voice Onset Time', *Phonetica*, 77(3), pp. 163-185.

Kulikov, V. (2020) 'Laryngeal Contrast in Qatari Arabic: Effect of Speaking Rate on Voice Onset Time', *Phonetica*, 77(3), pp. 163-185.

Ladd, D.R. and Schmid, S. (2018) 'Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops', *Journal of Phonetics*, 71, pp. 229-248.

Ladefoged, P. (2006) 'Representing linguistic phonetic structure', *unfinished MS, http://www. linguistics. ucla. edu/people/ladefoge/PhoneticStructure.*

Lavoie, L.M. (2001) *Consonant strength: Phonological patterns and phonetic manifestations*. Routledge.

Lehiste, I. (1970) Temporal organization of spoken language, In Form and Substance: Phonetic and Linguistic Papers Presented to Eli Fischer-Jorgensen. Copenhagen: Akademisk Foriag, pp. 159-169.

Lehiste, I. and Peterson, G.E. (1961) 'Some basic considerations in the analysis of intonation', *The Journal of the Acoustical Society of America*, 33(4), pp. 419-425.

Lehn, W. (1967) 'Vowel contrasts in Najdi Arabic', *Linguistics Studies in Memory of Richard Salade Harell, Graham Stuart*, pp. 123-31.

Lisker, L. (1957) 'Closure duration and the intervocalic voiced-voiceless distinction in English', *Language*, 33(1), pp. 42-49.

Lisker, L. (1986) '“Voicing” in English: A catalogue of acoustic features signaling/b/versus/p/in trochees', *Language and speech*, 29(1), pp. 3-11.

Lisker, L. and Abramson, A.S. (1964) 'A cross-language study of voicing in initial stops: Acoustical measurements', *Word*, 20(3), pp. 384-422.

Lombardi, L. (1991) *Laryngeal features and laryngeal neutralization*. University of Massachusetts Amherst.

Lousada, M., Jesus, L.M. and Hall, A. (2010a) 'Temporal acoustic correlates of the voicing contrast in European Portuguese stops', *Journal of the International Phonetic Association*, 40(3), pp. 261-275.

Lousada, M., Jesus, L.M.T. and Hall, A. (2010b) 'Temporal acoustic correlates of the voicing contrast in European Portuguese stops', *Journal of the International Phonetic Association*, 40(3), pp. 261-275.

Luce, P.A. and Charles-Luce, J. (1985) 'Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production', *The Journal of the Acoustical Society of America*, 78(6), pp. 1949-1957.

Mabrouk, F. (1981) *A linguistic study of gulf phonology: An articulatory and acoustic investigation of contiguous Kuwaiti stops and vowels*. University of Exeter.

Mack, M. (1982) 'Voicing-dependent vowel duration in English and French: Monolingual and bilingual production', *The Journal of the Acoustical Society of America*, 71(1), pp. 173-178.

Magloire, J. and Green, K.P. (1999) 'A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish', *Phonetica*, 56(3-4), pp. 158-185.

Miller, J.L., Green, K.P. and Reeves, A. (1986) 'Speaking rate and segments: A look at the relation between speech production and speech perception for the voicing contrast', *Phonetica*, 43(1-3), pp. 106-115.

Misnadin, M., Kirby, J.P. and Remijsen, B. (2015) Temporal and spectral properties of Madurese stops, in *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK.

Mitleb, F. (2001) 'Voice onset time of Jordanian Arabic stops', *Journal of the Acoustical Society of America*, 109(5), p. 2474.

Mitleb, F.M. (1984) 'Voicing effect on vowel duration is not an absolute universal', *Journal of Phonetics*, 12(1), pp. 23-27.

Munro, M.J. (1993) 'Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings', *Language and Speech*, 36(1), pp. 39-66.

Nagao, K. and de Jong, K. (2007) 'Perceptual rate normalization in naturally produced rate-varied speech', *The Journal of the Acoustical Society of America*, 121(5), pp. 2882-2898.

Nearey, T.M. and Rochet, B.L. (1994) 'Effects of place of articulation and vowel context on VOT production and perception for French and English stops', *Journal of the International Phonetic Association*, 24(1), pp. 1-18.

Ohala, J. J. (1981). Articulatory constraints on the cognitive representation of speech. In T. Myers, J. Laver & J. Anderson (Eds.), The cognitive representation of speech (pp. 111-127). Amsterdam: North-Holland Publishing Company.

Ohala, J.J. (1983) 'The origin of sound patterns in vocal tract constraints', in *The production of speech*. Springer, pp. 189-216.

Ohala, J.J. (1990) 'There is no interface between phonology and phonetics: a personal view', *Journal of phonetics*, 18(2), pp. 153-171.

Ohala, J.J. (1997). Aerodynamics of phonology. In *Proceedings of the Seoul International Conference on Linguistics*, 92, p. 97).

Ohala, J.J. (1999) ‘Phonetics in phonology’ *Linguistics in the Morning Calm 4. Seoul: The Linguistic Society of Korea*, pp. 105-113.

Ohala, J. J. (2011). Accommodation to the aerodynamic voicing constraint and its phonological relevance. Proceedings of the 17th International Congress of Phonetic Sciences (pp. 64-67). Hong Kong.

Ohde, R.N. (1984) 'Fundamental frequency as an acoustic correlate of stop consonant voicing', *The Journal of the Acoustical Society of America*, 75(1), pp. 224-230.

Perkell, J.S. (1969) 'Physiology of speech production: Results and implication of quantitative cineradiographic study', *Monograph*, 53.

Petrova, O., Plapp, R., Ringen, C. and Szentgyörgyi, S. (2006) 'Voice and aspiration: Evidence from Russian, Hungarian, German, Swedish, and Turkish'.

Pierrehumbert, J. and Beckman, M. (1988) 'Japanese tone structure', *Linguistic inquiry monographs*, (15), pp. 1-282.

Piroth, H.G. and Janker, P.M. (2004) 'Speaker-dependent differences in voicing and devoicing of German obstruents', *Journal of Phonetics*, 32(1), pp. 81-109.

Port, R.F. and Rotunno, R. (1979) 'Relation between voice-onset time and vowel duration', *The Journal of the Acoustical Society of America*, 66(3), pp. 654-662.

Prochazka, T. (1988) 'The Spoken Arabic of Abū Thōr in al-Ḥasa', *Zeitschrift für arabische Linguistik*, (18), pp. 59-76.

Radwan, M. (1996) *An experimental investigation of the acoustical temporal correlates of voicing contrast in stop consonants (with reference to Arabic)*. PhD thesis. University of Essex.

R Core Team (2014). ‘R: A Language and Environment for Statistical Computing’. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ringen, C. and Helgason, P. (2004) 'Distinctive [voice] does not imply regressive assimilation: evidence from Swedish', *International Journal of English Studies*, 4(2), pp. 53-71.

Ringen, C. and Kulikov, V. (2012) 'Voicing in Russian stops: Cross-linguistic implications', *Journal of Slavic Linguistics*, pp. 269-286.

Ringen, C. and van Dommelen, W.A. (2013) 'Quantity and laryngeal contrasts in Norwegian', *Journal of Phonetics*, 41(6), pp. 479-490.

Ringo, C.C. (1988) 'Enhanced amplitude of the first harmonic as a correlate of voicelessness in aspirated consonants', *The Journal of the Acoustical Society of America*, 83(S1), pp. S70-S70.

Roettger, T.B. (2019) 'Researcher degrees of freedom in phonetic research', *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 10(1).

Rojczyk, A. (2009) 'The voicing contrast in Polish', *슬라브어 연구*, 14(2), pp. 1-12.

Schwarz, M., Sonderegger, M. and Goad, H. (2019) 'Realization and representation of Nepali laryngeal contrasts: Voiced aspirates and laryngeal realism', *Journal of Phonetics*, 73, pp. 113-127.

Slis, I.H. and Cohen, A. (1969) 'On the complex regulating the voiced-voiceless distinction I', *Language and speech*, 12(2), pp. 80-102.

Smith, B.L. (1978) 'Temporal aspects of English speech production: A developmental perspective', *Journal of Phonetics*, 6(1), pp. 37-67.

Solé, M.-J. (2011) 'Voice-initiating gestures in Spanish: Prenasalization', *UC Berkeley PhonLab Annual Report*, 7(7).

Solé, M.J. and Estebas, E. (2000). Phonetic and phonological phenomena: VOT: A cross-language comparison. In *Proceedings of the 18th AEDEAN Conference*, pp. 437-44.

Sole, M.J. (2018) 'Articulatory adjustments in initial voiced stops in Spanish, French and English', *Journal of Phonetics*, 66, pp. 217-241.

Stevens, K.N. (2000) *Acoustic phonetics*. MIT press: Cambridge.

Stevens, K.N. and Hanson, H.M. (1995) 'Classification of glottal vibration from acoustic measurements', *Vocal fold physiology: Voice quality control*, pp. 147-170.

Suomi, K. (1980) *Voicing in English and Finnish stops: a typological comparision with an interlanguage study of the two languages in contact*. University of Turku.

Sweel, A.I.A. (1990) 'Some aspects of Najdi Arabic phonology', *Zeitschrift für Arabische Linguistik*, (21), pp. 71-82.

Theodore, R.M., Miller, J.L. and DeSteno, D. (2009) 'Individual talker differences in voice-onset-time: Contextual influences', *The Journal of the Acoustical Society of America*, 125(6), pp. 3974-3982.

Turk, A. and Nakai, S. (2006) 'Acoustic Segment Durations in Prosodic Research: A Practical Guide', *Methods in Empirical Prosody Research*, 3, p. 1.

Unal-Logacev, O., Fuchs, S. and Lancia, L. (2018) 'A multimodal approach to the voicing contrast in Turkish: Evidence from simultaneous measures of acoustics, intraoral pressure and tongue palatal contacts', *Journal of Phonetics*, 71, pp. 395-409.

Van Alphen, P.M. and Smits, R. (2004) 'Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing', *Journal of phonetics*, 32(4), pp. 455-491.

Van den Berg, J. (1958) 'Myoelastic-aerodynamic theory of voice production', *Journal of speech and hearing research*, 1(3), pp. 227-244.

Van Rooy, B. and Wissing, D. (1998) 'On the relation between voicing, aspiration and tenseness in stop classification 1', *South African Journal of Linguistics*, 16(sup36), pp. 101-124.

Vaux, B. and Samuels, B. (2005) 'Laryngeal markedness and aspiration', *Phonology*, 22(3), pp. 395-436.

Volaitis, L.E. and Miller, J.L. (1992) 'Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories', *The Journal of the Acoustical Society of America*, 92(2), pp. 723-735.

Watson, J.C.E. (2002) *The phonology and morphology of Arabic*. Oxford ; New York: Oxford University Press.

Weismer, G. (1979) 'Sensitivity of voice-onset time (VOT) measures to certain segmental features in speech production', *Journal of Phonetics*, 7(2), pp. 197-204.

Wells, J.C. (1982) *Accents of English: Volume 3: Beyond the British Isles*. Cambridge University Press.

Westbury, J.R. (1983) 'Enlargement of the supraglottal cavity and its relation to stop consonant voicing', *J Acoust Soc Am*, 73(4), pp. 1322-36.

Westbury, J.R. and Keating, P.A. (1986) 'On the naturalness of stop consonant voicing', *Journal of linguistics*, 22(1), pp. 145-166.

Wetzels, W.L. and Mascaró, J. (2001) 'The typology of voicing and devoicing', *Language*, pp. 207-244.

Williams, L. (1977) 'The perception of stop consonant voicing by Spanish-English bilinguals', *Perception and Psychophysics*, 21(4), pp. 289-297.

Wissing, D. and Roux, J. (1995) *Proceedings of the 13th International Conference of the Phonetic Sciences*.

Yao, Y. (2009) 'Understanding VOT variation in spontaneous speech', *UC Berkeley PhonLab Annual Report*, 5(5).

Yeni-Komshian, G.H., Caramazza, A. and Preston, M.S. (1977) 'A study of voicing in Lebanese Arabic', *Journal of Phonetics*, 5(1), pp. 35-48.

Zawaydeh, B.A. and de Jong, K. (2011). 'The phonetics of localising uvularisation in Ammani-Jordanian Arabic'. *Instrumental studies in Arabic phonetics*, 319, p.257.

Zee, E. (1980) 'The effect of aspiration on the F0 of the following vowel in Cantonese', *UCLA Working Papers in Phonetics*, 49, pp. 90-97.

Zue, V.W. (1976) *Acoustic characteristics of stop consonants: A controlled study*. PhD thesis. Massachusetts Institute of Technology.

Appendix A

Demographic Survey

1. General information:

Item	Answer
➢ What is your gender?	___ Male ___ Female
➢ What is your age?	___19-25 ___26-35 ___35-45
➢ What is your education level?	
➢ Were you born and raised in Riyadh?	___ Yes ___ No
➢ Do you have any difficulty speaking or hearing?	___ Yes ___ No
➢ Do you speak any language(s) beside Arabic? Specify.	

2. Dialect information:

Item	Answer
➢ What is your native dialect?	
➢ Do you speak any other dialect beside your native one?	
➢ Are you originally from Najd? Which area?	
➢ What is your parents' native dialect?	
➢ Do your parents speak any other dialect beside their native one?	
➢ Are your parents originally from Najd? Which area?	

