THE PHONETICS AND PHONOLOGY OF ONSET CLUSTERS: THE CASE OF MODERN HEBREW

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Rina Kreitman

May 2008

© 2008 Rina Kreitman

THE PHONETICS AND PHONOLOGY OF ONSET CLUSTERS: THE CASE OF MODERN HEBREW

Rina Kreitman, Ph. D.

Cornell University 2008

The focus of this work is the phonetics and phonology of word initial bi-consonantal onset clusters. I begin by reporting on a cross-linguistic study of the typological distribution of two features, [sonorant] and [voice], in word initial clusters. There are 4 logical possibilities for combining [–sonorant] (O) and [+sonorant] (S) in bi-consonantal onset clusters, OO, SS, OS and SO, and 16 logical combinations of these clusters. A survey of 62 languages was conducted and it was found that of the 16 logically possible language types, only 4 emerge as occurring: {OS}, {OS, OO}, {OS, OO, SS} and {OS, OO, SS, SO}. The following implications were found: SO ⇒ SS ⇒ OO ⇒ OS. The sub-typologies of OO and SS clusters were examined separately.

Next I address the feature [voice] and the distribution of this feature in bi-consonantal onset clusters, focusing on obstruents. Out of the 4 logically possible combinations of [–voice] and [+voice] in obstruent clusters, and 16 possible language types, only 6 emerge as occurring: {[–v][–v]}, {[–v][–v], [+v][+v]}, {[–v][–v], [–v][+v]}, {[–v][–v], [–v][+v], [+v][–v]}, {[–v][–v], [–v][+v], [+v][+v]} and {[–v][–v], [+v][+v], [–v][+v], [+v][–v]} with the following implicational relations:

$$\begin{array}{rcl} & & [+v][-v] \\ & & \Downarrow \\ & & [-v][+v] \\ & & \Downarrow \\ [+v][+v] & \Rightarrow & [-v][-v] \end{array}$$

The different typological patterning of the two features implies that it is impossible to predict the typological patterning of clusters of one of these features, based on the other. A language can be of one type in terms of [sonorant] but of a different type in terms of [voice]. The typological patterning of clusters based on the feature [sonorant] does not provide clues about the phonological patterning of the feature [voice]. In particular, I show that languages treat [+v][–v] and SO clusters differently and argue that solutions that have been proposed for the special phonological representation of SO clusters such as appendix, cannot account for [+v][–v] clusters.

In the second part, I present an acoustic phonetic study of word initial clusters with different laryngeal specifications in Modern Hebrew. I show that all four types of voicing clusters are realized phonetically. I further show that voicelessness is not always an underlying target, but can be a result of voicing failure in an unfavorable phonetic context.

BIOGRAPHICAL SKETCH

Rina Kreitman was born in Hadera, Israel on January 21, 1976. She spent most of her life in Israel and in the US. She spent a year in the US as a child and returned to Israel to complete school. After graduating high-school, she served in the military and upon completion of her military service, she attended the University of Tel-Aviv. She majored in Linguistics as a single major, graduating Magna Cum Laude. She returned to the US to attend graduate school. During 2004, she taught one semester at Bar-Ilan University. She is currently living in Atlanta and working as the Hebrew language coordinator at Emory University.

To my family

ACKNOWLEDGMENTS

Going through the process and learning experience of a Ph.D. – including classes, actual research and writing of the dissertation – is an amazing life experience. In addition to the enormous intellectual progress and the development of analytical, scientific thinking, one experiences personal growth. The personal and individual development was not a lesser experience than the achievement of the actual doctorate degree. The life lessons that I have learned which accompanied my intellectual growth, were no less important and at times much more difficult than the actual studies. Many individuals contributed to my growth and helped shape me into the person I am today. Some people had a more positive influence on my life, others taught me that the world is not all good, but our growth as individuals involves negative experiences as well as positive ones. I would like to take the time to thank the individuals who contributed to my life during these past few years.

First and foremost I would like to thank my family, my father Haim (חיים), my mother Ora (אורה) and my brother Noam (נועם). You are, and always will be my heart and soul. You got me to where I am and you were always a source of strength for me. You gave me everything I have in life and made me who and what I am today. I love you beyond what words will ever be able to convey. To my grandmother Esfir (אספיר), I am sorry you are not here to witness this. To Nelly who came in later, thanks for listening.

To my family, my uncle Wolf, I never got a chance to say how much I owe you and to thank you for all you have done for me, thank you. To my aunt Tanya, who, in her own right, is a remarkable woman, I thank for her role in providing me with support through some tough times. To my cousins Paul and Katya, thank you for your support through the rough times. You and the little ones, Alexei and Mihael – the most precious little people in the world – are my rock. To my cousin Yana, I wish you were closer.

I would like to thank Outi Bat-El for being my inspiration and for planting the love of phonology in me. It is because of you that I went to grad school and it is because of you that I chose to do what I do and to begin this venture.

I thank Draga Zec for picking up the baton, for helping me stick to phonology and bringing me to the finish line. You challenged me in many ways and without you I would not have been able to finish my degree. Nor would I have been able to produce this piece of work which brings me great pride. Thank you for walking me through this difficult road and for bringing me to where I am. You are both an amazing woman and an amazing scholar and I can only look up to you in admiration and be inspired by you. Thank you for your dedication, your hard work, the nights, the weekends and for the effort that you put into this work. It is all greatly appreciated.

To Abigail Cohn, Lisa Zsiga, Sue Hertz and Amanda Miller I thank for their guidance and for helping me through the past several years. Although many times you did not have to help me, you were always there to answer my questions and to sharpen my theoretical understanding of the material. Thank you for your time.

To my friends who accompanied me through this long journey… To those who are still near me and to those who are further away and a little more busy with life. I was blessed with the good fortune of knowing each and every one of you and I am very proud and grateful for the chance to be part of your lives.

To Amiram Ben-Yaacov (עמי בן-יעקב) and Boaz Nachshon (בועז נחשון) for their incredible support system in the first few years.

To Gil Paz (גיל פז), thanks for sticking by me all these years and being an amazing listener. Thanks for helping me psycho-analyze and understand things so much better. You single handedly managed to be my entire support system through some very rough experiences, all the while going through your own rough times and battling your own demons. Thank you for that and congratulations on your own successes.

To Oren Kurland (אורן קורלנד), knowing you is certainly an honor. You are a unique individual and despite our very interesting interactions you were, and still remain an important part of my life. You really helped me grow up and offered me incredible support when I needed it.

To Rimon Barr, I am honored and humbled to know someone as smart as you are. Your realism and sober/calculated view of life was the source of many stimulating conversations, despite our many disagreements, I did manage to convince you on more than one occasion ☺.

To Tali and Aharon Blank (טלי ואהרון בלנק), thank you for all your help during the rough times and thank you for taking care of things for me. To Anat Kapelnikov (ענת קפלניקוב) it was a pleasure knowing you and living with you. You are an amazing individual and I wish you great luck in the future.

It is an honor to know each and every one of you. You are fine individuals and I hope we will all continue to be part of each other's lives and have fruitful relationships.

To the rest of the gang at Cornell, Nir Ezion (ניר עציון), Regina and Eli Barzilai (רגינה ואלי ברזילאי), Ron Hose (רן הוז) Eran Dvir (ערן דביר) – I wish you had more time, Carmel and Veronique Domshlack (כרמי וורו דומשלק), thanks for all the great Friday night dinners. For the other Friday night dinners and other inspirational conversations, I thank Rabbi Eli Silverstein.

To my gang back home… it seems like a different life time and yet you were always there… Uri Karaso (אורי קרסו), Itzik Vakshi (איציק וקשי), Lior Bardugo (ליאור ברדוגו), Dotan Orenshtein (דותן אורנשטיין), Yariv Tidhar (יריב תדהר), Noam Cohen (נועם כהן), Adi Kosta (עדי קוסטה), Erez Stern (ארז שטרן), David Levitan (דוד לויתן)

...הריישים כבר לא ריישים. You have always accepted me for who and what I am and you taught me that I do not need to be anyone but myself.

To my long time friends Keren Fuchs (קרן פוקס), Keren Yoavi (קרן יואבי) and Roneet Gutman (רונית גוטמן) I value your friendship despite the distance. To my friend Karen Petersen, we always proved that closeness is in the heart, here's to the next 14.

To my colleagues, Anastasia K. Riehl, Marc Brunelle, Johanna Brugman, Jiahong Yuan, I thank for their help in shaping my scientific thinking, and for all your advice during the talks and various presentations. Thank you for your comments on this work and all the work that preceded this and led up to the complete product. Eric, thank you for always being in the lab and helping whenever help was needed. The lab was certainly a very important component and thank you for keeping it running smoothly.

Additional mentions of individuals who played a bigger or smaller part in my life in the past several years: Steve, Elana Chipman, Antonio Montalban, Samy Fridman, Andre LeClaire, Miguel, Holly Manslack, Nikos Chalkias, Dianah Angert, Asaf Abramovich, Michal Kuriat.

The following people were an inseparable part of my learning experience and contributed greatly to the "school of life," I would like to thank them for their contribution to my life: Lauren Aronson and Hadar Koter. I thank Sigal Bekker (סיגל בקר) for her enormous contribution to my growth and confidence.

To Jackie, Felipe and Josh I thank for teaching me the tougher lessons in life. Three completely different individuals, yet you all share the same fundamental trait. I also thank David (דוד), for showing me that things can be different.

To Sergeant Joshua D. Blevins I thank for his contribution to this endeavor. You will probably never know or be able to appreciate how important you were to this. Despite the tumultuous times and experiences, without you I probably would not have been able to ever finish this work, you kept me focused at times I needed it and for that I will be forever indebted to you.

I thank you all and am honored to have known such great individuals.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SPECTROGRAMS

CHAPTER 1

INTRODUCTION

1.0. Phonological features of voicing

In this work I focus on the distribution of the feature [voice] in word initial obstruent clusters, focusing both on its phonological distribution and its phonetic realization. The basic assumption guiding this work is that phonological features are abstract units which model and characterize the nature of speech sounds. Phonological features are also, at least partly, anchored by phonetic implementation. But the outcome of phonetic implementation is not always a reliable testimony of the underlying representation of a particular phonological feature. In this work I will explore the interface between abstract phonological features and their concrete phonetic implementation, as well as cases of misalignment between the phonetics and the phonology. In other words, while phonological features have phonetic correlates that result from their physical and physiological implementation, the phonetic realization of abstract phonological features is not always transparent.

I investigate the cross-linguistic distribution of the feature [voice] in word initial bi-consonantal obstruent clusters and propose a typology of possible sequencing of values of the feature [voice] in word initial obstruent onset clusters, which is based on 62 languages from 22 language families. In earlier proposals in the literature, voicing combinations in word initial position were assumed to be restricted to [–voice][–voice], [+voice][+voice] and [–voice][+voice] as the only cluster types that occur in natural language. The fourth logical possibility [+voice][–voice] was excluded as impossible both on phonetic and phonological grounds. Lindblöm (1983) claims that the [+voice][–voice] obstruent onset cluster is not possible phonetically, while Lombardi (1991, 1999) says it is prohibited phonologically and compares it to sonority reversed clusters. However, as I show in chapters 2 and 5, the [+voice][-voice] obstruent cluster has been attested both phonetically and phonologically in at least Khasi, Tsou and Modern Hebrew.

Once I establish that all four types of obstruent clusters, [–voice][–voice], [+voice][+voice], [–voice][+voice] and [+voice][–voice], are occurring clusters, I provide a cross-linguistic typology of these clusters. I also provide a detailed phonetic study of Modern Hebrew in which I investigate how voicing is realized in bi-consonantal word initial obstruent clusters in [–voice][–voice], [+voice][+voice], [–voice][+voice] and [+voice][–voice] sequences.

In this chapter I first give a general background on the feature [voice]. I begin with an overview of proposals to characterize voicing both on phonetic and phonological grounds, with features proposed by Jakobson, Fant and Halle (1952), Chomsky and Halle (1968) and Halle and Stevens (1971), and phonetic implementations of the phonological feature [voice] proposed by Keating (1984). After presenting a detailed discussion of the feature [voice], I focus specifically on the place of [voice] in the phonetics and phonology of Modern Hebrew. I discuss the relevant aspects of the phonology of Modern Hebrew and, due to its complex history, also address its relation with Biblical Hebrew. Finally, I provide an overview of the entire dissertation and its organization.

1.1. The phonology and phonetics of voicing

I begin with the early work by Jakobson, Fant and Halle (1952), in which voicing is characterized as follows:

> The voiced or "buzz" phonemes /b, d, z, v/ vs. the voiceless or "hiss" phonemes are characterized by the superposition of the harmonic sound source upon the noise source of the latter… The most striking manifestation of 'voicing' is the appearance of a strong low component which is represented by the voice bar along the base line of the spectrogram (Jakobson, Fant and Halle 1952:26).

The voiced/voiceless distinction was taken to characterize laryngeal contrasts only for some of the languages while other languages were described as having an opposition between tense and lax: "In consonants, tenseness is manifested primarily by the length of their sounding period, and in stops, in addition, by the greater strength of the explosion" (Jakobson, Fant and Halle 1952:36). In English, German and Welsh, the orthographic *b, d,* and *g* stand for voiced segments but they are realized as phonetically voiceless in most positions and as voiced only intervocalically. Therefore, in these languages the opposition is not between voiced and voiceless consonants, rather, the feature tense/lax is used. Jakobson, Fant and Halle (1952) propose 'voiced' and 'voiceless' as acoustically based features that remain in the phonetic realm and have no impact on the phonology.