Appendix B

The list of words used in the analysis:

	word	transcription	gloss		word	transcription	gloss
1	بات	/ba:t/	slept	48	فوق	/fo:g/	up
2	بيت	/be:t/	house	49	سوق	/su:g/	market
3	بون	/bo:n/	difference	50	بقيق	/bge:g/	city
4	بوك	/bu:k/	your dad	51	تابت	/ta:bat/	she repented
5	بيق	/bi:g/	stolen	52	دوبك	/du:bak/	you just now
6	تاب	/ta:b/	repented	53	غيبه	/ɣi:bah/	backbiting
7	توت	/tu:t/	blueberry	54	فادت	/fa:dat/	benefited
8	تين	/ti:n/	figs	55	دوده	/du:dah/	worm
9	توق	/to:g/	longing	56	سيده	/si:dah/	straight
10	تيس	/te:s/	goat	57	ساقت	/sa:gat/	she drove
11	داس	/da:s/	stepped	58	سوقك	/su:gak/	your market
12	دوم	/do:m/	always	59	بيقت	/bi:gat/	stolen
13	دين	/de:n/	debt	60	ماتت	/ma:tat/	she died
14	دود	/du:d/	worms	61	حبى	/ħaba/	crawled
15	ديك	/di:k/	rooster	62	ربى	/ruba/	Name (F)
16	كال	/ka:l/	weigh	63	هبه	/hiba/	you want
17	كوم	/ko:m/	group of	64	ندى	/nada/	Name (F)
18	كيس	/ki:s/	bag	65	هدى	/huda/	Name (F)
19	كيف	/ke:f/	how	66	فدى	/fida/	no worries
20	كود	/ku:d/	almost	67	سقى	/saga/	water (v)
21	قوم	/go:m/	tribe	68	بقع	/bugaʕ/	spots
22	قام	/ga:m/	woke up	69	بقى	/biga/	remained
23	قوت	/gu:t/	food	70	متى	/mata/	when
24	قيل	/gi:l/	said	71	غتر	/ɣutar/	shemaghs
25	قيد	/ge:d/	chain	72	فتن	/fitan/	difficulties
26	باب	/ba:b/	door	73	بكى	/baka/	remained
27	سيب	/si:b/	hall	74	بكا	/buka/	behind
28	شيب	/ʃe:b/	white hair	75	سكت	/sikat/	hushed
29	ثوب	/θo:b/	dress	76	قوتك	/gu:tak/	your food
30	دوب	/du:b/	just now	77	جيتك	/dʒi:tak/	I came
31	فات	/fa:t/	passed	78	شاكت	/ʃa:kat/	she sued
32	شوت	/ʃo:t/	kicking	79	حاكت	/ħa:kat/	she knitted
33	بيت	/be:t/	house	80	حيكت	/ħi:kat/	was knitted
34	جيت	/dʒ i:t/	I came	81	شباب	/ʃaba:b/	youth
35	توت	/tu:t/	blueberry	82	تبوك	/tabu:k/	city
36	فاد	/fa:d/	benefitted	83	تبيه	/tabi:h/	you want it
37	فود	/fo:d/	benefit	84	سداد	/sada:d/	pay
38	كيد	/ke:d/	conspiracy	85	سدود	/sadu:d/	dams
39	كيد	/ki:d/	was tricked	86	جديد	/dʒadi:d/	new
40	دود	/du:d/	worms	87	لقاك	/laga:k/	he met you
41	جاك	/dʒa:k/	came	88	نقوم	/tagu:m/	stand
42	فيك	/fi:k/	in you	89	ثقيل	/θagi:l/	heavy
				90	مكين	/maki:n/	robust
43	بوك	/bu:k/	your dad	91	شتات	/ʃata:t/	chaos
44	شوك	/ʃo:k/	thorns	92	كتوم	/katu:m/	secretive
45	بريك	/bre:k/	name	93	متين	/mati:n/	fat
46	ذاق	/ða:g/	tasted	94	مكان	/maka:n/	place
47	بيق	/bi:g/	stolen	95	شكوك	/ʃaku:k/	doubts

Appendix C

Tables for the statistical output of LMM models and the pairwise tests for the acoustic correlates in initial, medial, final, and stop-stop clusters.

1. Aspiration in utterance-initial stops

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	23.23	1.27	18.31	**<0.001**
sex_c	1.17	1.39	0.85	0.398
place_c	6.80	1.58	4.30	**<0.001**
rate_c	-6.60	0.35	-19.07	**<0.001**
vowel_c	-6.59	2.72	-2.42	**0.015**
voicing_c	-34.54	1.49	-23.23	**<0.001**

Table C-1. LMM results for aspiration in utterance-initial stops as a function of voicing (Voiceless-Voiced), place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	Vl FA – vd FA	<0.001

Table C-2. Pairwise comparison output for aspiration in utterance-initial stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		
vd FA B – vd FA A	0.008		

Table C-3. Pairwise comparison output for aspiration in utterance-initial stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar)

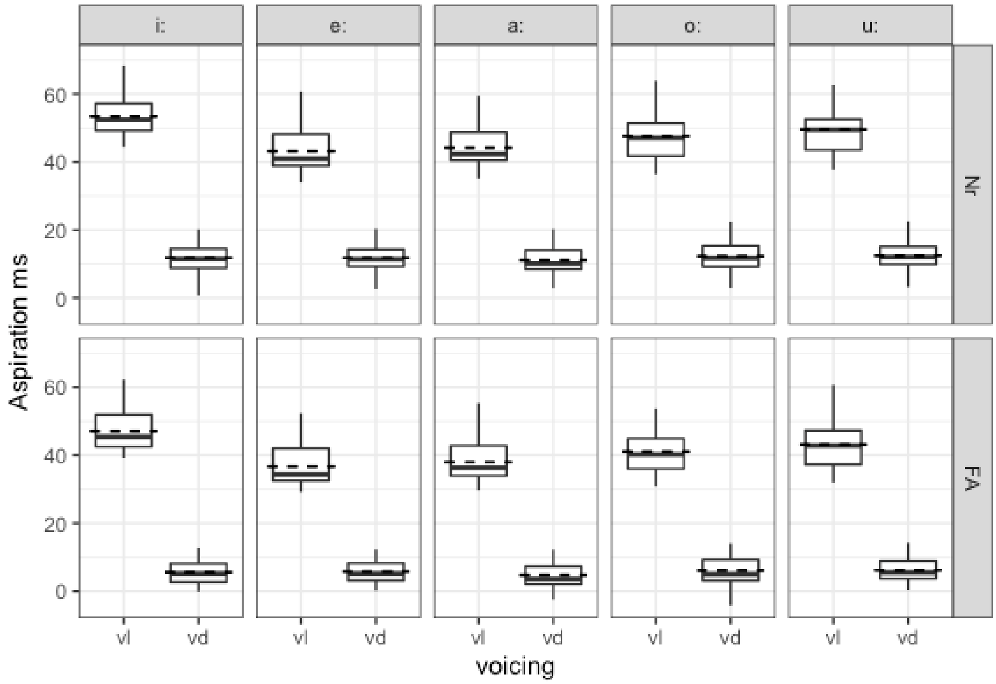


Figure C.1 Boxplots of aspiration in utterance-initial stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal) and vowel type. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
/i:/	53.4	5.68	11.8	4.02	47.1	5.89	5.52	3.32
/e:/	43.2	5.62	11.7	3.73	36.8	5.75	5.75	3.17
/a:/	44.2	5.51	11.1	3.72	38	5.97	4.73	3.73
/o:/	47.7	6.81	12.3	4.17	41.1	6.8	6.09	3.63
/u:/	49.6	7.07	12.5	3.84	43.1	7.07	6.2	3.43

Table C.4. Means and standard deviations of aspiration for Voiceless and Voiced utterance-initial stops grouped by speech rate and vowel type.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr i: – vl Nr e:	<0.001	vl Nr i: - vd Nr i:	<0.001
vl Nr i: – vl Nr a:	<0.001	vl Nr e: - vd Nr e:	<0.001
vl Nr i: – vl Nr o:	<0.001	vl Nr a: - vd Nr a:	<0.001
vl Nr i: – vl Nr u:	<0.001	vl Nr o: - vd Nr o:	<0.001
vl Nr e: – vl Nr a:	0.0484	vl Nr u: - vd Nr u:	<0.001
vl Nr e: – vl Nr o:	<0.001	vl FA i: - vd FA i:	<0.001
vl Nr e: – vl Nr u:	<0.001	vl FA e: - vd FA e:	<0.001
vl Nr a: – vl Nr o:	<0.001	vl FA a: - vd FA a:	<0.001
vl Nr a: – vl Nr u:	<0.001	vl FA o: - vd FA o:	<0.001
vl Nr o: – vl Nr u:	<0.001	vl FA u: - vd FA u:	<0.001
vd Nr i: – vd Nr e:	0.966		
vd Nr i: – vd Nr a:	0.156		

vd Nr i: – vd Nr o:	0.271
vd Nr i: – vd Nr u:	0.136
vd Nr e: – vd Nr a:	0.163
vd Nr e: – vd Nr o:	0.25
vd Nr e: – vd Nr u:	0.123
vd Nr a: – vd Nr o:	0.01
vd Nr a: – vd Nr u:	0.003
vd Nr o: – vd Nr u:	0.681
vl FA i: – vl FA e:	<0.001
vl FA i: – vl FA a:	<0.001
vl FA i: – vl FA o:	<0.001
vl FA i: – vl FA u:	<0.001
vl FA e: – vl FA a:	0.02
vl FA e: – vl FA o:	<0.001
vl FA e: – vl FA u:	<0.001
vl FA a: – vl FA o:	<0.001
vl FA a: – vl FA u:	<0.001
vl FA o: – vl FA u:	<0.001
vd FA i: – vd FA e:	0.628
vd FA i: – vd FA a:	0.1
vd FA i: – vd FA o:	0.233
vd FA i: – vd FA u:	0.163
vd FA e: – vd FA a:	0.031
vd FA e: – vd FA o:	0.487
vd FA e: – vd FA u:	0.366
vd FA a: – vd FA o:	0.003
vd FA a: – vd FA u:	0.002
vd FA o: – vd FA u:	0.832

Table C-5. Pairwise comparison output for aspiration in utterance-initial stops based on voicing, rate, and vowel type.

2. Aspiration in utterance-medial stops (Trochaic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	18.58	0.77	24.22	**<0.001**
voicing_c	-22.44	1.31	-17.08	**<0.001**
vowel_c	0.34	1.16	0.29	0.769
rate_c	-3.79	0.28	-13.40	**<0.001**
Place_c	3.51	1.46	2.40	**0.017**
sex_c	0.61	1.08	0.57	0.571

Table C-6. LMM results for aspiration in utterance-medial stops (Trochaic) stops as a function of voicing (Voiceless-Voiced), place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	Vl FA – vd FA	<0.001

Table C-7. Pairwise comparison output for aspiration in utterance-medial stops (Trochaic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		
vd FA B – vd FA A	<0.001		

Table C-8. Pairwise comparison output for aspiration in utterance-medial stops (Trochaic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

3. Aspiration in utterance-medial stops (Iambic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	23.38	1.93	12.12	**<0.001**
voicing_c	-31.13	3.75	-8.29	**<0.001**
vowel_c	-7.96	4.22	-1.89	0.059
rate_c	-5.80	0.40	-14.60	**<0.001**
place_c	5.04	4.94	1.02	0.308
sex_c	-0.14	1.19	-0.12	0.906

Table C-9. LMM results for aspiration in utterance-medial stops (iambic) stops as a function of voicing (Voiceless-Voiced), place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	Vl FA – vd FA	<0.001

Table C-10. Pairwise comparison output for aspiration in utterance-medial stops (iambic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.0120	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.0419	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	0.0075	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	0.0256		
vd FA A – vd FA V	<0.001		
vd FA B – vd FA A	<0.001		

Table C-11. Pairwise comparison output for aspiration in utterance-medial stops (iambic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: vealr)

4. Aspiration in utterance-final stops

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	40.57	2.41	16.87	**<0.001**
voicing_c	-54.12	2.99	-18.12	**<0.001**
vowel_c	0.41	4.36	0.09	0.925
rate_c	-24.96	1.52	-16.42	**<0.001**
place_c	2.80	2.52	1.11	0.265
sex_c	7.00	3.23	2.17	**0.030**

Table C-12. LMM results for aspiration in utterance-final stops in the function of voicing (Voiceless-Voiced), place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	Vl FA – vd FA	<0.001

Table C-13. Pairwise comparison output for aspiration in utterance-final stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.0751	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.3213	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.007		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	0.037		
vd FA B – vd FA A	<0.001		

Table C-14. Pairwise comparison output for aspiration in utterance-final stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: vealr)

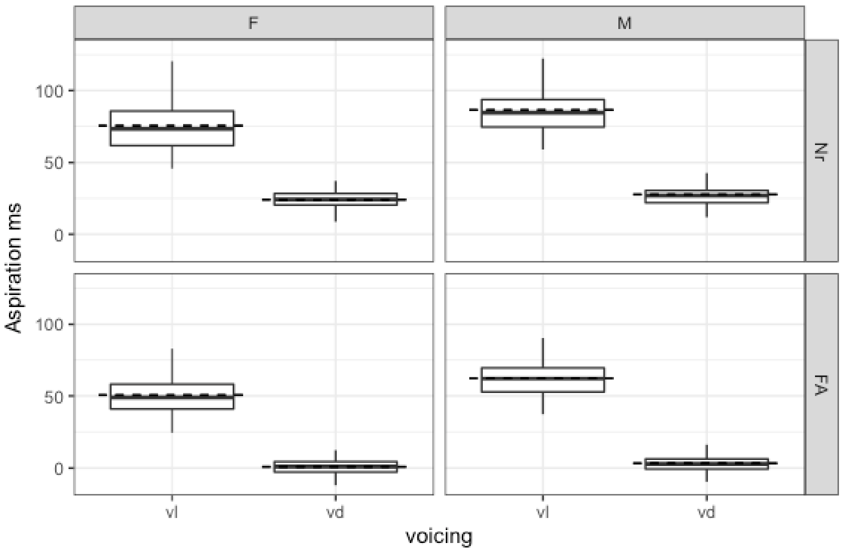


Figure C.2 Boxplots of aspiration in utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal), and gender (F = female, M = male). The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	57.4	17.3	24.2	5.65	50.6	13.5	0.86	5.01
male	86.6	16	27.7	8.1	62.3	12.5	3.27	5.45

Table C.15 Means and standard deviations of aspiration for Voiceless and Voiced utterance-final stops grouped by speech rate and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-16. Pairwise comparison output for aspiration in utterance-final stops based on voicing, rate, and gender.