Lisker and Abramson (1964) showed that the split of the voiced/voiceless and tense/lax distinctions into two separate independent dimensions is not necessary. They showed that, acoustically, the same cues are responsible for the voiced/voiceless and tense/lax contrasts in stops. In fact, their complementary distribution suggests that they are more closely linked than had previously been thought:

> The two features correlated with voicing and aspiration… have an interesting relation to one another, at least in the case of stops in English; each feature tends to be prominent in the spectrograms only where the other is absent. (Lisker and Abramson 1964:387).

Consequently, they propose to define the degree of voicing in a stop by "the duration of the time interval by which the onset of periodic pulsing either precedes or follows release" (Lisker and Abramson 1964:387). Thus, a single dimension, Voice Onset Time (VOT), can serve to distinguish between voiced/voiceless and tense/lax segments. If the vibration of the vocal folds precedes the release burst, the segment is perceived as voiced and is said to have a negative VOT. If the vibration of vocal folds begins after the burst, the sound may be perceived as voiceless unaspirated (short lag VOT); if the period of the initiation of the vocal fold vibration is quite lengthy, then a segment is said to be aspirated (long lag VOT). Thus, VOT is a continuum and the voicing of the stop depends solely on where along the continuum the burst occurs in relation to the initiation of vocal fold vibration. Lisker and Abramson's proposal to quantify voicing using VOT captured the acoustic phonetic reality of the production of voicing. However, this phonetic characterization had no implications in terms of phonological or featural categorization of voiced and voiceless segments. The ability to measure, or quantify, voicing had little bearing on the actual phonological representation of voicing and voicelessness.

Chomsky and Halle (1968) translate the phonetic realization of voicing into binary features and propose the features [voice] and [spread glottis], each of which can have either a negative or a positive specification. The phonological features are part of the grammar and grammatical knowledge of the speaker. These underlying phonological features, which are part of the linguistic knowledge of the speaker, are translated into the output phonetic form. Although Chomsky and Halle (1968) propose binary values for these features, their proposed features are grounded in phonetics. Phonological rules are responsible for the output form and therefore can change the binary value of the feature, but it is the phonetic rules which determine the quantitative values of the features along a phonetic scale.

Halle and Stevens (1971) translate the phonetic differences between voiced and voiceless segments into phonological categories by proposing four phonetically motivated features, each of which can be specified as negative or positive. They propose the following features, all grounded in the physiological realization of voicing: [±spread glottis], [±constricted glottis], [±stiff vocal cords] and [±slack vocal cords]. These features describe the state of the vocal folds at the moment of the release of the stop. They yield a typology of nine distinct phonological categories,[1] which characterize the glottal specifications for all segment classes including vowels and sonorants.[2] While these four features and their binary values produced all possible laryngeal distinctions for all stop consonants, they failed to account for the observation that no language distinguishes more than four categories of laryngeal states, voiceless aspirated, voiceless unaspirated, voiced and voiced aspirated stops. Their system produced a large range of distinctions that are not attested in natural languages. That is, although languages can distinguish a variety of glottal settings, including breathy voice and creaky voice, Halle and Stevens fail to account for the fact that many distinctions cannot combine within a single language. No language distinguishes murmur, creaky voice and breathy voice. Thus, while Halle and Stevens' features describe a phonetic reality grounded in articulatory phonetics, they do not account for the much narrower range of phonological distinctions that languages actually exploit. While they address the phonetic properties of sets of sounds, they do not address their phonological distribution or their phonological categorization.

---

[1] Although there are more logical possibilities for the combinations of these features, only nine combinations are physiologically possible to produce.

[2] For Halle and Stevens (1971) their proposed features can also account for glottal settings such as creaky and breathy voice in vowels as well as glottalized segments.

Keating (1984) draws the much needed line between *phonetic categories*, which are, among other things, used to characterize allophones, and *phonological representations*, which reflect the organization of sounds into natural classes. In the mapping from phonology to phonetics, a single phonological feature may be phonetically implemented in a number of different ways. She adopts a proposal by Liberman (1970, 1977) to use the feature [±voice] "as a binary phonological feature which can be implemented differently in different languages along the continuous dimension of V[oice] O[nset] T[ime]" (Keating 1984:290). Keating demonstrates how cross-linguistic differences in phonetic implementation are merely differences in realization of the same phonological category. Different languages choose different strategies to implement phonetically the same phonological distinctions. She contrasts Polish where the feature [+voice] is realized as negative VOT and the feature [−voice] is realized as a short lag VOT, with English, where the same phonological features are granted different phonetic implementation. In English, Keating claims, there is a phonological distinction between [+voice] and [−voice], but while the former is realized as a short lag VOT, the latter is realized as a long lag VOT. That is, the voicing distinction in English is between voiceless unaspirated and voiceless aspirated segments. Thus, languages choose where to draw their own categorical boundaries along the VOT continuum. The phonology of the language determines how to map the feature [voice] onto the VOT continuum, or in other words, where to draw the phonological boundaries. The diagram in figure (1.1) illustrates the mapping of the phonetic implementation of the phonological features according to Keating's proposal:

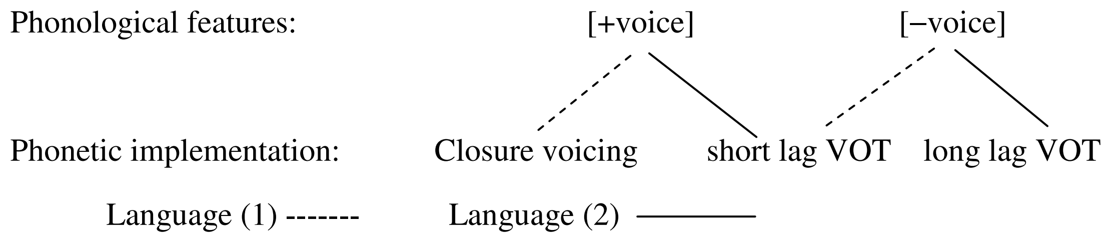


Figure 1.1: Mapping of voicing in two different languages.

While a language may distinguish only two categories, there may be several ways to realize the two phonological categories phonetically, as shown in Figure (1.1). In language (1), marked with a dashed line, voiced segments are realized with a negative VOT, while voiceless segments are realized with short lag VOT. In this language long lag VOT may also be mapped to [−voice]. A language of this type is Polish. In language (2), marked with a solid line, segments with short lag VOT are considered voiced while segments with long lag VOT are voiceless. German is claimed to be a language of this type. Languages may choose to have more than a two way distinction, and may have up to four [voice] distinctions. This issue, however, will not be pursued in this work.

According to Keating (1984), phonetic differences in the realization of phonological features should not be coded in the phonology. A phonological system should code only those distinctions necessary to create contrast in a language. Encoding phonetic differences yields an overly complex system, which fails to account for cross linguistic patterns and only reflects the phonetic reality of individual languages.

Following Keating (1984), distinctions in obstruent stops for languages such as German and English were understood to be the same as those in languages such as Dutch, Hungarian, Spanish and Russian. In all these languages the single phonological feature responsible for voicing distinction is [voice] but different languages have different phonetic implementation of this single feature. In German and English the phonological feature [voice] is phonetically realized as short lag VOT whereas voiceless consonants have long lag VOT. Hall (1993), Kingston and Diehl (1994), Lombardi (1991, 1995a, 1999) and Wiese (1996) all treat German as having a [voice] feature distinction. Kingston and Diehl (1994) argue, based on acoustic correlates of stop consonants established by experiments they had conducted, that even when the feature [voice] is distinguished by short lag VOT versus long lag VOT acoustically, phonologically the short lag VOT patterns with negative VOT and long lag VOT patterns with positive VOT. In other words, phonologically, segments that are realized with a short lag VOT in English belong to the same class as segments realized with negative VOT in other languages, while segments realized with long lag VOT belong to the same class as segments realized with positive VOT in other languages.

However, the view that German stops are distinguished by the feature [voice] has been challenged in a body of literature. If, in German and in English, segments that are transcribed as *b*, *d* and *g* are rarely realized with closure voicing, and if voicing in these languages surfaces only in a small subset of environments, such as intervocalically, why should the feature [voice] be assumed for these languages? That is, it is possible that in these languages some feature other than [voice] is responsible for the distinction between *p, t, k* and *b, d, g*. Moreover, does the feature [voice] correctly capture the phonological distinctions and sound patterning in this case?

Iverson and Solomon (1995), Jessen (2001), Jessen and Ringen (2002) and Petrova *et al*. (2006), as opposed to Keating (1984), claim that the feature [voice] does not reflect the phonological and phonetic reality of laryngeal distinctions in stops for languages such as German. They postulate the feature [spread glottis] for German, as the relevant distinctive phonological feature.[3] That is, according to Jessen (2001) and Jessen and Ringen (2002), it is not the case that all languages have a phonological opposition between [+voice] and [–voice]. Some languages use a different feature to distinguish the set *p, t, k* from the set *b, d, g*. In these languages, the distinctive feature is [spread glottis], not [voice].

Jessen (2001) and Jessen and Ringen (2002) thus, support the view[4] that for English and German, *b*, *d*, *g* and *p*, *t*, *k* are distinguished by the feature [spread glottis], originally proposed by Halle and Stevens (1971), rather than by the feature [voice]. Languages are divided into "voice" languages and "spread glottis" languages. Languages that use negative VOT as the basic cue for voicing are "voice" languages and languages that use aspiration as the basic cue are "spread glottis" languages. If intervocalic voicing occurs in "spread glottis" languages, it is interpreted by Jessen and Ringen (2002) as "passive voicing" and the vibration of the vocal cords is attributed to involuntary action on the speaker's side. That is, the vocal folds vibrate naturally in intervocalic position because, in vowels, voicing is more natural due to the pressure differences between sub-glottal and supra-glottal cavity characteristic of vowels. Therefore, the speaker does not need to invest any effort in vibrating the vocal cords; the vibration carries on from the vowel into the obstruent and is entirely passive. Jessen and Ringen explain that

> [T]he categorical difference between [spread glottis] and non-[spread glottis] stops in English and German is not so much in terms of presence *vs.* absence of glottal opening as in terms of active (and large) glottal opening *vs.* passive (and slight) glottal opening (Jessen and Ringen 2002:192).

[3] The claims about English are slightly different. Petrova *et al*. (2006) claim that English is a mixed language and that both [voice] and [speard glottis] features are necessary to accurately describe the phonological system and the phonological contrasts in English.
[4] Other proponents of this view include Iverson and Salmons (1995) and Petrova *et al*. (2006) *among others*.

However, Jessen's (2001) and Jessen and Ringen's (2002) proposal that the distinguishing phonological feature in German (among other Germanic languages) is [spread glottis] rather than [voice], and that voicing in intervocalic position is "passive voicing" requiring no laryngeal adjustment, encounters problems in light of a wide body of literature which offers experimental counter evidence to this proposal.

In this body of literature several authors demonstrate muscular activity in the production of *b, d, g* in intervocalic and utternace initial position, counter to Jessen and Ringen's claims. Kingston and Diehl (1994:427) show that "F0 is consistently depressed in vowels next to [+voice] stops, regardless of whether a language, or context employs pre-voiced or short lag stop as the realization of this phonation type." They further claim that "… the majority of Swedish and English speakers close the glottis substantially before the stop release in initial [+voice] stops" (Kingston and Diehl 1994:429), referring to the findings by Lindqvist (1972) and Flege (1982). Sawashima and Hirose (1983:23) note that "[t]he reciprocity between the PCA and the adductors has been observed for different languages, including American English, Japanese, Danish and French." Sawashima and Hirose (1983) show that in intervocalic position the phoneme *b* is produced with muscular activity of the INT muscle; a fact which counters Jessen and Ringen's claims that intervocalic voicing is produced with no active muscular movement but rather passively relying on a favorable phonetic environment. Fujimura and Sawashima (1971) show muscular activity in the inter-vocalic sequences of [tt], [td] and [dd] in English, which offers further counter evidence to Jessen and Ringen's claims about "passive voicing" in intervocalic position.[5]

[5] It must be noted that the [tt], [td] and [dd] sequences that Fujimura and Sawashima (1971) tested, contained a morpheme boundary within all sequences.

Jessen and Ringen's (2002:191) claims stem from their fundamental assumption that "phonological features should be grounded in phonetic reality." However, their specific view on this important issue mandates the use of a large number of phonetically based features which do not reflect the phonological reality and may yield more phonological distinctions and contrasts than are typologically attested. They claim that such an analysis is able to account for the lack of voicing in utterance initial and final position in languages such as German and accounts for the presence of voicing intervocalically. However, the use of the additional phonetic feature [spread glottis] is not the only way to account for the lack of voicing in English stops word initially. A phonological [+voice] target may be realized as phonetically voiceless in environments which are phonetically less favorable. That is, although the phonological target for an onset segment may be [+voice], the failure to implement voicing phonetically, due to insufficient pressure, can result in a phonetically voiceless segment. This type of voicelessness stems from failure to reach a minimum pressure threshold which is required for initiation of vocal fold vibration. This proposal will be discussed in detail in chapter 4.