5. % voicing in utterance-medial stops (trochaic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	51.08	0.40	128.60	**<0.001**
vowel_c	-0.83	0.60	-1.39	0.166
rate_c	0.23	0.30	0.77	0.441
place_c	-1.86	0.93	-2.00	**0.045**
voicing_c	96.21	0.76	127.29	**<0.001**
sex_c	0.67	0.62	1.09	0.277

Table C-17. LMM results for % voicing in utterance-medial stops (trochaic) as a function of voicing (Voiceless-Voiced), place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.083	vl Nr – vd Nr	<0.001
vd Nr – vd FA	0.060	Vl FA – vd FA	<0.001

Table C-18. Pairwise comparison output for % voicing in utterance-medial stops (trochaic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	0.0906
vd Nr B – vd Nr A	0.00109	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		
vd FA B – vd FA A	<0.001		

Table C-19. Pairwise comparison output for % voicing in utterance-medial stops (trochaic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: vealr)

6. % voicing in utterance-medial stops (iambic)

Predictors	*Estimates*	*std. Error*	*Statistic*	*p*
(Intercept)	50.45	0.28	183.29	**<0.001**
vowel_L_c	-0.56	0.50	-1.11	0.265
rate_c	0.43	0.32	1.35	0.176
place_c	-1.20	0.57	-2.10	**0.036**
voicing_c	98.14	0.46	212.98	**<0.001**
sex_c	0.71	0.57	1.25	0.211

Table C-20. LMM results for % voicing in utterance-medial stops (iambic) as a function of voicing (Voiceless-Voiced), place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	Vl FA – vd FA	<0.001

Table C-21. Pairwise comparison output for % voicing in utterance-medial stops (iambic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	0.119	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	0.0623		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-22. Pairwise comparison output for % voicing in utterance-medial stops (iambic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar)

7. % voicing in utterance-final stops

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	9.32	1.63	5.71	**<0.001**
vowel_c	2.36	1.91	1.23	0.217
rate_c	3.18	1.26	2.53	**0.011**
place_c	-0.76	2.06	-0.37	0.713
voicing_c	17.33	3.06	5.65	**<0.001**
sex_c	-4.79	3.01	-1.59	0.111

Table C-23. LMM results for % voicing in utterance-final stops as a function of voicing (Voiceless-Voiced), place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-24. Pairwise comparison output for % voicing in utterance-final stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.978	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.920	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	0.004	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	0.103	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.216		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		
vd FA B – vd FA A	<0.001		

Table C-25. Pairwise comparison output for % voicing in utterance-final stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

8. Prevoicing in utterance-initial stops

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	52.69	1.58	33.43	**<0.001**
vowel_c	-0.64	2.11	-0.30	0.761
rate_c	-24.59	1.13	-21.82	**<0.001**
place_c	-3.22	2.14	-1.50	0.132
sex_c	-3.33	2.84	-1.17	0.240

Table C-26. LMM results for prevoicing in utterance-initial Voiced stops as a function of place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Place		Rate	
Pair	p-value	Pair	p-value
Nr B – Nr A	0.11	Nr B – FA B	<0.001
Nr B – Nr V	<0.001	Nr A – FA A	<0.001
Nr A – Nr V	<0.001	Nr V – FA V	<0.001
FA B – FA A	0.12		
FA B – FA V	<0.001		
FA A – FA V	<0.001		

Table C-27. Pairwise comparison output for prevoicing in utterance-initial Voiced stops based on place (B: bilabial/A: alveolar/V: velar) and rate.

9. Voicing duration in utterance-medial stops (trochaic)

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	24.39	0.78	31.27	**<0.001**
vowel_c	-1.39	1.47	-0.95	0.343
rate_c	-5.22	0.38	-13.69	**<0.001**
place_c	-3.44	1.82	-1.89	0.059
voicing_c	44.91	1.45	31.03	**<0.001**
sex_c	-0.30	0.88	-0.34	0.731

Table C-28. LMM results for voicing duration in utterance-medial stops (trochaic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-29. Pairwise comparison output for voicing duration in utterance-medial stops (trochaic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		
vd FA B – vd FA A	<0.001		

Table C-30. Pairwise comparison output for voicing duration in utterance-medial stops (trochaic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar)

10. Voicing duration in utterance-medial stops (iambic).

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	24.28	1.14	21.22	**<0.001**
vowel_c	2.22	2.45	0.9	0.382
rate_c	-5.53	0.48	-11.41	**<0.001**
place_c	1.29	2.96	0.43	0.669
voicing_c	48.72	2.23	21.78	**<0.001**
sex_c	-0.24	0.83	-0.28	0.776

Table C-31. LMM results for voicing duration in utterance-medial stops (iambic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-32. Pairwise comparison output for voicing duration in utterance-medial stops (iambic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		
vd FA B – vd FA A	<0.001		

Table C-32. Pairwise comparison output for voicing duration in utterance-medial stops (iambic) based on voicing, rate, and place of articulation (B: bilabial/A: alveolar/V: velar).

11. Voicing duration in utterance-final stops

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	6.92	1.21	5.71	**<0.001**
vowel_c	1.28	1.21	1.06	0.288
rate_c	-0.56	0.89	-0.63	0.531
place_c	-1.11	1.43	-0.77	0.440
voicing_c	12.70	2.25	5.64	**<0.001**
sex_c	-3.32	2.30	-1.44	0.149

Table C-34. LMM results for voicing duration in utterance-medial stops (iambic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.3	vl Nr – vd Nr	<0.001
vd Nr – vd FA	0.006	vl FA – vd FA	<0.001

Table C-35. Pairwise comparison output for voicing duration in utterance-final stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.987	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.986	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	0.002	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.455		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	0.787		

Table C-36. Pairwise comparison output for voicing duration in utterance-final stops based on voicing, rate, and place of articulation (B: bilabial/A: alveolar/V: velar).

12. Closure duration for utterance-medial stops (trochaic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	49.31	1.10	44.94	**<0.001**
vowel_c	0.98	2.06	0.48	0.632
rate_c	-8.78	0.64	-13.78	**<0.001**
place_c	1.67	2.57	0.65	0.517
voicing_c	-3.77	1.97	-1.92	0.055
sex_c	-3.17	1.29	-2.45	**0.014**

Table C-37. LMM results for closure duration in utterance-medial stops (trochaic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-38. Pairwise comparison output for closure duration in utterance-medial stops (trochaic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-39. Pairwise comparison output for closure duration in utterance-medial stops (trochaic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar)

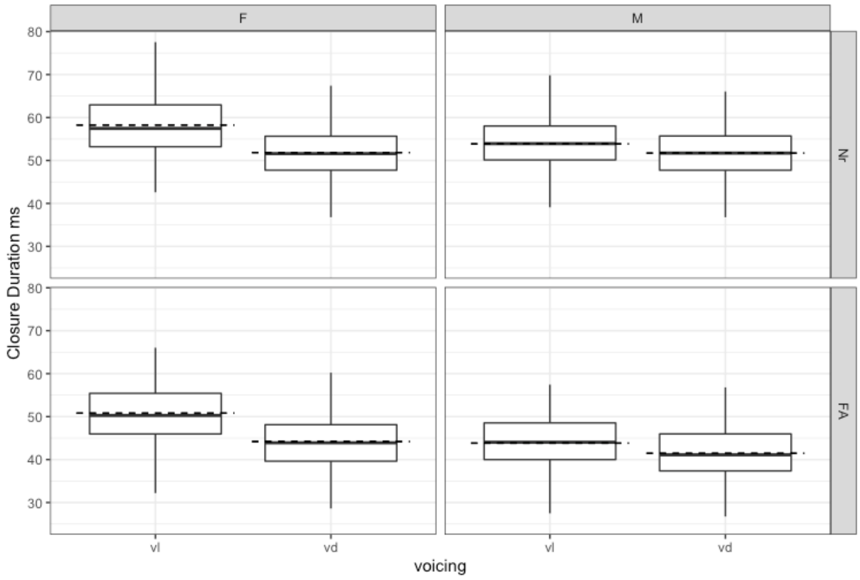


Figure C.3. Boxplots of the closure duration for utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate (FA = fast, Nr = Normal), and gender. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	58.2	7.04	51.8	5.97	50.8	7.61	44.2	6.63
male	53.9	6.04	51.7	5.62	43.8	6.03	41.5	6.01

Table C.41. Mean and standard deviation of closure duration for utterance-medial stops (Trochaic) grouped by speech rate and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	0.678	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-41. Pairwise comparison output for closure duration in utterance-medial stops (trochaic) based on voicing, rate, and gender.

13. Closure duration for utterance-medial stops (iambic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	50.90	1.54	33.02	**<0.001**
rate_c	-10.39	0.85	-12.19	**<0.001**
place_c	5.41	3.66	1.48	0.139
voicing_c	-3.28	2.64	-1.24	0.213
sex_c	-1.62	1.51	-1.07	0.285
vowel_c	1.74	2.50	0.70	0.486

Table C-42. LMM results for closure duration in utterance-medial stops (iambic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-43. Pairwise comparison output for closure duration in utterance-medial stops (iambic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-44. Pairwise comparison output for closure duration in utterance-medial stops (iambic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar)

14. Closure duration for utterance-final stops

Predictors	*Estimates*	*std. Error*	*Statistic*	*p*
(Intercept)	78.66	2.26	34.74	**<0.001**
vowel_c	1.40	3.74	0.37	0.709
rate_c	-25.54	1.52	-16.75	**<0.001**
place_c	-0.45	4.01	-0.11	0.911
voicing_c	9.83	3.05	3.22	**0.001**
sex_c	-1.31	3.61	-0.36	0.718

Table C-45. LMM results for closure duration in utterance-final stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-46. Pairwise comparison output for closure duration in utterance-final stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	0.642
vd Nr B – vd Nr A	0.00016	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	0.953
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	0.0194		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	0.0007		

Table C-47. Pairwise comparison output for closure duration in utterance-final stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar)

15. Burst duration for utterance-initial stops

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	2.54	0.17	14.54	**<0.001**
vowel_c_	0.29	0.20	1.49	0.137
rate_c	0.51	0.12	4.05	**<0.001**
place_c	0.71	0.21	3.35	**0.001**
voicing_c	-0.83	0.19	-4.31	**<0.001**
sex_c	-0.21	0.32	-0.66	0.507

Table C-48. LMM results for burst duration in utterance-initial stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-49. Pairwise comparison output for burst duration in utterance-initial stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-50. Pairwise comparison output for burst duration in utterance-initial stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

16. Burst duration for utterance-medial stops (trochaic)

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	2.43	0.13	18.61	**<0.001**
vowel_L_c	0.09	0.15	0.60	0.548
rate_c	0.46	0.12	3.96	**<0.001**
place_c	0.69	0.21	3.36	**0.001**
voicing_c	-1.03	0.15	-7.00	**<0.001**
sex_c	-0.31	0.23	-1.34	0.181

Table C-51. LMM results for burst duration in utterance-medial stops (trochaic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-52. Pairwise comparison output for burst duration in utterance-medial stops (trochaic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-53. Pairwise comparison output for burst duration in utterance-mediall stops (trochaic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

17. Burst duration for utterance-medial stops (iambic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	2.63	0.18	14.49	**<0.001**
vowel_c	-0.25	0.25	-0.99	0.320
rate_c	0.44	0.12	3.72	**<0.001**
place_c	0.92	0.32	2.93	**0.003**
voicing_c	-1.22	0.25	-4.81	**<0.001**
sex_c	-0.16	0.30	-0.53	0.597

Table C-54. LMM results for burst duration in utterance-medial stops (iambic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-55. Pairwise comparison output for burst duration in utterance-medial stops (iambic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-56. Pairwise comparison output for burst duration in utterance-medial stops (iambic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

18. Burst duration for utterance-final stops

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	3.74	0.21	18.03	**<0.001**
vowel_c	0.01	0.20	0.07	0.947
rate_c	-0.45	0.18	-2.54	**0.011**
place_c	0.65	0.20	3.18	**0.001**
voicing_c	-1.23	0.17	-7.05	**<0.001**
sex_c	0.04	0.39	0.10	0.918

Table C-57. LMM results for burst duration in utterance-final stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-58. Pairwise comparison output for burst duration in utterance-final stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	0.0018	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.005		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-59. Pairwise comparison output for burst duration in utterance-final stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

19. Following vowel duration for utterance-initial stops (FVD)

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	119.45	4.30	27.76	**<0.001**
vowel_c	4.62	8.45	0.55	0.584
rate_c	-35.89	2.92	-12.31	**<0.001**
place_c	0.26	9.08	0.03	0.977
voicing_c	6.63	6.66	1.00	0.319
sex_c	11.14	5.01	2.22	**0.026**

Table C-60. LMM results for FVD in utterance-initial stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-61. Pairwise comparison output for FVD in utterance-initial stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.451	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.385	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.714		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	0.687		

Table C-62. Pairwise comparison output for FVD in utterance-initial stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

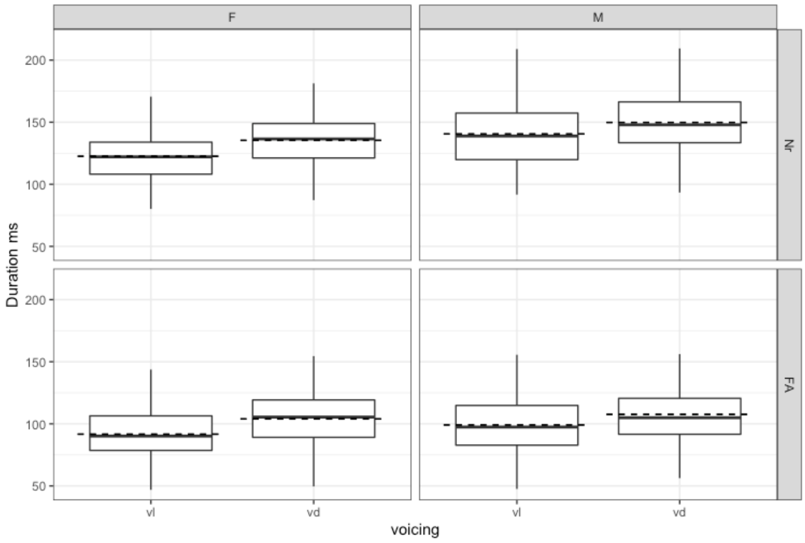


Figure C.4. Boxplots of FVD for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate and gender.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	123	18.1	135	19.5	91.7	19.9	104	21.3
male	141	25.3	150	25.1	99.1	23.1	108	23

Table C.63. Mean and standard deviation of FVD for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-64. Pairwise comparison output for FVD in utterance-initial stops based on voicing, rate, and gender.

20. Following vowel duration for utterance-medial stops (iambic)

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	106.91	5.54	19.29	**<0.001**
rate_c	-32.16	2.23	-14.44	**<0.001**
place_c	12.25	8.05	1.52	0.128
voicing_c	11.57	6.21	1.86	0.063
sex_c	7.09	4.75	1.49	0.136
vowel_c	-25.33	12.15	-2.08	**0.037**

Table C-65. LMM results for FVD in utterance-medial stops (iambic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-66. Pairwise comparison output for FVD in utterance-medial stops (iambic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.032	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.047	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.07		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	0.061		

Table C-67. Pairwise comparison output for FVD in utterance-medial stops (iambic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

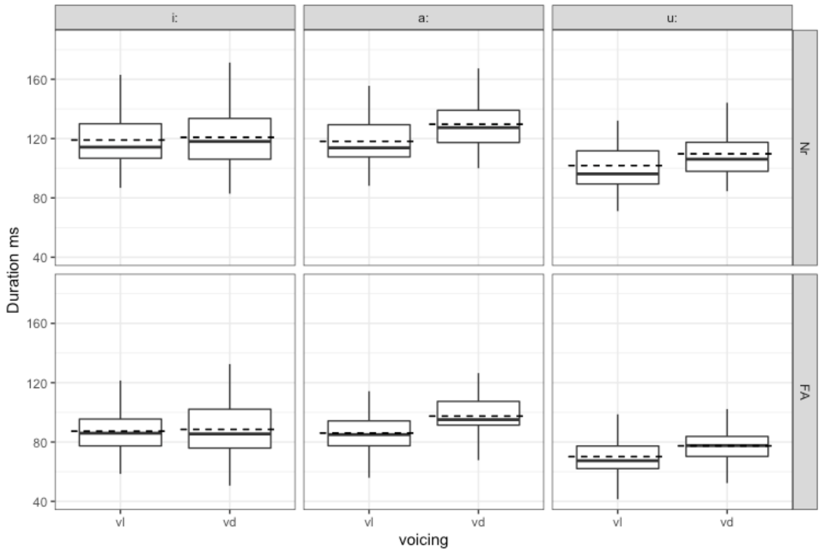


Figure C.5. Boxplots of FVD for Voiceless Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and vowel type.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
/i:/	119	19.3	121	21	87.3	14.6	88.5	17.4
/a:/	118	18.5	130	17.5	86.1	13.4	97.4	12.9
/u:/	102	19.2	110	16.2	70.1	13.3	77.4	11

Table C.68. Mean and standard deviation of FVD for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and vowel type.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr i: – vl Nr a:	0.623	vl Nr i: - vd Nr i:	0.266
vl Nr i: – vl Nr u:	<0.001	vl Nr a: - vd Nr a:	<0.001
vl Nr a: – vl Nr u:	<0.001	vl Nr u: - vd Nr u:	<0.001
vd Nr i: – vd Nr a:	<0.001	vl FA i: - vd FA i:	0.464
vd Nr i: – vd Nr u:	<0.001	vl FA a: - vd FA a:	<0.001
vd Nr a: – vd Nr u:	<0.001	vl FA u: - vd FA u:	<0.001
vl FA i: – vl FA a:	0.509		
vl FA i: – vl FA u:	<0.001		
vl FA a: – vl FA u:	<0.001		
vd FA i: – vd FA a:	<0.001		
vd FA i: – vd FA u:	<0.001		
vd FA a: – vd FA u:	<0.001		

Table C-69. Pairwise comparison output for FVD in utterance-medial stops (iambic) based on voicing, rate, and vowel type.

21. Preceding vowel duration for utterance-medial stops (trochaic)

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	116.92	9.53	12.27	**<0.001**
rate_c	-26.22	1.98	-13.23	**<0.001**
place_c	-6.57	14.11	-0.47	0.641
voicing_c	12.87	11.04	1.17	0.244
sex_c	7.49	4.78	1.57	0.117
vowel_c	-68.79	21.88	-3.14	**0.002**

Table C-70. LMM results for PVD in utterance-medial stops (trochaic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	0.45
vd Nr – vd FA	<0.001	vl FA – vd FA	0.17

Table C-71. Pairwise comparison output for PVD in utterance-medial stops (trochaic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.002	vl Nr A – vd Nr A	0.503
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	0.0002
vd Nr B – vd Nr A	0.906	vl FA A – vd FA A	0.563
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	0.735		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-72. Pairwise comparison output for PVD in utterance-medial stops (trochaic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

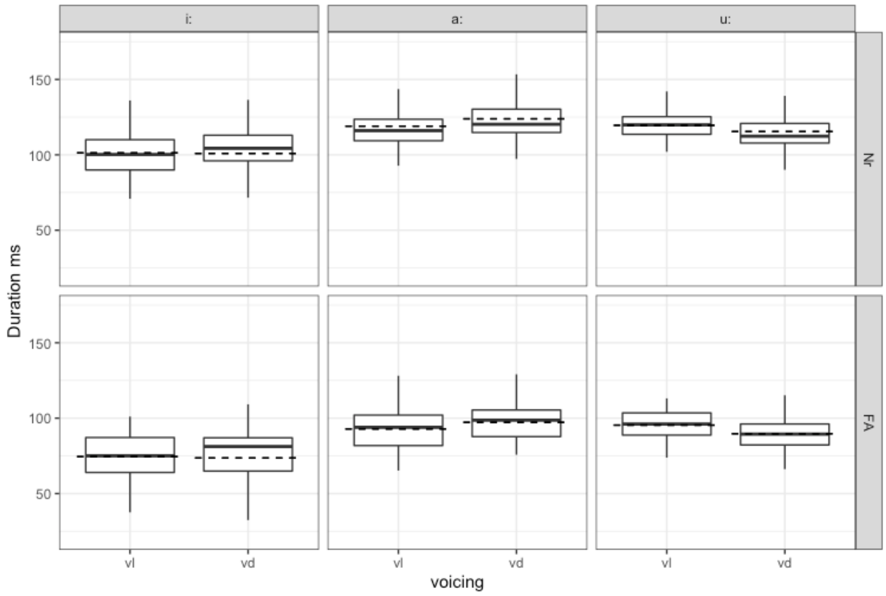


Figure C.6. Boxplots of PVD for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and vowel type.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
/i:/	101	14.8	101	21.2	74.5	14	73.6	21.1
/a:/	119	16.1	124	15	92.9	13.3	97.3	12.1
/u:/	120	15.8	115	13.4	95.4	11.4	89.7	10.5

Table C.73. Mean and standard deviation of PVD for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and vowel type.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr i: – vl Nr a:	<0.001	vl Nr i: - vd Nr i:	0.781
vl Nr i: – vl Nr u:	<0.001	vl Nr a: - vd Nr a:	0.0003
vl Nr a: – vl Nr u:	0.775	vl Nr u: - vd Nr u:	0.029
vd Nr i: – vd Nr a:	<0.001	vl FA i: - vd FA i:	0.557
vd Nr i: – vd Nr u:	<0.001	vl FA a: - vd FA a:	0.0016
vd Nr a: – vd Nr u:	<0.001	vl FA u: - vd FA u:	0.0033
vl FA i: – vl FA a:	<0.001		
vl FA i: – vl FA u:	<0.001		
vl FA a: – vl FA u:	0.193		
vd FA i: – vd FA a:	<0.001		
vd FA i: – vd FA u:	<0.001		
vd FA a: – vd FA u:	<0.001		

Table C-74. Pairwise comparison output for PVD in utterance-medial stops (trochaic) based on voicing, rate, and vowel type.

22. Preceding vowel duration for utterance-final stops

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	138.47	5.72	24.19	**<0.001**
vowel_c	2.28	6.28	0.36	0.716
rate_c	-26.55	3.69	-7.20	**<0.001**
place_c	4.21	6.61	0.64	0.524
voicing_c	3.46	5.15	0.67	0.501
sex_c	-10.78	10.15	-1.06	0.288

Table C-75. LMM results for PVD in utterance-final stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	0.94
vd Nr – vd FA	<0.001	vl FA – vd FA	0.88

Table C-76. Pairwise comparison output for PVD in utterance-final stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.68	vl Nr A – vd Nr A	0.70
vl FA A – vl FA V	0.75	vl Nr V – vd Nr V	0.47
vd Nr B – vd Nr A	0.9	vl FA A – vd FA A	0.75
vd Nr B – vd Nr V	0.49	vl FA V – vd FA V	0.75
vd Nr A – vd Nr V	0.49		
vd FA B – vd FA A	0.87		
vd FA B – vd FA V	0.60		
vd FA A – vd FA V	0.75		

Table C-77. Pairwise output for PVD in utterance-final stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

23. Burst intensity for utterance-initial stops

Predictors	*Estimates*	*std. Error*	*Statistic*	*p*
(Intercept)	52.04	0.68	76.95	**<0.001**
vowel_c	1.05	1.16	0.90	0.369
rate_c	1.32	0.36	3.70	**<0.001**
place_c	-4.17	1.22	-3.41	**0.001**
voicing_c	7.37	1.01	7.27	**<0.001**
sex_c	3.58	1.03	3.46	**0.001**

Table C-78. LMM results for burst intensity in utterance-initial stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-79. Pairwise comparison output for burst intensity in utterance-initial stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-80. Pairwise comparison output for burst intensity in utterance-initial stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

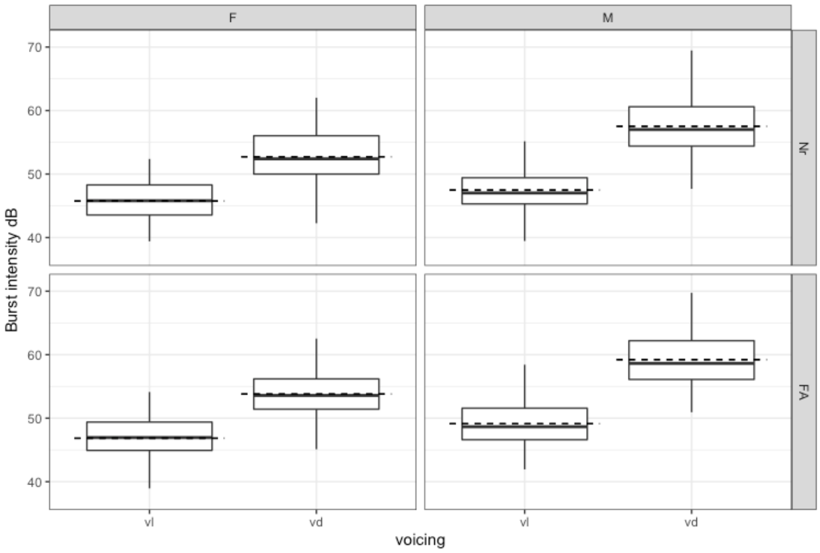


Figure C.7. Boxplots of the burst intensity for utterance-initial stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and gender. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	45.8	3.07	52.7	4.19	46.8	3.37	53.8	3.67
male	47.5	3.78	57.5	4.62	49.1	3.66	59.2	4.42

Table C.81. Mean and standard deviation of burst intensity for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-82. Pairwise comparison output for burst intensity in utterance-initial stops based on voicing, rate, and gender.

24. Burst intensity for utterance-medial stops (trochaic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	52.46	0.67	78.80	**<0.001**
vowel_c	1.08	0.77	1.41	0.159
rate_c	1.28	0.32	3.97	**<0.001**
place_c	-3.30	1.11	-2.98	**0.003**
voicing_c	8.32	0.91	9.19	**<0.001**
sex_c	4.70	1.08	4.33	**<0.001**

Table C-83. LMM results for burst intensity in utterance-medial stops (trochaic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female)**.**

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-84. Pairwise comparison output for burst intensity in utterance-medial stops (trochaic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-85. Pairwise comparison output for burst intensity in utterance-medial stops (trochaic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

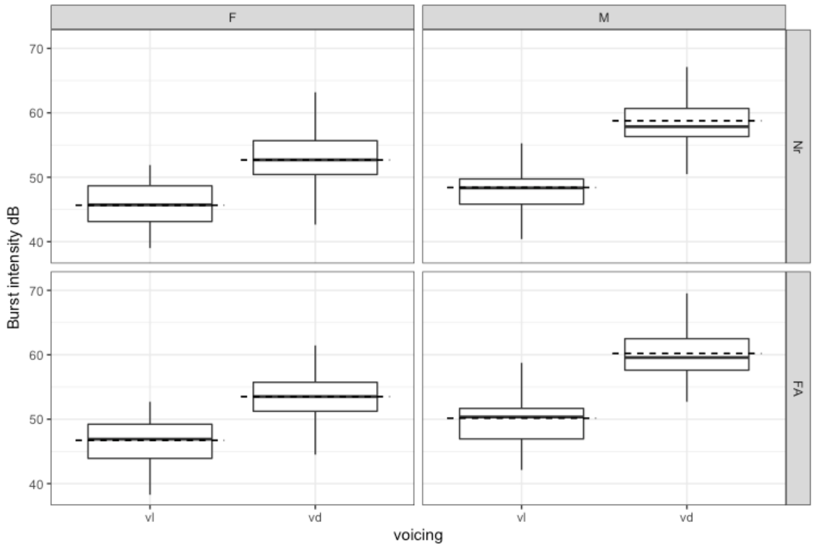


Figure C.8. Boxplots of the burst intensity for utterance-medial stops (Trochaic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and gender. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	45.7	0.826	52.7	0.659	46.7	1.16	53.5	0.949
male	48.4	0.755	58.8	0.522	50.2	0.875	60.2	0.617

Table C.86. Mean and standard deviation of burst intensity for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-87. Pairwise comparison output for burst intensity in utterance-medial stops (trochaic) based on voicing, rate, and gender.

25. Burst intensity for utterance-medial stops (iambic)

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	51.29	0.86	59.36	**<0.001**
vowel_c	1.07	1.24	0.86	0.390
rate_c	1.15	0.40	2.89	**0.004**
place_c	-3.24	1.59	-2.04	**0.042**
voicing_c	9.87	1.28	7.72	**<0.001**
sex_c	4.30	1.15	3.73	**<0.001**

Table C-88. LMM results for burst intensity in utterance-medial stops (iambic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	0.013	vl FA – vd FA	<0.001

Table C-89. Pairwise comparison output for burst intensity in utterance-medial stops (iambic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-90. Pairwise comparison output for burst intensity in utterance-medial stops (iambic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

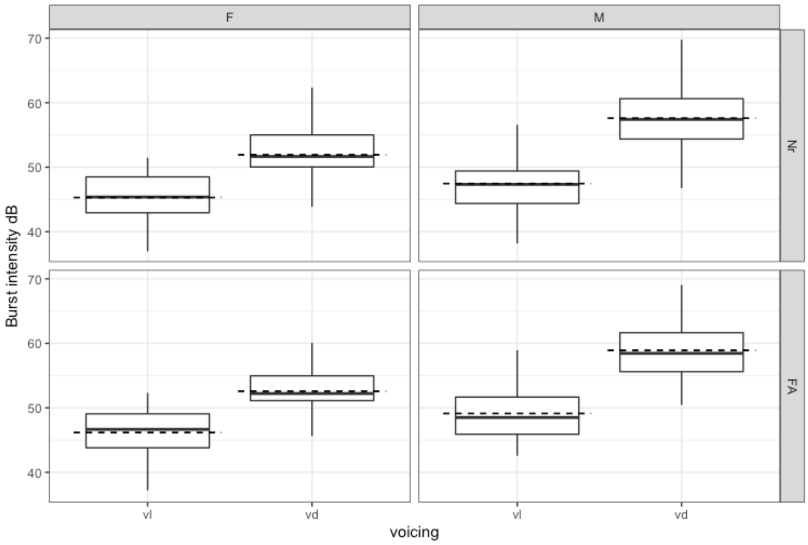


Figure C.9. Boxplots of the burst intensity for utterance-medial stops (Iambic) classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and gender. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	45.3	3.48	51.9	4.36	46.2	3.43	52.6	3.66
male	47.5	4.44	57.6	4.55	49.1	4.26	58.9	4.15

Table C.91. Mean and standard deviation of burst intensity for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-92. Pairwise comparison output for burst intensity in utterance-medial stops (iambic) based on voicing, rate, and gender.