The question whether the feature [voice] is sufficient to characterize all the voicing distinctions cross-linguistically or whether an additional feature, [spread glottis], is also necessary, is outside the scope of this work and will not be addressed here. Languages which may be controversial in regards to the underlying distinctive feature in obstruents have been excluded from the cross-linguistic survey for the feature [voice] presented in chapter 2. Modern Hebrew, as will be shown in this work, is not controversial in terms of the feature [voice]. Modern Hebrew is clearly a [voice] language and distinguishes between closure voicing prior to the burst in the phonetic realization of *b*, *d*, *g*, and aspiration, either short or long lag VOT, in the realization of *p*, *t* and *k*.

1.2. The place of voicing in Modern Hebrew sound system

An important focus of this dissertation is the phonetics and phonology of voicing in Modern Hebrew obstruent clusters, addressed in chapters 3 and 5. Modern Hebrew is to be understood as Modern Israeli Hebrew, the language spoken in the state of Israel and described by Blanc (1964) and Chayen (1972, 1973) as general Israeli Hebrew (also discussed in Bolozky 1972, 1978). I do not discuss the Oriental dialect of Modern Hebrew spoken in Israel (Devens 1980).

In this section, I outline the relevant aspects of the sound patterns of Modern Hebrew, in order to provide background for chapters 3 and 5. In section 1.2.1 I discuss the segmental inventory of Modern Hebrew, and in section 1.2.2 I discuss possible syllable structures in Modern Hebrew. In section 1.2.3 I discuss restrictions on onset clusters and present possible onset clusters in Modern Hebrew. To account for the onset clusters in Modern Hebrew which are an innovation, and did not exist in Biblical Hebrew, I briefly compare the two systems and point out the inconsistencies between Biblical and Modern Hebrew, and the difficulties in accounting for them. In section 1.2.4 I discuss the influence of Biblical Hebrew on Modern Hebrew. Finally, in section 1.3, I present the organization of the dissertation.

1.2.1. Segmental inventory of Modern Hebrew

We begin the discussion of the segmental inventory in Modern Hebrew with the vocalic inventory, which is presented in table (1.1).

Table 1.1: Phonemic vowels in Modern Hebrew

	Front	**Back**
High	i	u
Mid	e	o
Low		a

There are five phonemic vowels in Modern Hebrew: /i, e, a, o, u/. Vowel length is not phonologically distinctive although it occurs phonetically. Table (1.2) lists the consonantal phonemic inventory of Modern Hebrew.

Table 1.2: Phonemic consonants in Modern Hebrew.

	labials	**coronals [+ant]**	**coronals [–ant]**	**dorsals [+hi]**	**Glottal**
stop	p b	t d		k g	ʔ
fricative	f v	s z	š ž	x	(h)
trill				**r(?)**	
nasal	m	n			
affricates		ts	tš dž		
lateral		l			
approximant		j		**r(?)**	

Before turning to voicing in obstruents, I first provide some general comments on the consonantal inventory. The sounds *ž, tš* and *dž* are not native to Modern Hebrew and are found only in borrowed words. The sound *h* exists for some speakers but is frequently realized as the glottal stop. The nature of *r* in Modern Hebrew is unclear. It has been described in the literature as a velar or uvular, as a fricative, a trill or an approximant (Berman 1997, Blanc 1964, Bolozky 1978, Chayen 1973, Laufer 1983). This sound varies in its phonetic realization and the exact nature and place of articulation of *r* is yet to be determined (see Kreitman and Bolozky 2007). It seems to be behaving as a sonorant in certain phonological environments, but sometimes behaves as a fricative and participates in voicing assimilation, which is typical for obstruents. For convenience, it will be transcribed as *r* since the precise phonetic transcription has yet to be determined; this does not influence the discussion or arguments presented in this work.

Modern Hebrew distinguishes between voiced and voiceless obstruents. As shown in Table (1.2), all voiceless obstruents other than the uvular fricative *x* and the affricate *ts* have voiced counterparts. Generally, the [+voice] feature is realized as closure voicing in most environments. In utterance initial position the duration of closure voicing depends on place of articulation, but some closure voicing is usually present in all voiced stop consonants. Laufer (1995) found that, in Modern Hebrew stop consonants, the relevant feature for distinguishing *p*, *t*, *k* from *b*, *d*, *g* is [voice], since closure voicing preceding the burst of *b*, *d* and *g* is clearly present in all singleton forms. In fact, Laufer (1995) clearly demonstrates that the [spread glottis] distinction is inadequate for Modern Hebrew. His findings will be discussed in chapter 5.

1.2.2. Syllables in Modern Hebrew

Possible syllable types in Modern Hebrew are listed in (1). In the list below V stands for the syllable nucleus, and C for the syllable margin. As already noted, there are no phonologically long vowels in Modern Hebrew.

1)

Syllable type	Sample word	Meaning
CV	*ba*	‘came’
CVC	*kal*	‘easy’ (masc.)
CVCC	*gart*	‘you lived’
CCVC	*sfog*	‘sponge’

The nucleus of the syllable is always vocalic; there are no syllabic consonants. The nucleus is always bound on the left by an obligatory onset, and is sometimes bound on the right by an optional coda. There is no evidence to suggest that coda consonants contribute to syllable weight. Coda consonants never cause stress shift and do not participate in any other phonological processes. Therefore, I assume, following Bat-El (1989), that in Modern Hebrew all syllables are mono-moraic, with vowels occupying the single moraic position.

In Modern Hebrew any consonant can serve as an onset; *ž*, *tš*, and *dǰ* appear only in borrowed words. Both complex onsets and complex codas occur in Modern Hebrew. However, complex codas are subject to more restrictions than complex onsets. Possible complex onsets may contain up to two members in native words and up to three members in borrowed words. Complex onsets occur much more frequently in word initial positions although intervocalic consonantal sequences, which result in complex onsets and codas, exist in borrowed words and denominative verbs. An example of the word *control* 'ctrl key' where *t* and *r* form a word medial onset cluster is illustrated in figure (1.2).

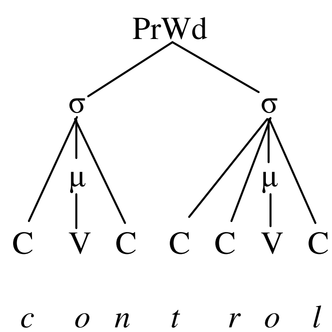


Figure 1.2: Prosodic structure of the borrowed word *control.*

Sequences of two consonants that appear intervocalically, for example in *kalmar* 'pencil box', are not tautosyllabic and are syllabified as $C_1VC_2.C_3VC_4$. Triconsonantal intervocalic sequences occur mostly in borrowed words and denominative verbs. When appearing word initially, these sequences are syllabified as onset clusters. In the borrowed word *šprits* ‘squirt’, the first three consonants are all syllabified as part of a complex onset.

1.2.3. Restrictions on onset clusters in Modern Hebrew

The only onset clusters permitted in Modern Hebrew are [–sonorant][+sonorant] and [–sonorant][–sonorant] clusters. *[+sonorant][+sonorant] and *[+sonorant][–sonorant] clusters are strictly banned in Modern Hebrew, as reflected in table (1.3); this will be detailed in chapter 2. Sonorants may combine with almost any obstruent, as long as the combination creates a rise in sonority.[6] When [+sonorant][+sonorant] or [+sonorant][–sonorant] sequences arise as a result of morphological processes such as affixation, they are broken by vowel epenthesis.

Table 1.3: Permissible clusters in Modern Hebrew.

Second member																		
First member	**p**	**t**	**k**	**f**	**s**	**š**	**x**	**b**	**d**	**g**	**v**	**z**	**ts**	**l**	**m**	**n**	**r**	**y**
p		+	+		+	+	+		+	+		+	+	+		+	+	
t			+	+	+	+	+			+	+	+		+	+	+	+	
k		+		+	+	+	+		+		+	+	+	+	+	+	+	
f		+												+		+	+	
s	+	+	+	+			+		+	+	+			+	+	+	+	
š	+	+	+	+			+	+	+	+	+	+		+	+	+	+	
x		+															+	
b		+	+		+	+	+		+	+		+		+		+	+	
d			+	+		+	+			+	+			+	+		+	
g				+	+	+	+		+		+	+		+	+	+	+	
v														+			+	
z			+				+	+	+	+	+			+	+	+	+	
ts				+			+	+	+		+			+	+	+	+	

[6] By rise in sonority I mean that the value for the feature [sonorant] must be positive in C2 and negative in C1. That is, C2 which is marked [+sonorant] is more sonorous than C1, which is marked [–sonorant].

However, despite rigid restrictions on onsets in terms of the feature [sonorant], onset clusters in Modern Hebrew are not restricted in terms of the feature [voice]. All possible voicing combinations exist in Modern Hebrew onsets; [–voice][–voice], [+voice][+voice], [–voice][+voice] and [+voice][–voice] clusters are all acceptable clusters. In other words, voiced obstruents may combine with voiceless obstruents as either first or second member of a cluster. Examples such as *dk* as in *dkalim* 'palms' and *gf* as in *gfanim* 'vines' are quite common and are not considered exceptional at all.

Place restrictions are evidenced in onsets, and exclude all onset clusters that share a labial place of articulation. When such clusters arise in borrowed words, one of the consonants is deleted as in the name *Pfeifer* borrowed as [faifer]. Coronals are not as restricted as labials and onset clusters which share coronal place of articulation are quite common. They include clusters such as *tn* in *tnuxa* 'position/posture', *tl* in *tlaʔi* 'patch' and *dl* in *dli* 'bucket', as well as the obstruent clusters *st* in *stani* 'devilish', *sd* in *sdurim* 'organized by number' and *zd* in *zdoni* 'evil'. Note the asymmetry in the behavior of coronals, where *t* may combine with all sonorants, as in *tmuna* 'picture', *tnuʔa* 'movement' and *tluš* 'receipt', while its voiced counterpart *d* may only combine with *m* in *dmaʔot* 'tears' or with *l* in *dlifa* 'leak' but not with *n* (**dn*). Clusters which share a dorsal place of articulation such as *kx* in *kxulim* 'blue (pl. masc.)' and *gx* in *gxonot* 'bellies', also exist.[7]

1.2.4. Historical Background

In this section we will address aspects of the phonology of Modern Hebrew crucial for the arguments made in this dissertation. I will also present the relevant historical

[7] The sequence *kx* is not an affricate in Modern Hebrew but is considered a cluster as defined in chapter 2.

background in the development of Modern Hebrew that sheds light on its phonological organization.

Modern Hebrew is classified as a West Semitic language belonging to the Canaanite branch of the central southern Semitic languages (Faber 1997). Modern Hebrew is the official language of the state of Israel spoken by approximately six million people natively, and is genetically classified together with Biblical Hebrew. It shows heavy influences from Indo-European languages such as Russian and Yiddish.

Assuming that Biblical Hebrew and Modern Hebrew are different stages in the development of the Hebrew language, we identify four historical periods, as in (2):

2)

(i). Biblical Hebrew (1300 to 200 BCE)
(ii). Mishnaic or Rabbinical Hebrew (until about 800 CE) crucial for
(iii) Medieval Hebrew the phonology of
(iv) Modern Hebrew Modern Hebrew.

There was a gap of roughly 2 millenia between the end of the Biblical period and the beginning of Modern Hebrew. During this long period, Hebrew was used as a literary and liturgical language and was not used as a spoken language. It was taught mostly for religious purposes but had no native speakers.

The historical period that is most relevant for the formation of Modern Hebrew phonemic inventory and syllable structure is Biblical Hebrew, which will be the focus of this discussion. The intermediate periods, Mishnaic and Medieval Hebrew, had a less prominent effect on the structure of Modern Hebrew, and while they are mentioned here, they will not be discussed in detail.

1.2.4.1. The origins of Modern Hebrew

The dialect of Modern Hebrew discussed in this work, sometimes referred to as General Israeli Hebrew (Blanc 1964, Chayen 1972), evolved with the nationalist Zionist movement that began to resettle the area that roughly corresponds to modern day Israel. With the establishment of the State of Israel in 1948, Hebrew became the official national language of the state. Prior to becoming the official language of the state of Israel, Hebrew was considered a liturgical language spoken only for religious purposes in the Diaspora.[8] For all other purposes it was extinct. As a liturgical language Hebrew had no native speakers and did not have a spoken continuum for several hundred years.

Towards the end of the 19th century the Hebrew language was revived by the Zionist movement, which began to resettle Israel, mostly for political, nationalistic and ethnic reasons. With migrations mainly from Eastern Europe (Russia, and later Poland and Germany) the Zionist movement began as a reaction to various political changes and activities that were taking place in Europe. With increasing persecution of Jews across Europe, Jewish people began to flee their native countries looking for a place where they can be safe. They chose to flee to Israel, the historical land of the Jewish people. Upon arrival in Israel, the immigrants felt they needed a stronger bond with their Jewish roots as well as a common language in which Jews migrating from various countries could communicate – a lingua franca. Some of the immigrants believed that Hebrew should be the official language of the Jewish people, mostly for historical reasons, and thus various groups and organizations began the task of reviving the Hebrew language. The linguistic background of the revivalists played a fundamental role in the shaping of the structure of the language. Since the new

[8] While all dialects of Hebrew were based on reconstructed Biblical Hebrew, they varied across local Jewish communities. For example, the dialect spoken in France was different from the dialect spoken in Poland or that spoken in Morocco.

immigrants to Israel were of Eastern European decent and spoke mostly Russian or Polish, these languages served as a substratum for the new language. Many phonetic and phonological characteristics of those languages seem to carry over to Modern Hebrew, particularly the wide variety of consonant clusters that are characteristic of these languages.