26. Burst intensity for utterance-final stops

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	43.14	0.66	65.75	**<0.001**
vowel_c	1.23	0.37	3.35	**0.001**
rate_c	1.03	0.35	2.92	**0.003**
place_c	5.15	0.63	8.18	**<0.001**
voicing_c	2.26	0.38	5.99	**<0.001**
sex_c	1.23	1.30	0.94	0.347

Table C-93. LMM results for burst intensity in utterance-final stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-94. Pairwise comparison output for burst intensity in utterance-final stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-95. Pairwise comparison output for burst intensity in utterance-final stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

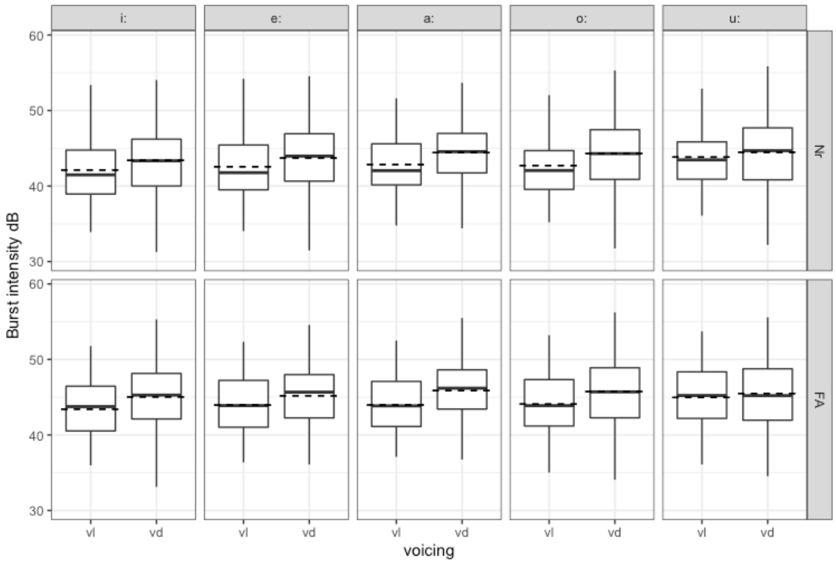


Figure C.10. Boxplots of the burst intensity for Voiced utterance-final stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and vowel type. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
/i:/	42.1	3.89	43.4	4.76	43.4	3.76	45	4.62
/e:/	42.6	4.05	43.7	4.84	44	3.88	45.2	4.5
/a:/	42.8	4.08	44.4	4.72	44	3.9	45.9	4.38
/o:/	42.7	4.28	44.3	4.91	44.1	4.08	45.7	4.65
/u:/	43.8	4.19	44.5	4.64	45	4.01	45.5	4.62

Table C.96. Mean and standard deviation of burst intensity for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and vowel type.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr i: – vl Nr e:	0.419	vl Nr i: - vd Nr i:	0.005
vl Nr i: – vl Nr a:	0.181	vl Nr e: - vd Nr e:	0.016
vl Nr i: – vl Nr o:	0.270	vl Nr a: - vd Nr a:	0.0011
vl Nr i: – vl Nr u:	<0.001	vl Nr o: - vd Nr o:	<0.001
vl Nr e: – vl Nr a:	0.606	vl Nr u: - vd Nr u:	0.187
vl Nr e: – vl Nr o:	0.787	vl FA i: - vd FA i:	0.0015
vl Nr e: – vl Nr u:	0.012	vl FA e: - vd FA e:	0.024
vl Nr a: – vl Nr o:	0.806	vl FA a: - vd FA a:	<0.001
vl Nr a: – vl Nr u:	0.064	vl FA o: - vd FA o:	<0.001
vl Nr o: – vl Nr u:	0.028	vl FA u: - vd FA u:	0.332
vd Nr i: – vd Nr e:	0.49		
vd Nr i: – vd Nr a:	0.017		
vd Nr i: – vd Nr o:	0.036		
vd Nr i: – vd Nr u:	0.012		
vd Nr e: – vd Nr a:	0.109		
vd Nr e: – vd Nr o:	0.187		
vd Nr e: – vd Nr u:	0.087		
vd Nr a: – vd Nr o:	0.805		

vd Nr a: – vd Nr u:	0.926
vd Nr o: – vd Nr u:	0.737
vl FA i: – vl FA e:	0.367
vl FA i: – vl FA a:	0.344
vl FA i: – vl FA o:	0.233
vl FA i: – vl FA u:	0.003
vl FA e: – vl FA a:	0.981
vl FA e: – vl FA o:	0.817
vl FA e: – vl FA u:	0.075
vl FA a: – vl FA o:	0.827
vl FA a: – vl FA u:	0.0705
vl FA o: – vl FA u:	0.102
vd FA i: – vd FA e:	0.783
vd FA i: – vd FA a:	0.079
vd FA i: – vd FA o:	0.141
vd FA i: – vd FA u:	0.375
vd FA e: – vd FA a:	0.156
vd FA e: – vd FA o:	0.256
vd FA e: – vd FA u:	0.561
vd FA a: – vd FA o:	0.805
vd FA a: – vd FA u:	0.454
vd FA o: – vd FA u:	0.624

Table C-97. Pairwise comparison output for burst intensity in utterance-final stop based on voicing, rate, vowel type

27. F0 onset for utterance-initial stops

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	198.72	3.47	57.21	**<0.001**
vowel_c	1.99	1.90	1.04	0.297
rate_c	4.96	1.30	3.82	**<0.001**
place_c	-1.35	2.14	-0.63	0.527
voicing_c	-22.91	2.34	-9.80	**<0.001**
sex_c	-75.57	6.77	-11.16	**<0.001**

Table C-98. LMM results for F0 onset in utterance-initial stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.0024	vl Nr – vd Nr	<0.001
vd Nr – vd FA	0.0041	vl FA – vd FA	<0.001

Table C-99. Pairwise comparison output for F0 onset in utterance-initial stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.303	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.788	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	0.541	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	0.23	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.543		
vd FA B – vd FA A	0.455		
vd FA B – vd FA V	0.233		
vd FA A – vd FA V	0.654		

Table C-100. Pairwise comparison output for F0 onset in utterance-initial stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

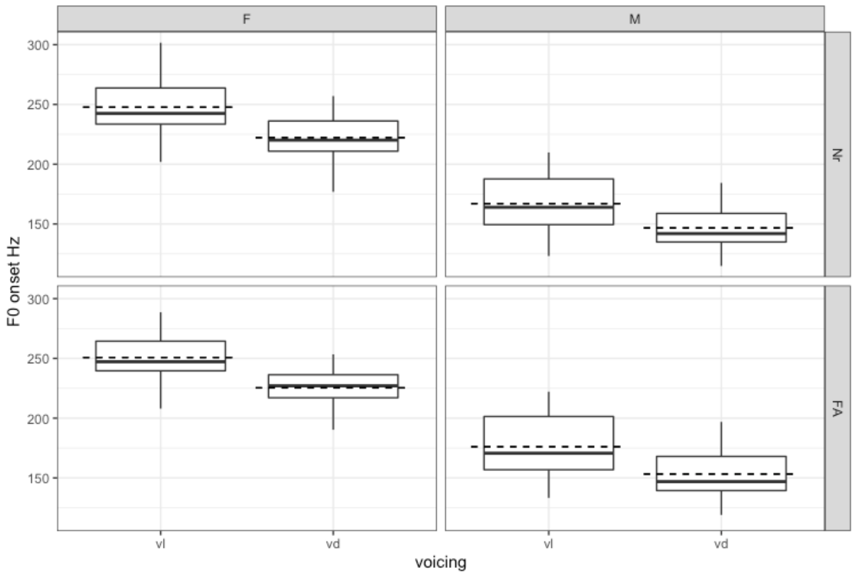


Figure C.11. Boxplots of F0 onset for Voiceless and Voiced utterance-initial stops classified by voicing (vl = Voiceless, vd = Voiced), speech rate, and gender. The dashed line represents the mean.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	248	20.2	222	17.4	251	17.1	225	14.5
male	167	22.5	147	17.3	176	25.9	153	20.2

Table C.101. Mean and standard deviation of F0 onset for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-102. Pairwise comparison output for F0 onset in utterance-initial stops based on voicing, rate, and gender.

28. F0 onset for utterance-medial stops (trochaic)

Predictors	*Estimates*	*std. Error*	*Statistic*	*p*
(Intercept)	196.11	3.91	50.21	**<0.001**
rate_c	2.68	1.19	2.26	**0.024**
]place_c	-2.48	2.90	-0.86	0.392
voicing_c	-13.39	2.64	-5.07	**<0.001**
sex_c	-73.73	7.30	-10.10	**<0.001**

Table C-103. LMM results for F0 onset in utterance-medial stops (trochaic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.0013	vl Nr – vd Nr	<0.001
vd Nr – vd FA	0.009	vl FA – vd FA	<0.001

Table C-104. Pairwise comparison output for F0 onset in utterance-medial stops (trochaic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.95	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.613	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	0.406	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	0.781	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.613		
vd FA B – vd FA A	0.71		
vd FA B – vd FA V	0.912		
vd FA A – vd FA V	0.796		

Table C-105. Pairwise comparison output for F0 onset in utterance-medial stops (trochaic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

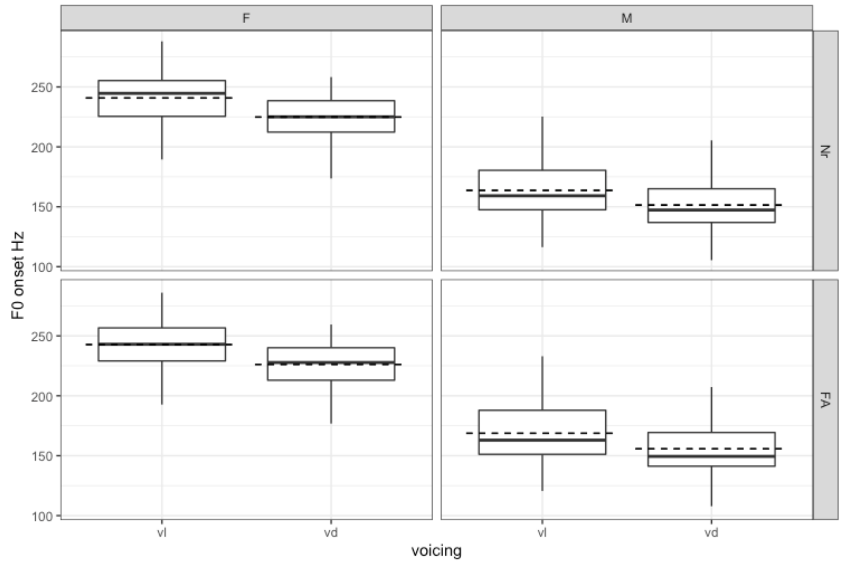


Figure C.12. Boxplots of F0 onset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing speech rate, and gender.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	241	20.2	225	17.5	243	19.3	226	17.5
male	164	23.3	151	22.8	169	26.2	156	24.3

Table C.106. Mean and standard deviation of F0 onset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-107. Pairwise comparison output for F0 onset in utterance- medial stops (trochaic) based on voicing, rate, and gender.

29. F0 onset for utterance-medial stops (iambic)

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	199.41	4.14	48.22	**<0.001**
rate_c	2.57	1.47	1.75	0.081
place_c	-3.44	2.26	-1.53	0.127
voicing_c	-23.95	2.21	-10.82	**<0.001**
sex_c	-72.12	7.80	-9.24	**<0.001**
vowel_L_c	2.20	3.23	0.68	0.496

Table C-108. LMM results for F0 onset in utterance-medial stops (iambic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.37	vl Nr – vd Nr	<0.001
vd Nr – vd FA	0.22	vl FA – vd FA	<0.001

Table C-109. Pairwise comparison output for F0 onset in utterance-medial stops (iambic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.67	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.85	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	0.81	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	0.62	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.78		
vd FA B – vd FA A	0.67		
vd FA B – vd FA V	0.47		
vd FA A – vd FA V	0.78		

Table C-110. Pairwise comparison output for F0 onset in utterance-medial stops (iambic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

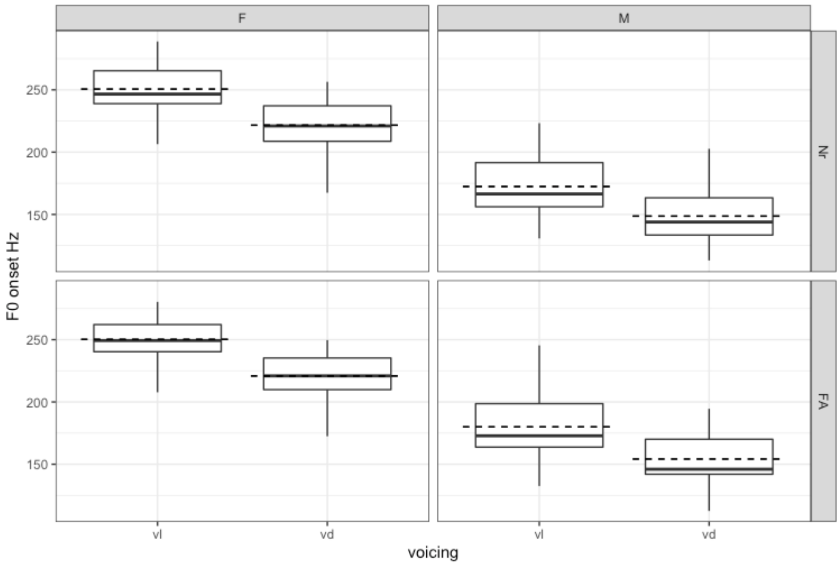


Figure C.13. Boxplots of F0 onset for Voiceless and utterance-medial stops (Iambic) grouped by voicing speech rate, and gender.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	251	18.8	222	18.8	250	16.4	221	17.2
male	172	24	149	21.3	180	28.4	154	24.9

Table C.111. Mean and standard deviation of F0 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-112. Pairwise comparison output for F0 onset in utterance- medial stops (iambic) based on voicing, rate, and gender.

30. F0 offset for utterance-medial stops (trochaic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	181.34	3.99	45.39	**<0.001**
rate_c	2.84	1.72	1.65	0.099
place_c	-6.63	2.33	-2.85	**0.004**
voicing_c	-5.37	2.41	-2.23	**0.026**
sex_c	-71.11	7.51	-9.47	**<0.001**
vowel_c	16.30	3.92	4.15	**<0.001**

Table C-113. LMM results for F0 offset in utterance-medial stops (trochaic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.59	vl Nr – vd Nr	0.42
vd Nr – vd FA	0.41	vl FA – vd FA	0.59

Table C-114. Pairwise comparison output for F0 offset in utterance-medial stops (trochaic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.168	vl Nr A – vd Nr A	0.374
vl FA A – vl FA V	0.374	vl Nr V – vd Nr V	0.645
vd Nr B – vd Nr A	0.675	vl FA A – vd FA A	0.578
vd Nr B – vd Nr V	0.168	vl FA V – vd FA V	0.656
vd Nr A – vd Nr V	0.421		
vd FA B – vd FA A	0.645		
vd FA B – vd FA V	0.168		
vd FA A – vd FA V	0.421		

Table C-115. Pairwise comparison output for F0 offset in utterance-medial stops (trochaic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

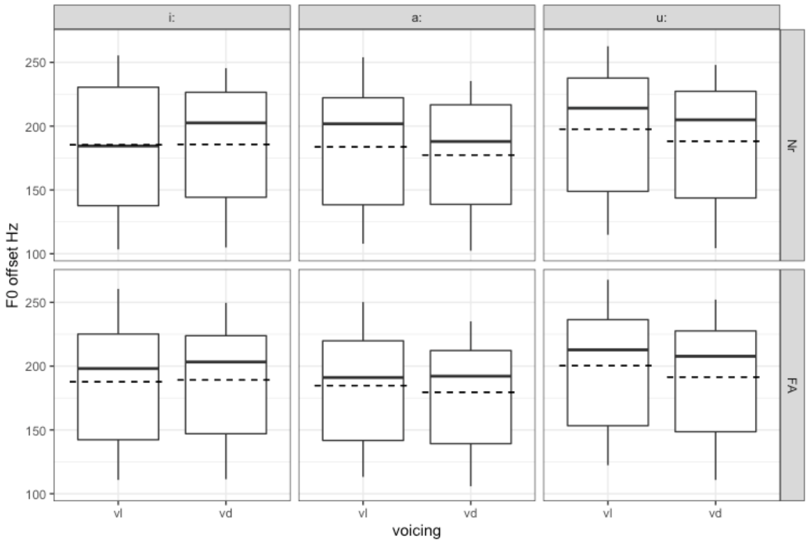


Figure C.14. Boxplots of F0 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and vowel type.

	Voiceless normal		Voiced normal		Voiceless fast		Voiced fast	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
/i:/	185	46.9	186	42.4	188	43.8	189	40.1
/a:/	184	44.5	177	40.9	185	41.6	180	38.7
/u:/	198	45.4	188	42.9	200	43.8	191	40.8

Table C.116. Mean and standard deviation of F0 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and vowel type.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr i: – vl Nr a:	0.772	vl Nr i: - vd Nr i:	0.93
vl Nr i: – vl Nr u:	0.063	vl Nr a: - vd Nr a:	0.156
vl Nr a: – vl Nr u:	0.026	vl Nr u: - vd Nr u:	0.135
vd Nr i: – vd Nr a:	0.051	vl FA i: - vd FA i:	0.784
vd Nr i: – vd Nr u:	0.616	vl FA a: - vd FA a:	0.265
vd Nr a: – vd Nr u:	0.016	vl FA u: - vd FA u:	0.156
vl FA i: – vl FA a:	0.618		
vl FA i: – vl FA u:	0.063		
vl FA a: – vl FA u:	0.016		
vd FA i: – vd FA a:	0.026		
vd FA i: – vd FA u:	0.647		
vd FA a: – vd FA u:	0.01		

Table C-117. Pairwise comparison output for F0 offset in utterance-medial stops (trochaic) based on voicing, rate, vowel type.

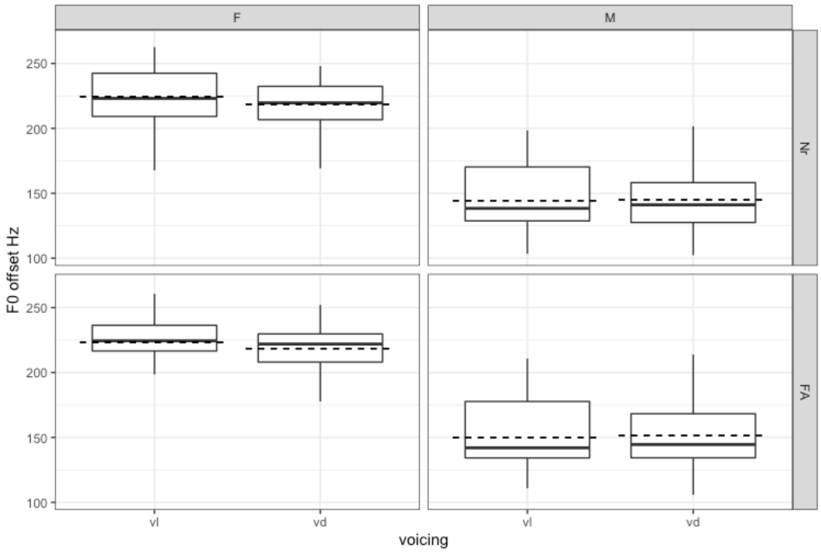


Figure C.15. Boxplots of F0 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and gender.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	224	21.8	218	17.1	223	19.7	218	16
male	144	22.2	145	24.7	150	25.4	151	27.6

Table C.118. Mean and standard deviation of F0 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	0.0004
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	0.693
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	0.0048
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	0.426

Table C-119. Pairwise comparison output for F0 offset in utterance-medial stops (trochaic) based on voicing, rate, and gender.

31. F0 offset for utterance-medial stops (iambic)

Predictors	*Estimates*	*std. Error*	*Statistic*	*p*
(Intercept)	170.18	3.33	51.11	**<0.001**
rate_c	3.53	0.97	3.63	**<0.001**
place_c	-3.78	1.72	-2.19	**0.028**
voicing_c	1.34	1.53	0.87	0.382
sex_c	-68.85	6.57	-10.48	**<0.001**

Table C-120. LMM results for F0 offset in utterance-medial stops (iambic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.433	vl Nr – vd Nr	0.59
vd Nr – vd FA	0.067	vl FA – vd FA	0.373

Table C-121. Pairwise comparison output for F0 offset in utterance-medial stops (iambic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.88	vl Nr A – vd Nr A	0.88
vl FA A – vl FA V	0.75	vl Nr V – vd Nr V	0.98
vd Nr B – vd Nr A	0.75	vl FA A – vd FA A	0.75
vd Nr B – vd Nr V	0.59	vl FA V – vd FA V	0.84
vd Nr A – vd Nr V	0.75		
vd FA B – vd FA A	0.88		
vd FA B – vd FA V	0.59		
vd FA A – vd FA V	0.61		

Table C-122. Pairwise comparison output for F0 offset in utterance-medial stops (iambic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

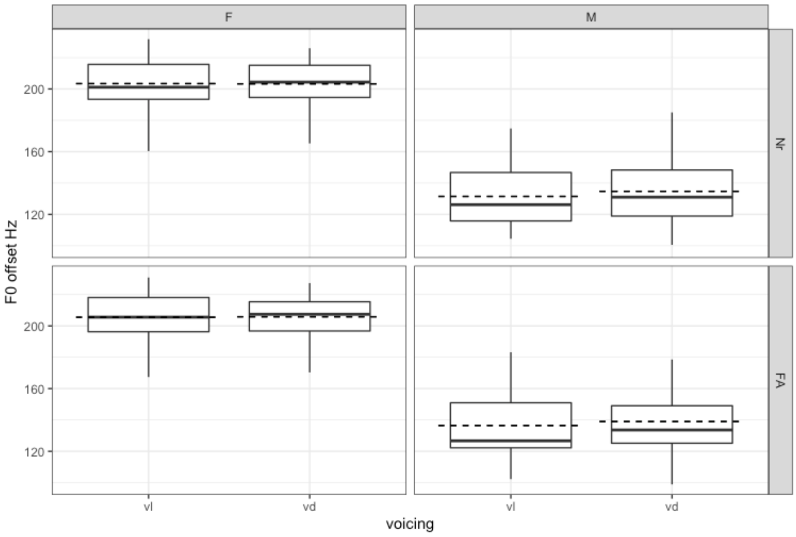


Figure C.16. Boxplots of F0 offset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing speech rate, and gender.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	203	16.1	203	14.5	206	14.2	206	12.8
male	131	20.7	135	21.2	136	22.2	139	22.9

Table C.123. Mean and standard deviation of F0 offset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	0.921
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	0.044
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	0.889
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	0.149

Table C-124. Pairwise comparison output for F0 offset in utterance-medial stops (iambic) based on voicing, rate, and gender.

32. F0 offset for utterance-final stops

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	175.77	4.83	36.39	**<0.001**
vowel_c	5.28	3.61	1.46	0.144
rate_c	-3.73	2.65	-1.41	0.160
place_c	3.48	3.71	0.94	0.349
voicing_c	-3.01	3.33	-0.90	0.367
sex_c	-91.89	9.30	-9.88	**<0.001**

Table C-125. LMM results for F0 offset in utterance-final stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.0742	vl Nr – vd Nr	0.618
vd Nr – vd FA	0.0014	vl FA – vd FA	0.177

Table C-126. Pairwise comparison output for F0 offset in utterance-final stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.23	vl Nr A – vd Nr A	0.452
vl FA A – vl FA V	0.053	vl Nr V – vd Nr V	0.226
vd Nr B – vd Nr A	0.33	vl FA A – vd FA A	0.049
vd Nr B – vd Nr V	0.053	vl FA V – vd FA V	0.4
vd Nr A – vd Nr V	0.43		
vd FA B – vd FA A	0.971		
vd FA B – vd FA V	0.331		
vd FA A – vd FA V	0.3314		

Table C-127. Pairwise comparison output for F0 offset in utterance-final stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

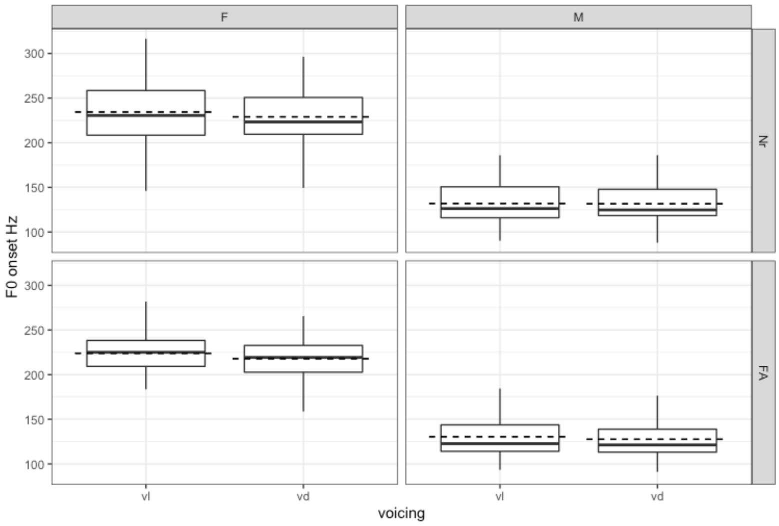


Figure C.17. Boxplots of F0 offset for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and gender.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	234	36.9	229	29.5	224	28.2	218	23
male	132	25.2	132	22.9	130	24.1	128	21.8

Table C.128. Mean and standard deviation of F0 offset for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	0.0036
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	0.865
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	0.0005
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	0.143

Table C-129. Pairwise comparison output for F0 offset in utterance-final stops based on voicing, rate, and gender.

33. F1 onset for utterance-initial stops

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	418.79	15.21	27.53	<**0.001**
vowel_L_c	62.73	37.30	1.68	0.093
rate_c	11.57	1.75	6.60	<**0.001**
place_c	-29.76	37.48	-0.79	0.427
voicing_c	-38.69	28.98	-1.33	0.182
sex_c	-44.27	9.60	-4.61	<**0.001**

Table C-130. LMM results for F1 onset in utterance-initial stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-131. Pairwise comparison output for F1 onset in utterance-initial stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.273	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.29	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.066		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	0.144		

Table C-132. Pairwise comparison output for F1 onset in utterance-initial stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

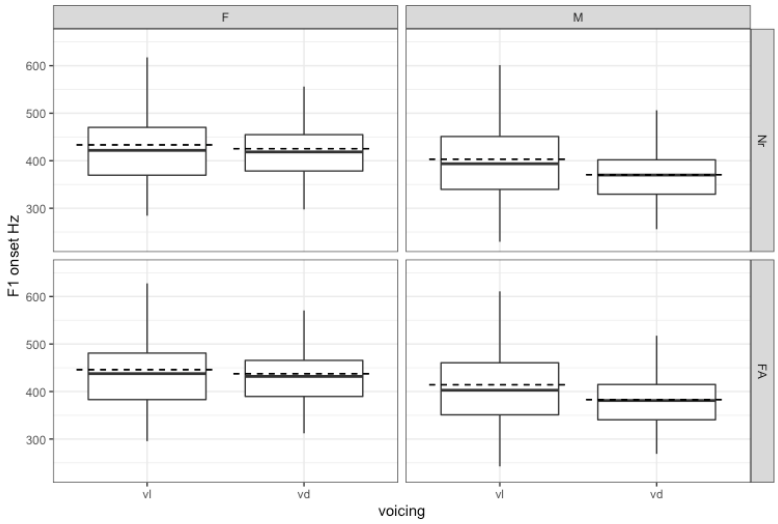


Figure C.18. Boxplots of F1 onset for Voiceless and Voiced utterance-initial stops grouped voicing, speech rate and gender.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	433	84.5	425	63.5	446	84.4	437	62.2
male	403	83.2	371	57.6	414	82.8	383	59

Table C.133. Mean and standard deviation of F1 onset for Voiceless and Voiced utterance-initial stops grouped by voicing, speech rate and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	0.053
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	0.042
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-134. Pairwise comparison output for F1 onset in utterance-initial stops based on voicing, rate, and gender.