Already at the onset of the revival project, it was clear that Biblical Hebrew could not be used as a daily language as it was "too poor" and missing much of the vocabulary needed for daily life. Language enthusiasts began the process of language revival through language planning and set up guidelines by which they planned to enrich the Hebrew vocabulary. The language planners established principles for the creation of new words, and for borrowing new vocabulary items into the language (Aloni-Feinberg 1978).

Many of the linguistic principles and theories of language planning contemporary at the time of the revival were adopted by the language enthusiasts. Different suggestions for guidelines regarding new word creations and borrowings were put forth by prominent figures. Some of the suggestions were widely accepted, but some of the proposed principles were highly debatable. It was widely agreed that as much of the original biblical vocabulary as possible will be used in the new and revived Modern Hebrew. Missing vocabulary items, that is, words that did not exist in the Bible but were needed in the language were supplemented by other sources such as religious texts like the *Talmud* and the *Mishna*, as well as other later sources such as medieval poetry. Words that were missing from those additional sources needed to be created. New words were formed using already existent forms, i.e. nouns were created from verbs, and vice versa. Missing adjectives were composed from phonologically similar nouns or nouns close in meaning (for example, *ʔafor* 'grey' from *ʔafar* 'ashes').

Borrowing from outside sources was also encouraged by some enthusiasts as long as the borrowing could be altered to fit as closely as possible to the original Biblical Hebrew forms, that is, as long as they complied with the linguistic principles of Biblical Hebrew. The revivers were split as to the languages they preferred to borrow from. While borrowing from Arabic was greatly disfavored, borrowing from other Semitic languages such as Aramaic was less favored than borrowing from European languages such as English and other Jewish languages, Ladino and Yiddish. For a detailed discussion of Modern Hebrew language planning principles, see Aloni-Feinberg (1978).

Modern Hebrew, thus, is an amalgamation of old and new, native and foreign, resulting in linguistic inconsistency and at least some grammatical confusion. A quick survey reveals that although all four stages in the development of Modern Hebrew listed in (2) contribute to the lexicon, it is Biblical Hebrew and the revived Modern Hebrew that are most influential. In table (1.4) we see that the most influential linguistic layer is the Biblical Hebrew period. Although it contributes only 22% of the lexicon, Biblical Hebrew vocabulary exhibits a 65% frequency in written texts. These numbers suggest that the influence of Biblical Hebrew on Modern Hebrew is considerable and cannot be ignored. We also see from table (1.4) that the two intermediate periods, Scholars' literature and Medieval literature, had a minimal effect on the lexicon (Schwartzwald - Rodrigue 1995)

.

Table 1.4: Origin of Modern Hebrew lexicon.

	% of words originating at period	Frequency of words in written text
Biblical Hebrew	22%	65%
Scholars' literature (Talmud, Mishna)	21%	16%
Medieval Literature	17%	5%
Modern Hebrew	40%	14%

Recall that the principles guiding the language revivers included heavy reliance on Biblical Hebrew texts. As a consequence, they adopted the phonological principles of Biblical Hebrew. However, the phonemic inventory of Biblical Hebrew was only partially revived as the speakers of the newly revived language were all immigrants, mostly from Eastern Europe, and did not fully acquire the phonology of the revived language. As a result, many phonemic distinctions which existed in Biblical Hebrew were lost in the revived Modern Hebrew.

Of relevance to this work are two main differences between Biblical and Modern Hebrew. Firstly, Biblical Hebrew did not admit word initial consonantal clusters. When such clusters arose as a result of morphological processes, they were broken by vowel epenthesis. Modern Hebrew, on the other hand, tolerates a wide range of consonantal clusters word initially, possibly due to heavy influence of languages rich in clusters such as Russian and Yiddish. Therefore, we may deduce that clusters are an innovation of Modern Hebrew. The syllable type CCVC, which did not exist in Biblical Hebrew, is quite common in Modern Hebrew.

The emergence of word initial onset clusters interacts in a complex manner with spirantization rules that existed in Biblical Hebrew. In Biblical Hebrew the spirantization rule in (3) was responsible for the distribution of stop and fricative obstruents.

3) Spirantization (Bolozky 1978, Idsardi 1998)

A stop becomes a fricative after a vowel (except for geminates which are unaffected).

Thus, in Biblical Hebrew, stops never appeared finally and fricatives never appeared initially. Stops and fricatives were in complementary distribution and their distribution was determined by the spirantization rule. The spirantization rule was adopted into Modern Hebrew by the language revivers but the phonemic inventory of Modern Hebrew was different from that of Biblical Hebrew. Table (1.5) illustrates the differences in the phonemic inventory between Modern Hebrew and Biblical Hebrew. The phonemes that are distinct in either Biblical or Modern Hebrew are bolded. Observe that both *f* and *v* existed in Biblical Hebrew, but in Biblical Hebrew these two segments (as well as *x*) were allophonic variations of their stop counterparts, whereas in Modern Hebrew they both have a status of independent phonemes (Rosen 1957). Conversely, *q*, *w*, *ħ* and ʕ, all of which were distinct phonemes in Biblical Hebrew, no longer exist in Modern Hebrew.

Table 1.5: Phonemic consonants in Modern and Biblical Hebrew compared.

Modern Hebrew	**labials**	**coronals [+ant]**	**coronals [–ant]**	**dorsals [+hi]**	**Glottal**
stop	p b	t d		k g	ʔ
fricative	**f v**	s z	š ž	**x**	(h)
trill					
nasal	m	n			
affricates		ts	**tš dž**		
lateral		l			
approximant		j		r	

Biblical Hebrew	**labial**	**coronals [+ant]**	**coronals [-ant]**	**dorsals [+hi]**	**dorsals [+bk]**	**dorsals [+low]**	**glottal**
stop	p b	t d		k g	**q**		ʔ
fric.		s z	š			**ħ ʕ**	h
trill		r					
nasal	m	n					
affr.		ts					
lateral		l					
approx		j		**w**			

The distinct phonemic system of Modern Hebrew caused a breakdown in the spirantization rules. In Modern Hebrew spirantization rules could not apply in the same contexts in which they applied in Biblical Hebrew. With the breakdown of the spirantization rules, which could not apply in Modern Hebrew, the distribution of stops and fricatives changed. Fricatives now appear in environments they could not appear in Biblical Hebrew. In Modern Hebrew fricatives may appear word initially and stops may appear word finally, a distribution which was not tolerated in Biblical Hebrew.

Most importantly, it is highly relevant for this work that clusters are allowed in Modern Hebrew and that the distribution of stops and fricatives cannot be predicted straightforwardly as it could be in Biblical Hebrew. However, the significant influence of Biblical Hebrew on Modern Hebrew, coupled with the breakdown of the spirantization rules, causes some difficulty for a straightforward analysis of permissible clusters in Modern Hebrew. The interaction of manner of articulation within clusters is outside the scope of this work. Of interest here is the fact that in word initial clusters we find all possible voicing combinations [–voice][–voice], [+voice][+voice], [–voice][+voice] and [+voice][–voice] but only two allowable combinations for the feature sonorant [–sonorant][+sonorant] and [–sonorant][–sonorant].

Possible clusters in Modern Hebrew will be discussed briefly in chapter 2, where I provide a cross linguistic typology of onsets in terms of the features [sonorant] and [voice]. I also discuss Modern Hebrew in chapter 3, where I provide phonological evidence against treating [+voice][–voice] clusters as sonority reversed clusters, and in chapter 5, where I provide a detailed phonetic study of voicing in Modern Hebrew obstruent clusters.

1.3. Organization of the dissertation

The dissertation will be organized as follows: in chapter 2 I present a cross-linguistic typological study in which I investigate the distribution of the feature [sonorant] and the feature [voice]. Sixty two languages from twenty two language families, which allow word initial onset clusters, were included in the survey. Results of the survey show that the patterns of typological distribution of clusters in terms of the feature [sonorant] are different from the patterns of the typological distribution of the feature [voice]. Therefore the two features, although may interact in important ways, are not dependent on one another and cannot be reduced to a single pattern. In chapter 2 I also argue for the existence of the cluster [+voice][–voice], which has been claimed by some (Lindblöm 1983, Lombardi 1991) to be an impossible cluster. I provide both phonological evidence from grammatical descriptions of at least four unrelated languages as well as phonetic evidence for at least three of these four languages. Most importantly, I demonstrate that such clusters exist in Modern Hebrew.

In chapter 3, I argue for the cluster status of [+voice][–voice] obstruent clusters in Modern Hebrew. I discuss two morpho-phonological processes in Modern Hebrew that result in clusters, plural affixation in the Segolate noun class and diminutive reduplication. I show that [+voice][–voice] sequences are "real" clusters and must be treated as tauto-syllabic. I argue that proposals to analyze [+voice][–voice] clusters as appendices, in parallel with sonority reversed clusters, cannot account for the syllabification of [+voice][–voice] clusters in Modern Hebrew.

In chapter 4 I discuss the phonetics of voicing. That is, I give a detailed description of the physiology of voicing and argue that there is no phonetic motivation to exclude [+voice][–voice] clusters, as claimed by Lindblöm (1983). I argue that [+voice][+voice] clusters are physiologically more difficult to realize than [+voice][–voice] clusters, and that, therefore, there is no physiologically grounded reason to rule out [+voice][–voice] clusters from the phonetics. I discuss the physiological difficulty of producing voicing in various prosodic positions and show that word initial position is a physiologically disfavored environment for the production of voicing. I then discuss voicing in clusters and, more specifically, voicing in word initial obstruent clusters. Word initial clusters are a doubly disfavored environment for voicing for obstruents; I address the complexity in realizing voicing in each member of the cluster. I end the chapter with predictions about possible phonetic realizations of various voicing combinations in word initial obstruent clusters, taking into account the phonetic disadvantage of these phonological positions. These predictions are tested in chapter 5.

In chapter 5 I present the results of a detailed phonetic study conducted on voicing in various consonant clusters. I use Modern Hebrew as a case study since it allows all possible voicing combinations. I show the various phonetic realizations of each possible voicing combination. I further subclassify all possible voicing combinations and all possible stop and fricative combinations, and cross-reference them with each other. The resulting sixteen cases rule out the possibility that manner of articulation influences the overall existence of voicing in various clusters.

To measure voicing I use various phonetic cues associated with voicing such as voice bar and *f*0. I show that [+voice][–voice] are realized as such phonetically and that voicing can be, and is, realized in C1. I argue that in those cases where [+voice][–voice] clusters surface as phonetically voiceless, that is, in those cases where both members of the cluster are realized with no voicing, it is in fact failure to achieve voicing in word initial position that is responsible for the surface [–voice][–voice] realization, rather than systematic phonological assimilation. Thus, surface [–voice][–voice] clusters are not always underlyingly voiceless; rather, in some cases, the surface representation of these clusters must be attributed to phonetic failure to achieve voicing. In order to determine the underlying representation of a surface [–voice][–voice] clusters, such clusters must be considered in a more comprehensive context and as part of a larger paradigm.

CHAPTER 2

TYPOLOGY OF ONSET CLUSTERS

2.0. Introduction

In this chapter I focus on the typologies of biconsonantal onset clusters along two dimensions. First, I present the typology of onset clusters in terms of the feature [sonorant] in subsection 2.1. In addition to the broad typology based on the possible combinations of obstruents (O) and sonorants (S) within biconsonantal syllable onsets, I also focus on the subtypologies of obstruent-obstruent (OO) and sonorant-sonorant (SS) clusters. Next, I present a typology of voicing specifications within onset clusters in subsection 2.2. It is shown that despite claims in the literature to the contrary, the rare cluster type [+voice][−voice] does occur cross linguistically. The typologies I propose, are a result of a cross linguistic survey that I conducted, which includes 62 languages from 22 language families.

I chose to focus on these typologies because there have been claims in the literature that the patterning of onset clusters in terms of [sonorant] on the one hand, and in terms of [voice] on the other, are closely correlated (Lombardi 1991, Morelli 1999, Steriade 1997). My own findings, to be reported in this chapter, do not support this position. Rather, I show that the organization of onset clusters in terms of [sonorant] follows a different pattern from the organization of onset clusters in terms of [voice]. I further show that the claim that [+voice][−voice] clusters are closely correlated with SO clusters (Lombardi 1991, Morelli 1999) is untenable.