34. F1 onset for utterance-medial stops (iambic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	485.56	16.37	29.67	**<0.001**
rate_c	9.41	2.26	4.17	**<0.001**
place_c	-41.36	24.01	-1.72	0.085
voicing_c	-26.11	18.39	-1.42	0.156
sex_c	-34.79	13.49	-2.58	**0.010**
vowel_c	-263.96	36.64	-7.20	**<0.001**

Table C-135. LMM results for F1 onset in utterance-medial stops (iambic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.066	vl Nr – vd Nr	<0.001
vd Nr – vd FA	0.055	vl FA – vd FA	<0.001

Table C-136. Pairwise comparison output for F1 onset in utterance-medial stops (iambic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.638	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.841	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.904		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	0.982		

Table C-137. Pairwise comparison output for F1 onset in utterance-medial stops (iambic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

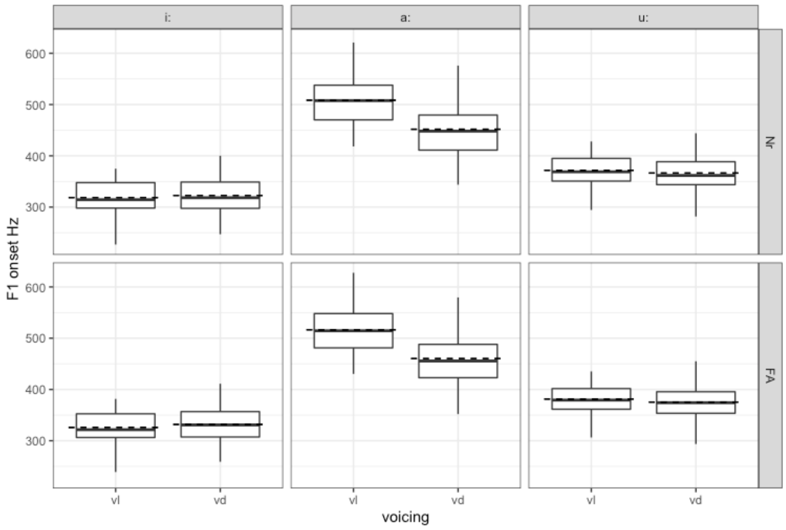


Figure C.19. Boxplots of F1 onset for Voiceless Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and vowel type.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
/i:/	319	32.6	323	33.5	326	30.7	332	33.6
/a:/	509	45.2	452	52.7	517	44.3	460	52.7
/u:/	372	31.4	366	33.6	381	30.3	375	33.3

Table C.138. Mean and standard deviation of F1 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and vowel type.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr i: – vl Nr a:	<0.001	vl Nr i: - vd Nr i:	0.276
vl Nr i: – vl Nr u:	<0.001	vl Nr a: - vd Nr a:	<0.001
vl Nr a: – vl Nr u:	<0.001	vl Nr u: - vd Nr u:	0.153
vd Nr i: – vd Nr a:	<0.001	vl FA i: - vd FA i:	0.133
vd Nr i: – vd Nr u:	<0.001	vl FA a: - vd FA a:	<0.001
vd Nr a: – vd Nr u:	<0.001	vl FA u: - vd FA u:	0.122
vl FA i: – vl FA a:	<0.001		
vl FA i: – vl FA u:	<0.001		
vl FA a: – vl FA u:	<0.001		
vd FA i: – vd FA a:	<0.001		
vd FA i: – vd FA u:	<0.001		
vd FA a: – vd FA u:	<0.001		

Table C-139. Pairwise comparison output for F1 onset in utterance-medial stops (iambic) based on voicing, rate, vowel type.

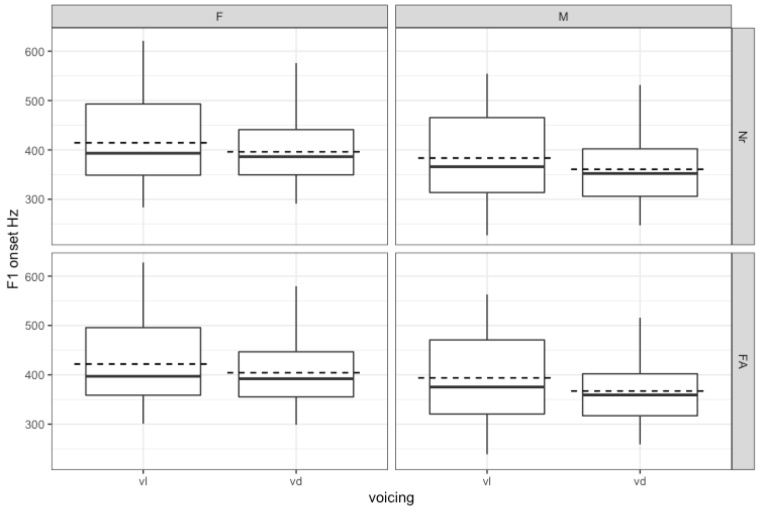


Figure C.20. Boxplots of F1 onset for Voiceless Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and gender.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	415	88.6	396	65.4	422	86.9	404	64.2
male	384	84.7	361	64.3	394	85.3	367	63.1

Table C.140. Means and standard deviations of F1 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	0.0017
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	0.0034
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-141. Pairwise comparison output for F1 onset in utterance-medial stops (iambic) based on voicing, rate, and gender.

35. F1 offset for utterance-medial stops (trochaic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	508.58	13.25	38.39	**<0.001**
rate_c	14.59	3.12	4.68	**<0.001**
place_c	-64.11	17.83	-3.60	**<0.001**
voicing_c	-9.15	14.80	-0.62	0.536
sex_c	-46.54	14.99	-3.11	**0.002**
vowel_c	-260.18	28.88	-9.01	**<0.001**

Table C-142. LMM results for F1 offset in utterance-medial stops (trochaic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.007	vl Nr – vd Nr	<0.001
vd Nr – vd FA	0.055	vl FA – vd FA	<0.001

Table C-143. Pairwise comparison output for F1 offset in utterance-medial stops (trochaic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	0.075
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	0.334
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-144. Pairwise comparison output for F1 offset in utterance-medial stops (trochaic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

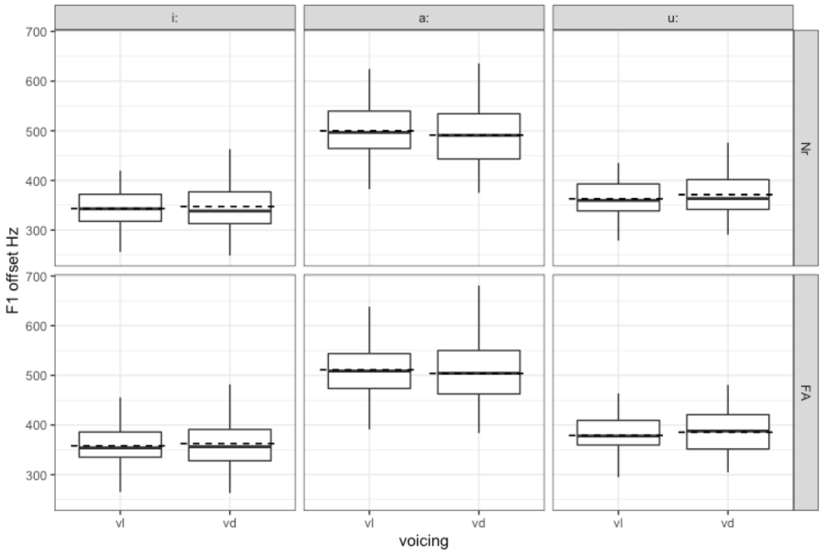


Figure C.21. Boxplots of F1 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and vowel type.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
/i:/	344	39.7	348	51	358	38.5	363	49.3
/a:/	500	54.5	492	63.1	511	50.5	504	60.4
/u:/	363	38.4	371	42.4	379	36.6	385	40

Table C.145. Mean and standard deviation of F1 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and vowel type.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr i: – vl Nr a:	<0.001	vl Nr i: - vd Nr i:	0.363
vl Nr i: – vl Nr u:	0.002	vl Nr a: - vd Nr a:	0.055
vl Nr a: – vl Nr u:	<0.001	vl Nr u: - vd Nr u:	0.18
vd Nr i: – vd Nr a:	<0.001	vl FA i: - vd FA i:	0.344
vd Nr i: – vd Nr u:	<0.001	vl FA a: - vd FA a:	0.111
vd Nr a: – vd Nr u:	<0.001	vl FA u: - vd FA u:	0.294
vl FA i: – vl FA a:	<0.001		
vl FA i: – vl FA u:	0.0015		
vl FA a: – vl FA u:	<0.001		
vd FA i: – vd FA a:	<0.001		
vd FA i: – vd FA u:	<0.001		
vd FA a: – vd FA u:	<0.001		

Table C-146. Pairwise comparison output for F1 offset in utterance-medial stops (trochaic) based on voicing, rate, and vowel type.

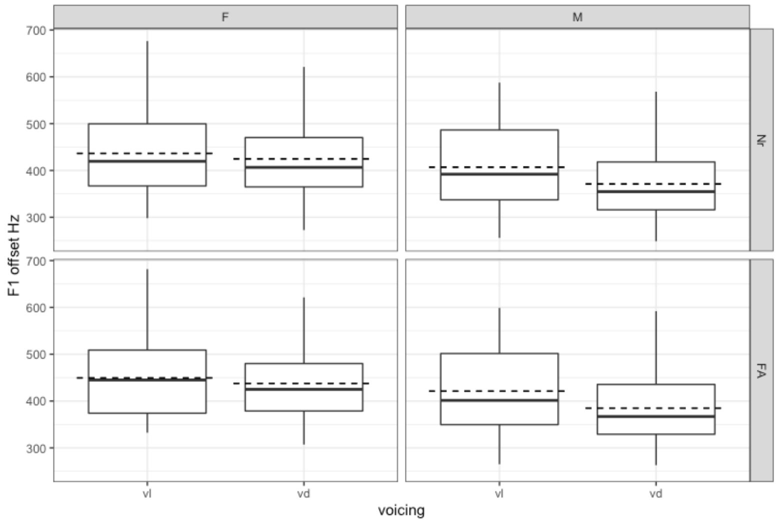


Figure C.22. Boxplots of F1 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and gender.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	437	88.7	425	81.6	449	83	438	78.3
male	407	86.3	371	71.9	421	86.5	385	71.4

Table C.147. Mean and standard deviation of F1 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	0.056
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	0.056
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-148. Pairwise comparison output for F1 offset in utterance-medial stops (trochaic) based on voicing, rate, and gender.

36. F1 offset for utterance-final stops

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	424.49	16.92	25.09	**<0.001**
vowel_L_c	11.75	35.21	0.33	0.739
rate_c	2.36	2.95	0.80	0.424
place_c	-90.70	34.48	-2.63	**0.009**
voicing_c	24.84	29.06	0.85	0.393
sex_c	-35.40	13.38	-2.65	**0.008**

Table C-149. LMM results for F1 offset in utterance-final stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.742	vl Nr – vd Nr	0.124
vd Nr – vd FA	0.333	vl FA – vd FA	0.044

Table C-150. Pairwise comparison output for F1 offset in utterance-final stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.008	vl Nr A – vd Nr A	0.928
vl FA A – vl FA V	0.023	vl Nr V – vd Nr V	0.002
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	0.419
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	0.002
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-151. Pairwise comparison output for F1 offset in utterance-final stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

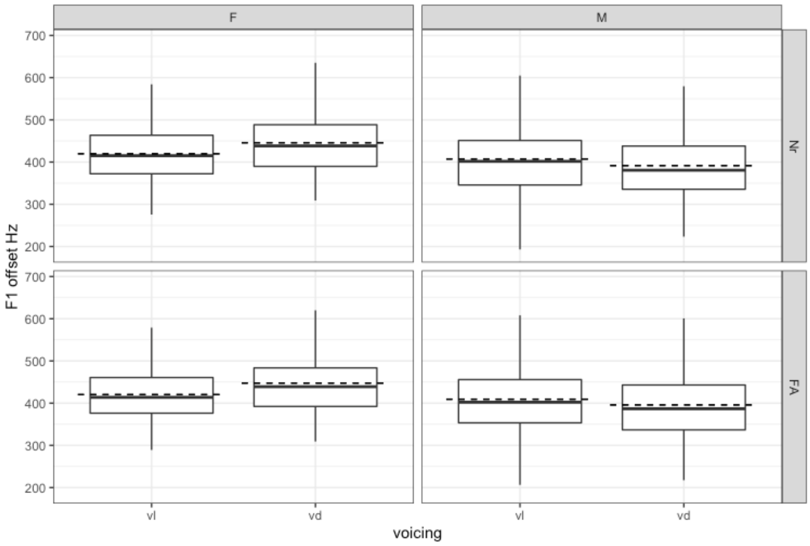


Figure C.23. Boxplots of F1 offset for Voiceless and utterance-final stops grouped by voicing, speech rate, and gender.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	420	68.2	446	75.9	420	64.7	448	74.1
male	407	86.4	392	79.2	409	85.1	396	80

Table C.152. Mean and standard deviation of F1 offset for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-153. Pairwise comparison output for F1 offset in utterance-final stops based on voicing, rate, and gender.

37. H1-H2 onset for utterance-initial stops

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	4.69	0.52	9.07	**<0.001**
rate_c	-0.57	0.16	-3.50	**<0.001**
place_c	1.04	0.77	1.35	0.177
voicing_c	-1.81	0.74	-2.44	**0.015**
sex_c	-1.35	0.76	-1.79	0.074
vowel_L_c	0.76	0.77	0.98	0.326

Table C-154. LMM results for H1-H2 onset in utterance-initial stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-155. Pairwise comparison output for H1-H2 onset in utterance-initial stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.022	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.04	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	<0.001	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	0.0045
vd Nr A – vd Nr V	<0.001		
vd FA B – vd FA A	<0.001		
vd FA B – vd FA V	<0.001		
vd FA A – vd FA V	<0.001		

Table C-156. Pairwise comparison output for H1-H2 onset in utterance-initial stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

38. H1-H2 onset for utterance-medial stops (iambic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	3.02	0.86	3.51	**<0.001**
rate_c	-0.36	0.18	-2.06	**0.040**
place_c	1.34	0.86	1.55	0.121
voicing_c	-0.91	0.91	-1.00	0.318
sex_c	-1.28	0.98	-1.31	0.191
vowel_L_c	3.78	1.37	2.77	**0.006**

Table C-157. LMM results for H1-H2 onset in utterance-medial stops (iambic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.095	vl Nr – vd Nr	0.886
vd Nr – vd FA	0.084	vl FA – vd FA	0.886

Table C-158. Pairwise comparison output for H1-H2 onset in utterance-medial stops (iambic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.0017	vl Nr A – vd Nr A	0.507
vl FA A – vl FA V	0.0033	vl Nr V – vd Nr V	0.233
vd Nr B – vd Nr A	0.097	vl FA A – vd FA A	0.469
vd Nr B – vd Nr V	0.71	vl FA V – vd FA V	0.236
vd Nr A – vd Nr V	0.223		
vd FA B – vd FA A	0.078		
vd FA B – vd FA V	0.507		
vd FA A – vd FA V	0.274		

Table C-159. Pairwise comparison output for H1-H2 onset in utterance-medial stops (iambic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

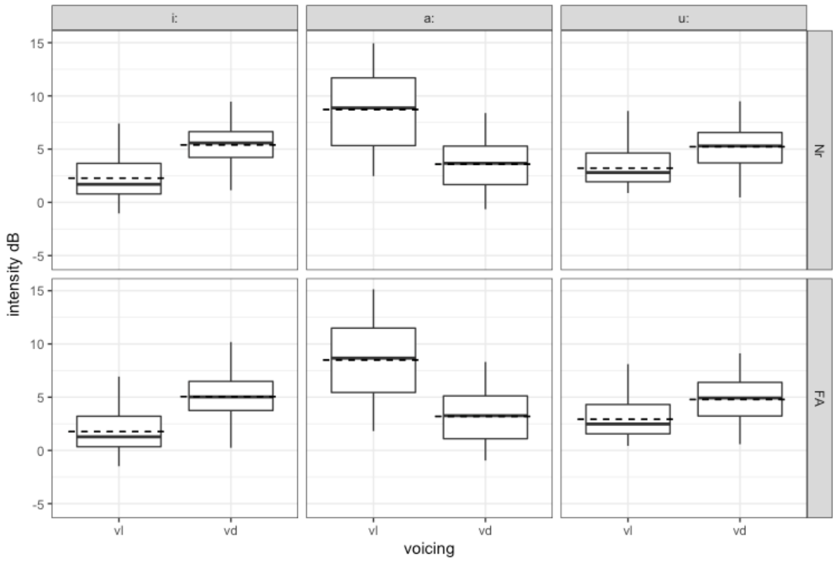


Figure C.24. Boxplots of H1-H2 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and vowel type.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
/i:/	2.28	2.25	5.4	1.96	1.76	2.24	5.06	2.03
/a:/	8.72	3.52	3.59	2.29	8.49	3.58	3.2	2.44
/u:/	3.22	2.49	5.23	2.12	2.92	2.30	4.78	2.19

Table C.160. Mean and standard deviation of H1-H2 onset for Voiceless and Voiced utterance-medial stops (Iambic) grouped by voicing, speech rate, and vowel type.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr i: – vl Nr a:	<0.001	vl Nr i: - vd Nr i:	<0.001
vl Nr i: – vl Nr u:	<0.001	vl Nr a: - vd Nr a:	<0.001
vl Nr a: – vl Nr u:	<0.001	vl Nr u: - vd Nr u:	<0.001
vd Nr i: – vd Nr a:	<0.001	vl FA i: - vd FA i:	<0.001
vd Nr i: – vd Nr u:	0.414	vl FA a: - vd FA a:	<0.001
vd Nr a: – vd Nr u:	<0.001	vl FA u: - vd FA u:	<0.001
vl FA i: – vl FA a:	<0.001		
vl FA i: – vl FA u:	<0.001		
vl FA a: – vl FA u:	<0.001		
vd FA i: – vd FA a:	<0.001		
vd FA i: – vd FA u:	0.198		
vd FA a: – vd FA u:	<0.001		

Table C-161. Pairwise comparison output for H1-H2 onset in utterance-medial stops (iambic) based on voicing, rate, vowel type

39. H1-H2 offset for utterance-medial stops (trochaic)

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	6.35	0.66	9.63	**<0.001**
rate_c	-0.39	0.19	-2.09	**0.036**
place_c	0.90	0.90	1.01	0.314
voicing_c	-2.58	0.73	-3.53	**<0.001**
sex_c	-3.88	0.73	-5.31	**<0.001**
vowel_c	-2.27	1.49	-1.53	0.127

Table C-162. LMM results for H1-H2 offset in utterance-medial stops (trochaic) as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/a:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	0.016	vl Nr – vd Nr	<0.001
vd Nr – vd FA	0.016	vl FA – vd FA	<0.001

Table C-163. Pairwise comparison output for H1-H2 offset in utterance-medial stops (trochaic) based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	<0.001	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	<0.001	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	0.604	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	<0.001	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.006		
vd FA B – vd FA A	0.693		
vd FA B – vd FA V	0.011		
vd FA A – vd FA V	0.036		

Table C-164. Pairwise comparison output for H1-H2 offset in utterance-medial stops (trochaic) based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

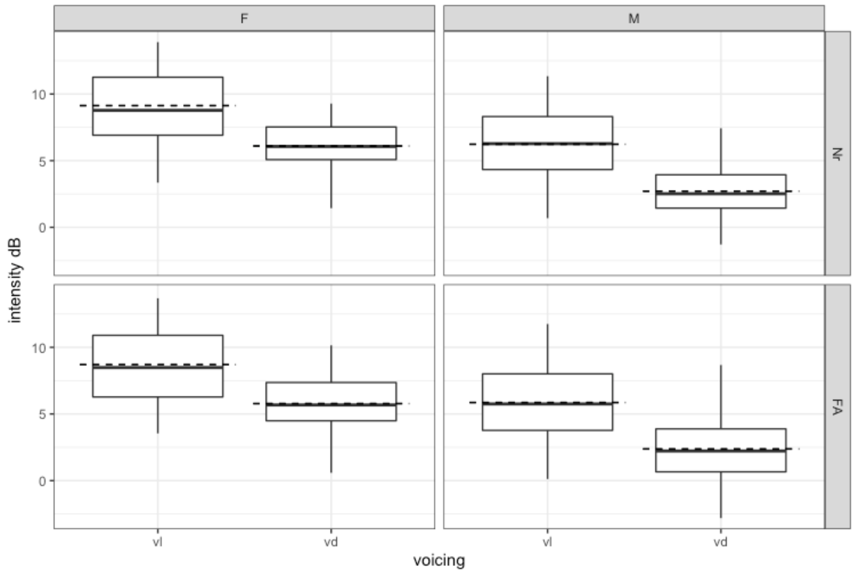


Figure C.25. Boxplots of H1-H2 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and gender.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	9.13	2.62	6.1	1.64	8.70	2.72	5.78	1.88
male	6.23	2.76	2.71	2.03	5.86	2.94	2.38	2.33

Table C.165. Mean and standard deviation of H1-H2 offset for Voiceless and Voiced utterance-medial stops (Trochaic) grouped by voicing, speech rate and gender.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-166. Pairwise comparison output for H1-H2 offset in utterance-medial stops (trochaic) based on voicing, rate, and gender.

40. H1-H2 offset for utterance-final stops

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	4.46	0.42	10.50	**<0.001**
rate_c	-0.65	0.29	-2.24	**0.025**
place_c	-0.08	0.65	-0.12	0.902
voicing_c	-3.29	0.64	-5.19	**<0.001**
sex_c	-3.47	0.73	-4.75	**<0.001**
vowel_c	0.13	0.60	0.21	0.833

Table C-167. LMM results for H1-H2 offset in utterance-final stops as a function of voicing, place (bilabial, alveolar, velar), rate (normal/fast), vowel (i:/e:/a:/o:/u:), and gender (male/female).

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr – vl FA	<0.001	vl Nr – vd Nr	<0.001
vd Nr – vd FA	<0.001	vl FA – vd FA	<0.001

Table C-168. Pairwise comparison output for H1-H2 offset in utterance-final stops based on voicing and rate.

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr A – vl Nr V	0.992	vl Nr A – vd Nr A	<0.001
vl FA A – vl FA V	0.193	vl Nr V – vd Nr V	<0.001
vd Nr B – vd Nr A	0.256	vl FA A – vd FA A	<0.001
vd Nr B – vd Nr V	0.88	vl FA V – vd FA V	<0.001
vd Nr A – vd Nr V	0.203		
vd FA B – vd FA A	0.0013		
vd FA B – vd FA V	0.509		
vd FA A – vd FA V	0.0101		

Table C-169. Pairwise comparison output for H1-H2 offset in utterance-final stops based on voicing, rate, place of articulation (B: bilabial/A: alveolar/V: velar).

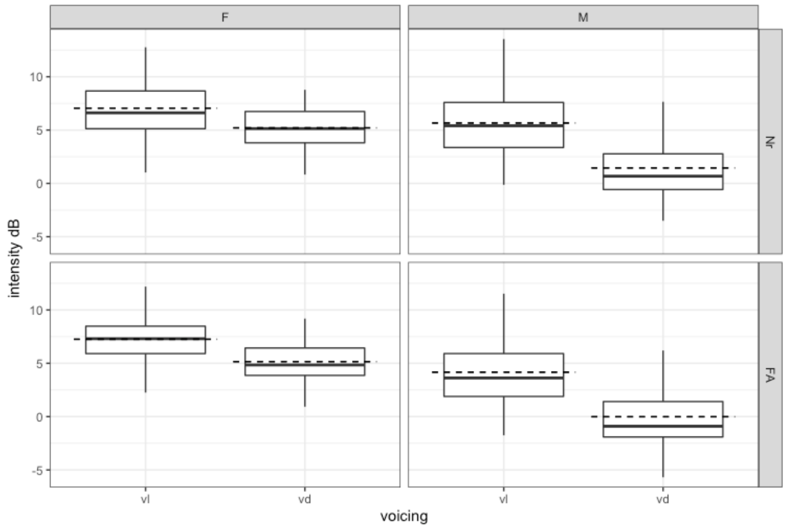


Figure C.26. Boxplots of H1-H2 offset for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and vowel type.

	Voiceless normal		**Voiced normal**		**Voiceless fast**		**Voiced fast**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	7.06	2.52	5.21	1.88	7.25	1.94	5.15	1.9
male	5.66	3.11	1.45	2.84	4.17	3.27	0.013	3.03

Table C.170. Mean and standard deviation of H1-H2 offset for Voiceless and Voiced utterance-final stops grouped by voicing, speech rate, and gender

Rate		Voicing	
Pair	p-value	Pair	p-value
vl Nr F – vl Nr M	<0.001	vl Nr F – vd Nr F	<0.001
vl FA F – vl FA M	<0.001	vl Nr M – vd Nr M	<0.001
vd Nr F – vd Nr M	<0.001	vl FA F – vd FA F	<0.001
vd FA F – vd FA M	<0.001	vl FA M – vd FA M	<0.001

Table C-171. Pairwise comparison output for H1-H2 offset in utterance-final stops based on voicing, rate, and gender.

41. Voicing duration in C1 vs C1 baseline

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	14.58	4.37	3.33	**0.001**
sex_c	-1.08	2.08	-0.52	0.605
context_c	15.66	7.74	2.02	**0.043**
place_c	-2.22	13.54	-0.16	0.870
voicing_c	27.36	8.43	3.25	**0.001**
cluster_c	-2.77	13.53	-0.20	0.838

Table C-172. LMM results for Voicing duration in stop-stop clusters as a function of voicing, place (bilabial, alveolar, velar), cluster (/bk/kb/dg/tg/gt), context (C1/C1 baseline), and gender (male/female).

42. F0 offset in C1 vs C1 baseline

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	172.30	4.77	36.11	**<0.001**
sex_c	-83.83	7.87	-10.65	**<0.001**
context_c	-11.20	6.80	-1.65	0.100
place_c	8.26	9.05	0.91	0.362
voicing_c	-4.05	5.85	-0.69	0.489
cluster_c	6.69	9.06	0.74	0.460

Table C-173. LMM results for F0 offset in stop-stop clusters as a function of voicing, place (bilabial, alveolar, velar), cluster (/bk/kb/dg/tg/gt), context (C1/C1 baseline), and gender (male/female).

43. F1 offset in C1 vs C1 baseline

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	414.14	9.63	42.99	**<0.001**
sex_c	-31.22	12.47	-2.50	**0.012**
context_c	-40.50	16.61	-2.44	**0.015**
place_c	-14.95	30.47	-0.49	0.624
voicing_c	-48.13	19.18	-2.51	**0.012**
cluster_c	-31.62	30.10	-1.05	0.293

Table C-174. LMM results for F1 offset in stop-stop clusters as a function of voicing, place (bilabial, alveolar, velar), cluster (/bk/kb/dg/tg/gt), context (C1/C1 baseline), and gender (male/female).

44. Burst intensity in C1 vc C1 baseline

Predictors	*Estimates*	*std. Error*	*t*	*p*
(Intercept)	42.92	0.83	51.55	**<0.001**
sex_c	2.90	1.53	1.90	0.058
context_c	1.51	0.94	1.60	0.110
place_c	6.75	1.53	4.40	**<0.001**
voicing_c	2.35	0.96	2.45	**0.014**
cluster_c	-1.54	1.53	-1.00	0.316

Table C-175. LMM results for burst intensity in stop-stop clusters as a function of voicing, place (bilabial, alveolar, velar), cluster (/bk/kb/dg/tg/gt), context (C1/C1 baseline), and gender (male/female).

45. Voicing duration in C2 vs C2 baseline

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	53.49	3.67	14.58	<**0.001**
sex_c	-7.78	3.93	-1.98	**0.048**
context_c	-9.44	5.33	-1.77	0.077
cluster_c	-11.11	13.24	-0.84	0.401
place_c	5.84	11.31	0.52	0.605

Table C-176. LMM results for voicing duration in stop-stop clusters as a function of place (bilabial, alveolar, velar), cluster (/bk/kb/dg/tg/gt), context (C2/C2 baseline), and gender (male/female).

46. F0 onset in C2 vs C2 basline

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	193.62	3.68	52.68	<**0.001**
sex_c	-74.22	7.00	-10.61	<**0.001**
context_c	-1.75	2.46	-0.71	0.477
cluster_c	-2.22	3.23	-0.69	0.491
voicing_c	-16.21	2.70	-6.00	<**0.001**
place_c	-9.46	2.86	-3.30	**0.001**

Table C-177. LMM results for F0 onset in stop-stop clusters as a function of voicing, place (bilabial, alveolar, velar), cluster (/bk/kb/dg/tg/gt), context (C2/C2 baseline), and gender (male/female).

47. F1 onset in C2 vs C2 baseline

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	500.30	16.51	30.31	**<0.001**
sex_c	-53.79	26.18	-2.05	**0.040**
context_c	-0.23	25.13	-0.01	0.993
cluster_c	6.60	34.75	0.19	0.849
voicing_c	-33.41	28.82	-1.16	0.246
place_c	-17.63	31.98	-0.55	0.582

Table C-178. LMM results for F0 onset in stop-stop clusters as a function of voicing, place (bilabial, alveolar, velar), cluster (/bk/kb/dg/tg/gt), context (C2/C2 baseline), and gender (male/female).

48. Burst intensity in C2 vs C2 baseline

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	51.62	0.74	69.90	**<0.001**
sex_c	4.66	1.27	3.66	**<0.001**
context_c	0.11	0.74	0.15	0.879
cluster_c	-7.52	0.84	-8.92	**<0.001**
voicing_c	8.20	0.87	9.39	**<0.001**
place_c	-5.31	0.77	-6.89	**<0.001**

Table C-179. LMM results for burst intensity in stop-stop clusters as a function of voicing, place (bilabial, alveolar, velar), cluster (/bk/kb/dg/tg/gt), context (C2/C2 baseline), and gender (male/female).

49. The interaction between F0 onset and aspiration

The model: F0 onset ~ rate_c + place_c + voicing_c + sex_c + Aspiration_c + context_c + (1 + rate_c + place_c + voicing_c + Aspiration_c + context_c || speaker) + (1 | word)

Predictors	*Estimates*	*std. Errort*		*p*
(Intercept)	194.81	3.54	55.06	<**0.001**
rate_c	3.52	1.26	2.79	**0.005**
place_c	-2.70	1.86	-1.46	0.145
voicing_c	-26.35	2.09	-12.62	<**0.001**
sex_c	-69.97	6.93	-10.10	<**0.001**
Aspiration_c	-4.99	1.32	-3.78	<**0.001**
context_c	-0.64	1.57	-0.41	0.682

Table C-180. LMM results for F0 onset as a function of voicing, place (bilabial, alveolar, velar), aspiration (UASP: unaspirated/MASP: moderate aspiration/HASP: heavy aspiration), context (initial/medial trochaic/medial iambic), rate (normal/fast), and gender (male/female).