2.1. Clustering of sonorants (S) and obstruents (O)

In word initial, bi-consonantal onset clusters there are four logical combinations of obstruents (O) standing for [–sonorant] consonants, and sonorants (S), standing for

[+sonorant] consonants. The four logical possibilities for combining obstruent (O) and sonorant (S) consonants in an onset cluster are as in (1):

1)

(a)	OS
(b)	OO
(c)	SS
(d)	SO

Logically, a language can have any of the clusters in (1), or any combination of them, or none. A language that has none of the clusters listed in (1) is, of course, a language that does not allow any consonant clusters. We examine only those languages which allow at least one of the clusters listed in (1). Given the cluster combinations in (1), a-priori there are fifteen logical possibilities for combining these clusters into groups of one to four cluster types. Therefore, a-priori there are fifteen logically possible language types, as in (2). If a language L only has one of the onset clusters listed in (1), it can, a-priori, be any one of them, as in (2a). If a language has two of the onset clusters in (1), it can, a-priori be any of the sets listed in (2b). If a language has three of the onset clusters in (1), it can have any of the sets listed in (2c). Finally, it is logically possible for a language to have all four onset clusters in (1), as in (2d). A language that has no onset clusters constitutes an empty group, { }, which is a sixteenth logically possible language type and is excluded from this study:

2)

(a) {OS}
{OO}
{SS}
{SO}
(b) {OS,OO}
{OS,SS}
{OS,SO}
{OO,SS}
{OO,SO}
{SS,SO}
(c) {OS,OO,SS}
{OS,OO,SO}
{OS,SS,SO}
{OO,SS,SO}
(d) {OS,OO,SS,SO}

In sum, in (2) are listed all fifteen logically possible language types (excluding the empty group, which represents a language with no clusters). The question arises, which of the logically possible language types in (2) are occurring language types. To address this, I conducted a cross-linguistic survey of languages that allow word initial onset clusters. The methodology of the survey is outlined in section 2.1.1 and the results of the survey are presented in section 2.1.2.

2.1.1. The survey - methodology

The cross-linguistic typological survey I conducted is based on 62 languages from 22 language families. A complete list of the languages included in the survey is provided in appendix I of this chapter. The survey includes languages which were included in Greenberg (1965), Levin (1985), Morelli (1999) and Steriade (1982) as well as 23 languages that had not been included in any earlier cross-linguistic typological studies. My survey includes only those languages that have onset clusters, which is not the case with Greenberg's survey. This automatically excluded Persian, for example, which is a language with no onset clusters included in Greenberg's survey. Moreover, when I consulted the sources for some of the languages included in Greenberg's, Steriade's, Levin's and Morelli's studies, I decided that some (e.g. those for Eggon and Nisqually) did not contain enough information to be safely included in my survey.

The survey relies on descriptive grammars and grammar books as well as additional research material where available. Multiple sources were used, and data from several sources compared, whenever possible. The basic criterion for including a language is whether it allows consonantal clusters word initially. To be precise, the word initial consonant sequence C_iC_j is taken to be an onset cluster if it does not contain a morpheme boundary or any intervening phonological material as stated in (3):

3) *Onset Cluster* – Let C_i C_j be a word initial sequence of consonants. The sequence C_iC_j is an onset *cluster* iff:
   (i) There is no morpheme boundary between C_i and C_j: (C_i and C_j are tauto-morphemic).
   (ii) There is no segment S_i such that $C_iS_iC_j$ (there is no intervening material between C_iC_j).
   (iii) C_iC_j are linked to the same syllable node.

It should be noted that, all sequences that conform to (3), including sequences of segments which violate the Sonority Sequencing Principle (Selkirk 1984) constitute regular onset clusters for the purposes of this survey. In this, I depart from proposals in the literature (Levin 1985, Steriade 1982 *among others*) which grant a special status to clusters that violate the Sonority Sequencing Principle, for example, by associating the first member of a cluster with declining sonority to prosodic levels higher than the syllable.

The Sonority Sequencing Principle (SSP) states the strong cross-linguistic tendency for syllables to rise in sonority towards the peak and fall in sonority towards the margins. The SSP as formulated in Selkirk (1984) is given in (4). However, while

(4) constitutes a strong tendency, it is not equally obeyed by all types of languages, as will be argued in this chapter:

4) *Sonority Sequencing Principle* (SSP)

In any syllable, there is a segment constituting a sonority peak that is preceded and/or followed by a sequence of segments with progressively decreasing sonority values. (Selkirk 1984:116)

A language was excluded from the survey if its clusters did not conform to one (or more) of the conditions listed in (3). Moreover, any of the circumstances in (5) would exclude a language from the survey:[1]

(i) A language was excluded if it had only obstruent + glide clusters. For example, Korean, which has obstruent + glide clusters such as *py* and *gw*, was not included.[2]

(ii) Also excluded from the survey were languages with only homorganic nasal + obstruent clusters such as *mb* and *nd*. For example Babungo (Schaub 1985) has only simplex onsets and pre-nasalized onsets and no other clusters. The status of pre-nasalized sequences is not transparent. Such

[1] Languages were also excluded for technical reasons, for example, if sources of data were incomplete or inconclusive. Some sources, for example, Matthews (1955) for Dakota, and Hoff (1968) for Carib, do not make a clear distinction between word initial and word medial clusters, which makes it impossible to distinguish them. For Dakota, different grammars listed different possible clusters. Also excluded were languages for which data from different sources were inconsistent. One such example is Chuckchee (Bogoras 1922, Kenstowitz 1981 and Levin 1985 *among others*). Some sources claim that Chukchee contains clusters (Levin 1985 following Bogoras 1922) while others (Kenstowiz 1981) claim that clusters in Chukchee are broken by vowel epenthesis. Skorik (1961) explains that in Chuckchee some words consonantal sequences can appear either with or without a vowel word initially but when the same sequence appears in an onset position word medially, it must appear with the vowel between the two segments or with a preceding vowel, suggesting that consonant sequences are not truly clusters underlyingly. This is confirmed in Asinovskii's (1991) acoustic data. Also excluded, were clusters with a suspected morpheme boundary within the target sequence.

[2] Clusters with glides as the second member are not included in this survey. Surface glides may have a different underlying status. They may be underlying glides that surface as glides or they may be underlyingly vowels that surface as glides (Levi 2004).

sequences can be a cluster or a pre-nasalized segment (Maddieson and Ladefoged 1993, Riehl 2008). Without more information about the phonological status of the sequence, it is impossible to determine whether these sequences are clusters or simply pre-nasalized segments. Languages which have non-homorganic nasal-obstruent sequences in addition to homorganic nasal-obstruent sequences were included in the survey. For example, if a language has *mb* clusters but also *mt* or *mk* clusters (Taba – Bowden 2001), then the language was included in the survey but the homorganic clusters were excluded (i.e. they were not counted as SO clusters since their underlying status is not always transparent, and they may or may not be clusters). The non-homorganic clusters were included in the survey.

(iii) Also excluded were languages that have only *h* + obstruent or *ʔ* + obstruent clusters, or obstruent + *ʔ* and obstruent + *h* clusters such as Comanche (Riggs 1949) since these may function as pre or post-aspiration or glottalization.[3]

(iv) Sequences which contained a morpheme boundary within the cluster were excluded since in (3) clusters were defined as tautosyllabic and tautomorphemic.

In sum, the survey focuses on languages which allow biconosnantal word initial onset clusters. Some of the languages included in the survey, such as Chatino (McKaughan 1954), Georgian (Butskhrikidze 2002), and Polish (Sawicka 1974), to

[3] Mazatec (Steriade 1994) and Temoayan Otomi (Andrews 1949) are examples of languages that have mostly pre and post-aspirated and pre and post-glottalized clusters as well as pre-nasalized clusters, therefore, they were excluded from the survey all together.

name a few, allow clusters longer than two consonants but those clusters were not the focus of this survey.

2.1.2. Results of survey

According to the survey, of the fifteen logically possible language types listed in (2) only four emerge as occurring language types, as in (5):

5)

{OS}	Type 1
{OS, OO}	Type 2
{OS, OO, SS}	Type 3
{OS, OO, SS, SO}	Type 4

Evident from table (2.1) are the implicational relations between the various clusters. If a language allows only one type of cluster it is OS. If a language has OO, it will also allow OS clusters. If a language has an SS cluster, it will also have OO and OS clusters. And lastly, if a language has SO clusters it will allow all other clusters: SS, OO and OS.

Table 2.1: Attested language types: obstruent and sonorant clusters.

Type	OS	OO	SS	SO	Language
Type 1	✓				Basque, Wa
Type 2	✓	✓			Kutenai, Modern Hebrew
Type 3	✓	✓	✓		Greek, Irish
Type 4	✓	✓	✓	✓	Georgian, Russian, Pashto

In sum, evident from table (2.1) are the implicational relations captured in (6):

6)

$$SO \Rightarrow SS \Rightarrow OO \Rightarrow OS$$

The implicational relations in (6) are all unidirectional and without exceptions in the languages of the survey. Next, I single out crucial asymmetries. First, there is an asymmetry between the right and left edge of the implicational relations. SO clusters imply the existence of all other clusters while OS clusters are implied by all other clusters. This asymmetry is expected given that SO is of falling sonority, that is, violates the SSP, while OS has a rise in sonority, i.e. conforms to the SSP. Based on the SSP we expect clusters with rising sonority to occur more frequently than clusters with reversed sonority.

Secondly, and more importantly, we observe an asymmetry between OO and SS clusters. Why is it that SS implies OO but neither OO implies SS nor OO and SS symmetrically imply each other (*OO⇔SS)? This results in {OS, OO} being an occurring language type but *{OS, SS} and *{OS, SS, SO} being a non-occurring language types. Given the SSP which demands a rise in sonority, it is not immediately transparent why {OS, OO} is an occurring language type but *{OS, SS} is not. Both OO and SS are of flat sonority so why is it that there is a language type that includes only OO clusters (type 2), but not a language type that includes only SS clusters?

To better understand the asymmetry between OO and SS clusters, a closer inspection of these cluster types and their respective subtypologies is necessary. In section 2.1.3.1, a subtypology of Obstruent clusters (OO) will be presented followed by a sub-typology of Sonorant clusters (SS) in sections 2.1.3.2 – 2.1.3.4.

2.1.3. Asymmetries between OO and SS clusters

In this section I discuss the asymmetries between OO and SS clusters. I will show that manner of articulation plays a crucial role in the implicational relations of both OO and SS clusters but in SS clusters, in addition to manner of articulation, place of articulation becomes crucial for implicational relations.

2.1.3.1. Obstruent-Obstruent (OO) clusters

Within OO clusters, we consider the four possible combinations involving stops and fricatives as in (7) (affricates are excluded because of their complex structure and their tendency not to combine easily into clusters):

7)

(a) $O_{stop}O_{fric}$

(b) $O_{stop}O_{stop}$

(c) $O_{fric}\ O_{stop}$

(d) $O_{fric}\ O_{fric}$

The clusters listed in (7), again, yield fifteen logically possible language types. However, Morelli (1999) reports that of the fifteen logically possibly language types, only six are attested in the world's languages (these results are similar to Greenberg's (1965) findings). If a language has an $O_{stop}O_{stop}$ cluster, it will also allow an $O_{stop}O_{fric}$ cluster as well as an $O_{fric}O_{stop}$ cluster. Nothing implies $O_{fric}O_{fric}$ but if a language allows fricatives to combine then it will also allow an $O_{stop}O_{fric}$ to combine. Table (2.2) presents Morelli's findings (Morelli 1999:42):

Table 2.2: Attested language types: obstruent clusters.

Type	$O_{fric}O_{stop}$	$O_{stop}O_{fric}$	$O_{stop}O_{stop}$	$O_{fric}O_{fric}$	Language
Type 1	✓				Haida
Type 2	✓			✓	Dutch
Type 3	✓	✓			Mawo
Type 4	✓	✓		✓	Pashto
Type 5	✓	✓	✓		Greek
Type 6	✓	✓	✓	✓	Seri

The implicational relations in (8) arise from table (2.2) (Morelli 1999:44):

8)

$$
\begin{array}{ccc}
 & & O_{stop}O_{stop} \\
 & & \Downarrow \\
 & & O_{stop}O_{fric} \\
 & & \Downarrow \\
O_{fric}O_{fric} & \Rightarrow & O_{fric}O_{stop}
\end{array}
$$

It is important to note that, for obstruent clusters, manner of articulation alone is sufficient to state implicational relations. Clusters with the same manner of articulation imply the presence of at least one cluster with varying manner of articulation. An $O_{stop}O_{stop}$ cluster implies the presence of both $O_{stop}O_{fric}$ and $O_{fric}O_{stop}$ clusters. $O_{fric}O_{fric}$ clusters imply the existence of at least one $O_{fric}O_{stop}$ cluster. However, a cluster with varying manner of articulation, $O_{fric}O_{stop}$ or $O_{stop}O_{fric}$, never implies clusters with the same manner of articulation. It should be noted that clusters beginning with *s* such as *st*, *sp* and *sk* are all considered $O_{fric}O_{stop}$ clusters. Given the special status of *s* and the fact that there are many languages in which the only obstruent cluster is *s*+Obstruent (Swedish, English, Zapotec),[4] $O_{fric}O_{stop}$ clusters are expected to be the least marked clusters and therefore implied by all other clusters.[5]

2.1.3.2. Sonorant-Sonorant (SS) clusters

In this section we look at the sub-typologies of SS clusters. For the purpose of this study only Nasal (N) and Liquid (L) combinations were tested. (Glides were not included in this study due to, their potentially non-transparent underlying status, as noted in section 2.1.1, footnote 2). When considering nasals I focus on *m* and *n*, as

[4] In these languages *š* +obstruent clusters also exist.

[5] For more on the special status of S+stop clusters see Morelli (2003).

they are the most common nasals. For the same reason I consider *r* and *l* as the most common liquids. Geminates are also not discussed in this work as the underlying status of geminates varies cross-linguistically (Chomsky and Halle 1968, Ham 1998 and references therein, Trubetzkoy 1939 *among others*) and they often behave as a single segment rather than a cluster. In this section I begin by looking at clusters whose first member is a nasal, and then turn to clusters whose first member is a liquid. Finally, I will give an overview and a summary of SS clusters in general.[6] Note that the survey of SS clusters is based on a subset of languages in the survey outlined in section 2.1.1. Of the 62 languages in the survey only 28 allow SS clusters and those are listed in appendix II.

I begin by looking at SS clusters whose first member is a nasal, that is, NN and NL clusters, which are presented in table (2.3).

Table 2.3: Sonorant clusters subtypology: first member of cluster: nasal.

Language	NN		NL			
	mn	nm	ml	mr	nr	*nl
Greek	✓					
Tsou	✓					
Ukranian			✓	✓	✓	
Yiddish			✓	✓		
Serbian	✓		✓	✓		
Taba	✓	✓	✓			
Russian	✓		✓	✓	✓	

Note that there are languages with only NN clusters (Greek and Tsou), languages with only NL clusters (Ukrainian and Yiddish) and languages with both NN and NL

[6] It should be noted that for some of the languages there are examples of SS clusters but there is not a complete list of all SS clusters. Such a case is Chatino which according to (McKaughan 1954) has the cluster *rm*. Therefore, although we are able to say that Chatino has an SS cluster, we are not able to give a conclusive list of the SS clusters which exist in that language because the sources provide no detailed lists of the clusters in the language. Languages for which a detailed list is missing are not included in the SS subtypology study.

clusters (Serbian, Taba, Russian). Therefore, we cannot draw any conclusions about the implicational relations between NN and NL cluster types based on manner of articulation alone. In fact we need to take a closer look at the subtypologies of NN and NL clusters separately.

In table (2.4) I provide a table of the subtypology of NN clusters:

Table 2.4: Subtypology of NN clusters.

Language	NN	
	mn	nm
Greek, Tsou	✓	
Taba	✓	✓

It is evident from table (2.4) that if a language has an *nm* cluster, it will have an *mn* cluster, or, as stated in (9):

9)

(i) $nm \Rightarrow mn$

(ii) $N_{[coronal]}N_{[labial]} \Rightarrow N_{[labial]}N_{[coronal]}$

The generalization formalized in (9i) and (9ii) is a weak generalization as it is based on only one language (out of the 29 languages in the survey with SS clusters), which allows the cluster *nm*. However, the generalizations in (9i) and (9ii) seem to reflect a broader generalization regarding preferences in the direction of airflow in clusters. Chitoran *et al*. (2002) have found that in Georgian stop-stop sequences, more gestural overlap was observed in those sequences with front-to-back order of articulation than in those with back-to-front order of articulation. That is, in sequences in which C1 was articulated in a more fronted position in the vocal tract than C2, more gestural overlap was observed. For example, an alveolar-velar sequence has more overlap than a velar-alveolar sequence because alveolars are produced in a more fronted place of articulation than velars, which are further back. This suggests that recoverability of sequences articulated in a front-to-back order of articulation is easier and therefore, it is preferable for sequences of consonants to be articulated in a front-to-back order of articulation than a back-to-front order of articulation. Release may become completely masked and inaudible by a more front closure but is still audible when masked by a more back closure. Therefore, it seems that there is a preference for a front-to-back order of articulation rather than a back-to-front order of articulation (thus a *pt* cluster is preferable to a *tp* cluster because *pt* has front-to-back order of articulation while *tp* has a back-to-front order of articulation). In our case, *mn* clusters are preferred to *nm* clusters simply because *mn* is articulated from a more fronted area of the vocal tract, the lips, to a more back area of the vocal tract, alveolar. Conversely, *nm* is articulated from a more back place of articulation, alveolar, to a more fronted place of articulation, the lips. The former, but not the latter, is more easily recoverable. The preference for the front-to-back order of articulation (the *mn* cluster) over the back-to-front order of articulation (*nm* cluster), is robust in the survey, with 17 out of 29 languages allowing *mn* clusters and only 1 language allowing *nm* clusters. Moreover, 16 out of the 17 languages which allow *mn* have no other NN cluster and only one language allows both *mn* and *nm.*

Next we turn to NL clusters presented in table (2.5):

Table 2.5: Subtypology of NL clusters.

Language	NL			
	ml	mr	nr	*nl
Yiddish, Khasi	✓	✓		
Ukrainian	✓	✓	✓	
Bilaan	✓			

Most obvious in table (2.5) is that no language in the survey has **nl* clusters. In (10) I state the implicational relations that arise from table (2.5). The only implicational relation pertaining to NL cluster type is that if a language has *nr* clusters it will also have *ml* and *mr* clusters as stated in (10). An apparent exception to this is Bilaan, which has *ml* clusters but no *mr* clusters nor a *nr* cluster. However, Bilaan does not have the phoneme *r* in its phonemic inventory and therefore cannot have the clusters *mr*, nor *nr*, hence, the implicational relations stated in (10) hold true for Bilaan as well.

10)

(i) $nr \Rightarrow ml, mr$

(ii) $N_{[coronal]}L_{[coronal]} \Rightarrow N_{[labial]}L_{[coronal]}$

We are now able to draw implicational relations for both NN and NL cluster types. In all nasal initial clusters the implications are based on place of articulation rather than on manner of articulation. From (9) and (10) it follows that if a nasal is the first member of the cluster, then [labial][coronal] clusters are implied by all other cluster types. This is in concurrence with the observation stated earlier that clusters with a front-to-back order of articulation are preferable to all others. A [labial][coronal] cluster is a cluster with a front-to-back order of articulation, and is therefore the preferred cluster implied by all other clusters in both NN and NL cluster types.

I draw the reader's attention to the apparent difference in the typology of SS clusters in comparison to that of OO clusters. If SS clusters behaved similarly to OO clusters, we would expect implicational relations to exist between NN and NL cluster types as NN and NL cluster types differ in manner of articulation. For OO clusters we found that same manner of articulation clusters always imply the existence of varying manner of articulation clusters. Hence, we were able to state that $O_{fric}O_{fric}$ cluster implies the presence of an $O_{fric}O_{stop}$ cluster, regardless of the place of articulation of the segments in the cluster. We could, for example, predict that if a language has the cluster *sf*, it can also have the cluster *sk* (Greek). The same cannot be said of SS clusters. The presence of a cluster with the same manner of articulation, for example, NN cluster, does not imply the presence of NL clusters, nor the presence of any other cluster, as we saw in table (2.3), there are languages which have only NN clusters, languages that have only NL clusters and languages that have both NN and NL clusters. Therefore no predications can be made based on manner of articulation in SS cluster types.

Next, we will explore consonant clusters whose first member is a liquid, LN and LL clusters. Table (2.6) lists possible clusters within the LN cluster type:

Table 2.6: Subtypology of LN clusters.

Language	LN			
	ln	rn	lm	rm
Polish	✓			
Khasi		✓		
Pashto			✓	
Slovak				✓
Bilaan	✓		✓	
Czech	✓			✓

In the LN cluster type each cluster can occur on its own: *ln* can be the only cluster in a language as in Polish, *rn* can be the only cluster as in Khasi, and Pashto and Slovak have *lm* and *rm* clusters respectively, as the only LN clusters in the language. There are also examples of languages that have more than one LN cluster: Bilaan has *lm* and *ln* clusters whereas Czech contains *ln* and *rm* clusters. Therefore, within the LN cluster type, no implicational relations can be drawn between different subtypes, based either on manner or on place of articulation.

A cluster composed of two liquids would be either **rl* or **lr*, however, as apparent in table (2.7) no language has been attested that has either one of these clusters:

Table 2.7: Subtypology: LL clusters.

Lanugage	***LL**	
	*rl	*lr
?		

To summarize, place of articulation plays a crucial role in implicational relations for SS clusters. Moreover, no implicational relations can be drawn based on manner of articulation between NN and NL cluster types. That is, no cluster that belongs to the NN type implies the presence of any cluster in the NL type, and no cluster in the NL type implies the presence of any cluster in the NN type. This is different from the situation in OO clusters where we could draw implicational relations based on manner of articulation alone. Recall that OO clusters with the same manner of articulation imply the presence of clusters with varying manners of articulation.

2.1.3.3. SS in general

Let us now make a comparison across all SS cluster types, those whose first member is a nasal as well as those whose first member is a liquid. We combine the information from three tables, table (2.3), which lists NN and NL clusters, table (2.6) which lists LN clusters as well as table (2.7) which lists LL clusters, into one table (2.8), which lists all SS clusters:

Table 2.8: SS clusters – summary table.

Language	**NN**		**NL**				**LN**				***LL**	
	mn	nm	ml	mr	nr	*nl	ln	rn	lm	rm	*rl	*lr
Greek	✓											
Tsou	✓											
Ukranian			✓	✓	✓							
Yiddish			✓	✓								
Serbian	✓		✓	✓								
Taba	✓	✓	✓									
Russian	✓		✓	✓	✓							
Polish	✓		✓	✓			✓					
Khasi			✓	✓				✓				
Pashto	✓		✓	✓					✓			
Slovak	✓		✓	✓						✓		
Bilaan	✓		✓				✓		✓			
Czech	✓		✓	✓			✓			✓		

In table (2.8) the non-occurring clusters are **nl* and *LL. I draw the reader's attention to the shaded columns which are clearly empty. As no language in the survey has these clusters, this fact strongly suggests cases of systematic gaps.

From table (2.8) we further see that if a language has LN clusters it also has NL clusters. At least one language (Khasi) has LN clusters but has no NN clusters. Therefore, based on manner of articulation we may say that the implicational relation in (11) holds true:

11) LN $\Rightarrow$ NL

2.1.3.4. Gaps in SS clusters

As already noted the following clusters do not surface in any of the languages surveyed:

12)

(i) **nl*

(ii) *LL (**lr*, **rl*)

Let us begin with the **nl* gap. Note that **nl* rises in sonority yet is unattested but *ln* which falls in sonority is an attested cluster.[7] This situation counters what is predicted by the SSP, that is, the requirement that onset clusters rise in sonority. Normally we would expect that if a cluster of falling sonority is present in a language then its counterpart which rises in sonority should also be present. Yet, the exact opposite is attested in this case.

One possible reason for the non-occurrence of **nl* clusters might be that they are acoustically disadvantageous. Acoustically *n* and *l* are similar and therefore difficult to distinguish. They both have: (1) low first formant (2) antiformants and (3) similar formant transitions, since both are alveolars. However, if this was the reason for the lack of **nl* clusters then we should not have *ln* clusters either. Yet *ln* clusters do occur while **nl* clusters do not.

The reason for this asymmetry may lie in the transitions between nasals and liquids. When a nasal precedes a liquid, there is a risk of strong nasal co-articulation. There is a great risk that the velum may not be raised in time to seal the nasal cavity

[7] Here we must refer to a more elaborate sonority scale which assumes: L < N < O.

and cease airflow through the nasal cavity. Thus, air may continue to flow through the nose even after the onset of the second segment *l*. This nasal co-articulation into the following segment, also known as a "nasal leak", makes **nl* indistinguishable from *nn* resulting in **nl* sounding like a geminate. This is not true of *ln*, since the liquid does not co-articulate with the nasal in the same manner. In the case of *ln* the velum must be lowered after *l*. Therefore, *l* is not likely to "leak" into the following nasal segment, making the cluster *ln* acoustically more advantageous than the cluster **nl*.

Moving to the *LL gap, there are two possible reasons why *LL clusters do not occur in natural language.[8] First, *r* and *l* are acoustically similar and are difficult to distinguish. More than just being acoustically similar to each other, they are very similar to vowels. Thus, having a sequence of *r* and *l* pre-vocalically is acoustically disadvantageous. Secondly, *r* and *l* do not combine into a cluster because they do not differ in either place or manner of articulation. I argued above that place of articulation plays a crucial role in the distinction of SS cluster types. Manner of articulation alone is not sufficient to predict the implicational relations between SS clusters. Manner and place of articulation must combine together to yield the types of SS clusters that occur in languages.

To summarize thus far, we have seen that place of articulation is crucial for stating implicational relations within SS clusters. There is only one set of manner based implicational relations and two sets of place based implicational relations, as in (13):

[8] It should be noted that in Taba there are surface *rl* sequences, but closer inspection of these sequences reveals that they contain a morpheme boundary between *r* and *l* and therefore do not count as clusters as defined in (3). Other than the one case of Taba, I am not aware of any cases of *LL clusters.

13) Manner based implicational relations:

LN ⇒ NL[9]

Place based implicational relations:

nm ⇒ *mn*

nr ⇒ *ml*, *mr*

2.1.3.5. Concluding remarks

As pointed out in section 2.1.3, there are several asymmetries between OO clusters and SS clusters. First, same manner OO clusters imply different manner OO clusters. Thus, $O_{fric}O_{fric}$ or $O_{stop}O_{stop}$ cannot be the only OO cluster in a language. In contrast, same manner SS clusters do not imply different manner SS clusters. NN can be the only SS cluster in a language (e.g. Greek, Tsou). On the other hand *LL clusters are not occurring clusters in any language. With the background provided in sections 2.1, we are now in a position to address the question raised in section 2.1.2: why do OO clusters imply SS clusters but SS clusters do not imply OO clusters? There are several reasons that make OO clusters less marked than SS clusters. One possible reason is the acoustic salience of obstruents as opposed to sonorants. Obstruents are perceptually more salient than sonorants, therefore their combinations are also more salient. Ohala (1983:193) notes:

> Obstruents, especially those that involve a transient burst due to the rapid equalization of an appreciable difference in air pressure, create more rapid spectral changes and thus are able to carry more information and make more distinctive sounds than nonobstruents.

[9] These findings are similar to Greenberg's findings (1965).

Obstruents, due to their acoustic attributes, when released, carry more information, especially in onset position (or word initial position) and are therefore easier to distinguish from non-obstruents. We may deduce that their combinations are also more acoustically salient than combinations of sonorants and are therefore perceptually more advantageous. This might also explain another cross-linguistic observation made by Lindblöm and Maddieson (1988): that phonemic inventories of languages tend to have a distribution of roughly 70% obstruents and 30% sonorants. This results in greater clustering possibilities for obstruents than for sonorants simply because there are more obstruents than sonorants. Finally, in SS clusters difference in place becomes crucial to reinforce their perceptual salience. For obstruents, however, manner of articulation alone is sufficient to distinguish between possible members of a cluster.

2.1.4. Further implications

We already noted in section 2.1.3.3, that if a language has LN clusters it also has NL clusters (repeated in (14a)). It is of interest that the presence for LN, which is the only sonority reversed SS cluster, also implies the presence of other sonority reversed clusters, in particular, the presence of SO clusters. Polish has *ln* as well as *lž* and *mž* clusters and Georgian has *lm* as well as *lb* and *rb*. In other words, if a language has LN, it will also have SO clusters, as stated in (14b):

14)

(a) LN $\Rightarrow$ NL

(b) LN $\Rightarrow$ SO

2.1.5. Predictions about language type shifts

Historical changes in a language's cluster inventory can cause a language to shift types. For example, a language that does not allow clusters at one stage but allows them at another stage, is said to shift types. Clusters may become part of the grammar in several ways: borrowings, morphological processes, or phonological processes such as syncope. Predictions regarding language type shifts follow from the implicational relations stated in (6). A language L1 of type T1, can change membership and become a member of another type, T2, by changing the inventory of clusters allowed by the language's grammar. It follows from (6) that if a language has no clusters then the first cluster type it will achieve is OS. Thus, a language with no clusters can shift to become a type 1 language, i.e. a language with OS clusters. Examples of languages that shifted types are West Greenlandic (Fortescue 1984) and Popoluca (Elson 1947). Both languages disallowed consonantal clusters word initially at an earlier point in history, and due to borrowing (from Danish and Spanish respectively), have shifted to become type 1 languages; both now allow OS clusters.

A language may also gain clusters through a process of vowel syncope. For example, a vowel may be consistently deleted in the first syllable of every word. That could result in a language gaining all types of clusters at once and becoming a type 4 language. However, a language cannot gain only {OO} or only {SS} clusters as languages with only {OO} or {SS} clusters are not empirically attested and are therefore not part of the typology.

It is also possible for a language to lose clusters. Once again it is predicted that if a language loses one cluster type, it will lose the cluster type which implies all other clusters. Thus, a language of type 4, which allows reversed sonority clusters, those that imply all other clusters, may disallow such clusters and shift to become a type 3 language.

The prediction is that no matter what stage the language is in, if it gains or loses clusters, it must become a language type that is predicted by the typology. A language will never gain only OO and SS clusters without having OS clusters as well, because the set *{OO, SS} cannot belong to an occurring language type.

2.2. Voicing typology

In section 2.1 I presented a typology of word initial onset clusters based on the feature [sonorant]. We now turn to the typology of word initial biconsonantal onset clusters based on the feature [voice], in which I will be focusing only on obstruent clusters.

2.2.1. Voicing combinations

In word initial, biconsonantal onset clusters there are four logical combinations of voiced ([+v]) and voiceless ([−v]) obstruents as in (15):

15)

(a) [−v][−v]
(b) [+v][+v]
(c) [−v][+v]
(d) [+v][−v]

Previous studies of voicing in clusters (Lindblöm 1983, Lombardi 1991, 1995, 1999, Wetzel and Mascaró 2001, Wheeler 2005, *among others*) accept (15a-c) as possible clusters, but there are conflicting reports in the literature regarding cluster (15d). While Blevins (2003), Greenberg (1965), and Steriade (1997) all report that (15d) is an attested cluster, Lombardi (1991, 1999) and Lindblöm (1983) categorically reject it from the set of occurring clusters. Let us review these conflicting claims in depth.

In phonological studies such as Lombardi (1991, 1999), clusters of type (15a-c) are assumed to be possible clusters but clusters of type (15d) are excluded from the set of occurring clusters. In other words, the configuration in figure (2.1), in which a voiced obstruent precedes a voiceless obstruent in pre-nuclear position, is taken to be ill-formed:

*Voiced Obstruent – voiceless obstruent – syllable nucleus

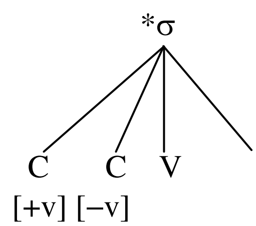


Figure 2.1: Possible onsets (Lombardi 1991).

Lombardi claims that the prohibition against [+v][−v] is universal. According to her, voiced segments may occur only before a sonorant segment, either a vowel or a sonorant consonant. Her argument continues that a [+v][−v] obstruent cluster cannot be an occurring cluster type because voiceless segments cannot intervene between a voiced obstruent and a vowel. She refers to figure (2.1) as a "Universal Sonority Constraint", "…an absolute universal which no language can violate." (1991:59). Moreover, Lombardi correlates the prohibition in figure (2.1) with the prohibition on sonority reversed clusters. For her "voicing reversals" are comparable to SO clusters. As we will see later in this chapter, this parallel is untenable.

Likewise, Lindblöm (1983) claims, based on the principle of gestural economy, that [+v][−v] clusters should be excluded on phonetic grounds. According to Lindblöm, [+v][−v] is an illicit structure for the following reasons:

> We should also mention here the absence of rapid intrasyllabic alternations of inspiration and expiration. The only universally preferred airstream mechanism appears to be expiratory. A similar no doubt energy-saving arrangement is observed in the distribution of phonation types… **Clusters do not allow *[+voiced C][–voiced C]V initially**, nor its mirror image finally. (Lindblöm 1983:240) (my emphasis).

However, Greenberg's (1965) survey based on 104 languages found the following statistics on voicing in initial clusters:

16) Greengberg's 1965 survey:

(a)	Voiceless + Voiceless	[–v][–v]	66.7%
(b)	Voiced + Voiced	[+v][+v]	21.65%
(c)	Voiceless + Voiced	[–v][+v]	10.68%
(d)	Voiced + Voiceless	[+v][–v]	0.97%

From (16) it is clear that the majority, or two thirds (66.7%) of the clusters in Greenberg's survey are sequences of voiceless obstruents [–v][–v]. All other cluster types constitute the remaining third. Of these, almost 22% are [+v][+v] and just under 11% are [–v][+v] clusters. Under 1% of all clusters are [+v][–v] clusters. However, the numbers are somewhat misleading since Greenberg does not separate obstruent clusters from sonorant clusters. A *tn* cluster, for example, is considered a [–v][+v] cluster. This skews the numbers of [–v][+v] clusters and [+v][+v] clusters, making it difficult to correctly decipher the statistical data.

Evident from this survey is that clusters in which both members are voiceless are preferable to clusters with any other voicing combination. Mixed voicing clusters are a great minority at just a little over 10% of all clusters but both [–v][+v] and [+v][–v] clusters exist. However, while Greenberg accepts the existence of [–v][+v] clusters, he doubts the existence of [+v][–v] clusters, although his survey lists two languages, Bilaan and Khasi, for which obstruent [+v][–v] clusters have been reported. Since Greenberg's sources for Khasi and Bilaan (Rabel 1961 and Dean 1955 respectively), presented no phonetic evidence for [+v][–v] obstruent clusters, Greenberg allows for the possibility that the reported [+v][–v] clusters are phonetically realized as [–v][–v]; clusters like *bt* reported for Khasi and *bs* reported for Bilaan, might actually be phonetically realized as *pt* and *ps* respectively.[10] Since Bilaan does not distinguish between *b* and *p*, and contains only *b* in its phonemic inventory, it is possible that the cluster *bs* listed in the grammar is phonetically realized as *ps*. With no phonetic evidence for Bilaan, it is impossible to determine how the cluster *bs* is realized.

Blevins (2003) and Steriade (1997) also mention the existence of [+v][–v] clusters. Both recognize the existence of these clusters in Khasi based on descriptive grammars but do not pursue their phonetics. They use [+v][–v] clusters to argue for licensing by cue. That is, voicing distinctions are more likely to occur in perceptually advantageous environments in which the perception of voicing is enhanced. In an OO environment, cues for voicing are rather poor and therefore voicing is less likely to occur in this environment (particularly if voicing is contrastive within a cluster as in the case of [–v][+v] or [+v][–v]). Steriade (1997:7) argues that lack of perceptual cues in an OO environment accounts for the rarity of languages with [+v][–v] clusters word initially. Although, both Blevins and Steriade examine the influence of perceptual cues on the distribution of voicing in obstruents they do not pursue the phonetic implementation of such clusters in languages that do have [+v][–v] clusters.

To summarize, we encounter conflicting claims in the literature concerning [+v][–v] sequencing. On the one hand, based on grammatical descriptions, Blevins (2003), Greenberg (1965, 1978) and Steriade (1997) document the existence of

---

[10] In fact, recent phonetic findings show that Khasi does indeed have [+v][–v] obstruent clusters (Henderson 1991).

[+v][–v] sequences word initially. However, none pursue the phonetic realization of these sequences. On the other hand, Lindblöm (1983) and Lombardi (1991, 1995) exclude [+v][–v] word initial sequences as clusters on theoretical grounds, both in the phonology and in the phonetics.

To address the conflict between phonetic and phonological theories on the one hand and scarce empirical evidence on the other, a cross linguistic typological study was conducted to establish the distributional facts of voicing in word initial onset obstruent clusters. The cross linguistic typological study is supported by an acoustic phonetic study, which will be presented in chapter 5, where I test the phonetic implementation of voicing in word initial obstruent clusters.

2.2.2. The survey

The languages included in the survey are the same languages used for the survey outlined in section 2.1. However, while the earlier survey included languages with clusters containing both obstruents and sonorants, the present survey includes only languages with obstruent clusters. That is, only 46 languages of the 62 surveyed for the feature [sonorant] were surveyed for the feature [voice]. Some languages such as German, Klamath and Welsh were also not included in this survey, although they do have OO clusters. However, in these languages the distinction between orthographic *p*, *t*, *k* and *b*, *d*, *g* is claimed to be based on the feature [spread glottis]. That is, they are claimed to have a distinction between unaspirated and aspirated stops rather than voiced/voiceless. For some of these languages (German), there are conflicting claims regarding the proper distinctive laryngeal feature. Since the nature of the distinctive feature in these languages is outside the scope of this work, these languages were excluded from the survey. The methodology I employed is the same as described in 2.1.1. The languages included in this section are listed in appendix III.

Results of the survey indicate that in reality clusters of the [+v][−v] type do occur albeit they are rare. Three cases have been documented with supporting phonetic evidence: (i) Khasi in which, *dk* in *dkar* 'tortoise' is distinct from *tk* in *tkor-tkor* 'plump and tender' (Henderson 1991); (ii) Tsou in which *ɓs* is distinct from *ps* (Wright 1996); and (iii) Modern Hebrew in which *dk* in *dkalim* 'palms' is distinct from *tk* in *tkarim* 'flat tires' and *dg* in *dgalim* 'flags.' Given these facts, Lombardi's cross-linguistic prohibition against [+v][−v] clusters and Linblöm's prediction that [+v][−v] clusters cannot be produced, have no empirical basis.

For Khasi and Tsou the following phonetic evidence is available supporting the existence of [+v][−v] clusters. In a phonetic study of Khasi, Henderson (1991) showed, using an oscilloscope, differences in the realization of Khasi initial clusters, presented in figure (2.2):

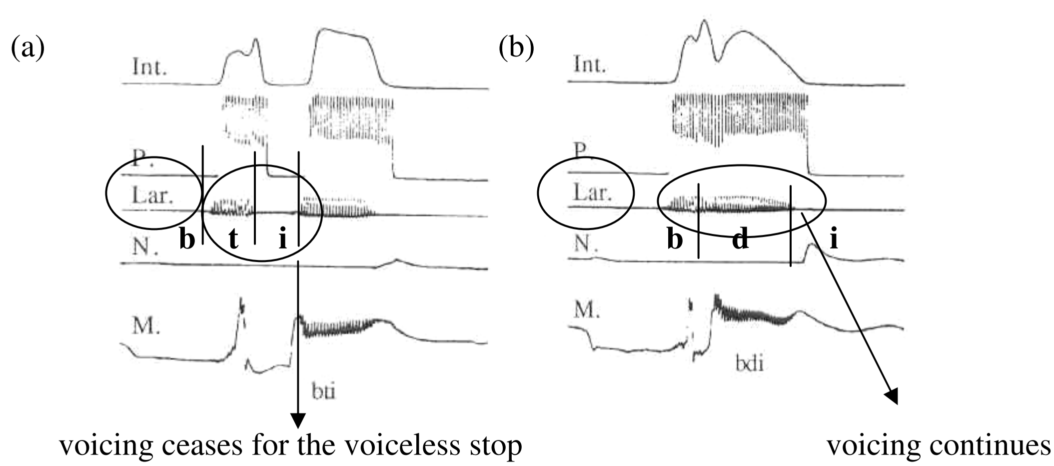


Figure 2.2: Khasi *bti* and *bdi* initial clusters (Henderson 1991).

In figure (2.2) the laryngeal trace in each picture is circled and demonstrates differences in the realization of the minimal pair: *bti* ~ *bdi*. Each laryngeal trace marks the laryngeal activity during the production of the cluster. In figure (2.2b) there is laryngeal activity through the entire cluster *bd*, which is not present in the laryngeal trace in figure (2.2a) for *bt*, where laryngeal activity ceases after *b* and does not resume until the vowel; there is no laryngeal activity during the production of *t* as expected for voiceless segments. The laryngeal traces in figure (2.2) provide phonetic evidence for the existence of [+v][−v] clusters in Khasi.

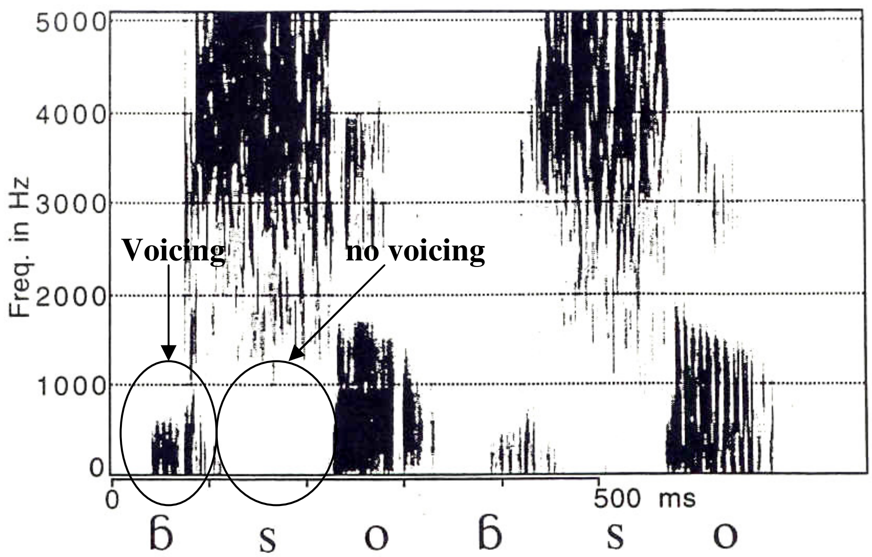


Figure 2.3: A spectrogram of the Tsou word *ɓsoɓso* ‘potato tuber.’ (Wright 1996)

In figure (2.3) a spectrogram of the word *ɓsoɓso* ‘potato tuber’ from Tsou is provided (Wright 1996). In the spectrogram closure voicing is evident prior to the burst and the frication noise for the following fricative, is adjacent to the burst. As can be seen in the spectrogram, the voiced stop is followed by a voiceless fricative with no intervening vowel between them.

My own findings for Modern Hebrew, which provide further acoustic phonetic evidence for the existence of [+v][−v], will be presented in chapter 5.

Table (2.9) summarizes the results of the survey for all cluster types listed in (17). Of all the fifteen logically possible language types that result from all possible groupings of the clusters in (17), only 6 occurring language types emerge, as in table (2.9):

17)

(a) {[−v][−v]}
{[+v][+v]}
{[−v][+v]}
{[+v][−v]}
(b) {[−v][−v], [+v][+v]}
{[−v][−v], [−v][+v]}
{[−v][−v], [+v][−v]}
{[+v][+v], [−v][+v]}
{[+v][+v], [+v][−v]}
{[−v][+v], [+v][−v]}
(c) {[−v][−v], [+v][+v], [−v][+v]}
{[−v][−v], [+v][+v], [+v][−v]}
{[−v][−v], [−v][+v], [+v][−v]}
{[+v][+v], [−v][+v], [+v][−v]}
(d) {[−v][−v], [+v][+v], [−v][+v], [+v][−v]}

Table 2.9: Attested cluster types according to possible voicing combinations.

	[−v][−v]	**[+v][+v]**	**[−v][+v]**	**[+v][−v]**	**sample language**
Type 1	✓				Dutch, Kutenai
Type 2	✓	✓			Greek, Rumanian,
Type 3	✓	✓	✓		Georgian
Type 4	✓	✓	✓	✓	M.H, Tsou, Khasi
Type 5	✓		✓		Biloxi, Camsa, Klamath
Type 6	✓		✓	✓	Bilaan, Amuesha

Clear implicational relations stated in (18) arise from table (2.9):

18)

$$\begin{array}{ccc} & & [+v][-v] \\ & & \Downarrow \\ & & [-v][+v] \\ & & \Downarrow \\ [+v][+v] & \Rightarrow & [-v][-v] \end{array}$$

A language type 1 has only [−v][−v] clusters. A language type 2 has [+v][+v] clusters and thus, by implication, also [−v][−v] clusters as in (19):

19)

$$[+v][+v] \Rightarrow [-v][-v]$$

A language type 3 has [+v][+v] clusters and also [–v][+v] clusters. Thus, by implication, it also has [–v][–v] clusters, as in (20):

20)

$$\begin{array}{ccc} & & [-v][+v] \\ & & \Downarrow \\ [+v][+v] & \Rightarrow & [-v][-v] \end{array}$$

A type 4 language has all possible voicing combinations as in (17) and a type 5 language has only [−v][+v] and [−v][−v] clusters as in (21):

21)

$$\begin{array}{c} [-v][+v] \\ \Downarrow \\ [-v][-v] \end{array}$$

A type 6 language has both possible varying voicing clusters [−v][+v] and [+v][−v] and therefore by implication also [−v][−v] clusters as in (22):

22)

[+v][−v]
⇓
[−v][+v]
⇓
[−v][−v]

A type 6 language is typologically predicted on the basis of the implicational relations in (22); in table (2.9) this is exemplified by Bilaan and Amuesha. The only grammatical description of Bilaan that we have (Dean 1955) lists [−v][−v], [−v][+v] and [+v][−v] as occurring clusters, making Bilaan a type 6 language. However, as mentioned previously, with lack of phonetic evidence, the cases of [+v][−v] in Bilaan are suspect. Without further phonetic investigation it is impossible to determine whether the [+v][−v] clusters are realized as such in Bilaan or whether some other phonetic properties are used to distinguish these clusters.

From these implications we may draw the following conclusions: (i) if a language has only one cluster with the same voicing it will be [−v][−v]; (ii) if a language has a mixed voicing cluster it must have at least one cluster with the same voicing; (iii) nothing implies the presence of [+v][+v] clusters; but (iv) the presence of a [+v][+v] cluster implies the existence of a [−v][−v] cluster; and does not imply the existence of any other cluster.

2.3. Comparing the two typologies

We now compare the implicational relations for the two typologies presented in 2.1 and 2.2. For [voice] in obstruents, the preferred cluster type, the cluster that is implied by all other clusters, is [−v][−v], the one in which both segments are voiceless as in (19), repeated in (24b). Thus, in the least marked cluster its members have the same voicing specification. For [sonorant] the least marked cluster, the one implied by all other clusters, is OS as in (6) repeated in (23a). Thus in the least marked cluster, its members have different specifications for the feature [sonorant].

23) (a) Sonority implicational relations:

$$\text{SO} \Rightarrow \text{SS} \Rightarrow \text{OO} \Rightarrow \text{OS}$$

(b)

$$\begin{array}{ccc} & & [+\text{v}][-\text{v}] \\ & & \Downarrow \\ & & [-\text{v}][+\text{v}] \\ & & \Downarrow \\ [+\text{v}][+\text{v}] & \Rightarrow & [-\text{v}][-\text{v}] \end{array}$$

It is clear from (23) that there is no basis for correlating the role of [voice] and [sonorant] in onset clusters. In other words, Lombardi's proposal to treat [+v][−v] as a sonority reversed cluster, comparable to SO, is not supported by the place of these clusters in their respective typologies. Even if we adopt Lombardi's assumption that difference in voicing is comparable to difference in sonority, the least marked cluster in terms of [voice] has no "rise" in voicing, while the least marked cluster in terms of [sonorant] does. That is, the least marked cluster in terms of voicing ([voice]) is that in which both members of the cluster have the same voicing specifications; the least marked cluster in terms of sonority ([sonorant]) is that in which the sonority values of each members of the clusters are different and the second member is more sonorous than the first member. Moreover, while the typology based on the feature [sonorant] in (23a) yields SO as the most marked cluster, the typology based on the feature [voice] does not yield a single most marked cluster. Both [+v][+v] and [+v][−v] are candidates for this status.

APPENDIX I: Table of clusters by the features [sonorant].

Language	OS	OO	SS	SO	Language family
1. Aguacatec	✓	✓	✓	✓	Mayan
2. Aleut	✓	✓			Eskimo-Aleut
3. Amuesha	✓	✓			Arawakan
4. Basque	✓				Basque
5. Belarusian	✓	✓	✓		Slavic (IE)
6. Bilaan	✓	✓	✓	✓	Austronesian
7. Biloxi	✓	✓			Siouan
8. Breton	✓	✓			Celtic (IE)
9. Bulgarian	✓	✓	✓		Slavic (IE)
10. Cambodian	✓	✓	✓	✓	Mon-Khmer
11. Camsa	✓	✓			lang. isolate.
12. Chami	✓				Choco
13. Chatino	✓	✓	✓	✓	Oto- Manguean
14. Cornish	✓	✓			Celtic (IE)
15. Czech	✓	✓	✓	✓	Slavic (IE)
16. Danish	✓	✓			Germanic (IE)
17. Dutch	✓	✓			Germanic (IE)
18. Embara – Catio	✓				Choco
19. Frisian	✓	✓			Germanic (IE)
20. Gaelic (Scots)	✓	✓	✓		Celtic (IE)
21. Georgian	✓	✓	✓	✓	Kartvelian
22. German	✓	✓			Germanic (IE)
23. Greek	✓	✓	✓		Greek (IE)
24. Hebrew (Modern)	✓	✓			Semitic
25. Hindi	✓	✓	✓		Indo-Iranian (IE)
26. Hixkaryana	✓	✓			Carib
27. Hungarian	✓	✓			Uralic(Fino-Ugric)
28. Icelandic	✓	✓			Germanic (IE)
29. Inga	✓	✓	✓	✓	Quechua (II)
30. Irish	✓	✓	✓		Celtic (IE)
31. Khasi	✓	✓	✓	✓	Mon-Khmer
32. Klamath	✓	✓	✓	✓	Penutian
33. Kobon	✓				Trans New Guinea
34. Kutenai	✓	✓			language isolate
35. Lithuanian	✓	✓			Baltic (IE)
36. Macedonian	✓	✓	✓		Slavic (IE)
37. Manx	✓	✓	✓		Celtic (IE)
38. Mon (Burmese)	✓				Mon-Khmer
39. Norwegian	✓	✓			Germanic (IE)
40. Pashto	✓	✓	✓	✓	Indo-Iranian (IE)
41. Polish	✓	✓	✓	✓	Slavic (IE)

APPENDIX I (Continued)

Language	OS	OO	SS	SO	Language family
42. Popoluca	✓				Oto- Manguean
43. Romani	✓	✓			Indo-Iranian (IE)
44. Romanian	✓	✓	✓		Romance (IE)
45. Russian	✓	✓	✓	✓	Slavic (IE)
46. Serbian	✓	✓	✓		Slavic (IE)
47. Seri	✓	✓			Hokan
48. Slovak	✓	✓	✓	✓	Slavic (IE)
49. Slovenian	✓	✓	✓		Slavic (IE)
50. Sorbian (lower)	✓	✓	✓	✓	Slavic (IE)
51. Sorbian (upper)	✓	✓	✓	✓	Slavic (IE)
52. Spanish	✓				Celtic (IE)
53. Swedish	✓	✓			Germanic (IE)
54. Taba	✓	✓	✓	✓	Austronesian
55. Totonac	✓	✓			Totonacan
56. Tsou	✓	✓	✓	✓	Austronesian
57. Ukrainian	✓	✓	✓	✓	Slavic (IE)
58. Wa	✓				Mon-Khmer
59. Welsh	✓	✓			Celtic (IE)
60. Yiddish	✓	✓	✓		Germanic (IE)
61. Zapotec(Isthmus)[11]	✓	✓			Oto- Manguean
62. Zoque	✓				Mixe-Zoque

[11] Mitla Zapotec shows the same patterns.

APPENDIX II: SS clusters.

Language	NL				LN				NN		*LL
	ml	mr	nr	*nl	lm	rm	ln	rn	mn	nm	*LL
1. Belarusian	✓	✓							✓		
2. Bilaan	✓				✓		✓		✓		
3. Bulgarian	✓	✓	✓						✓		
4. Cambodian	✓	✓							✓		
5. Czech	✓	✓				✓	✓		✓		
6. Gaelic(Scot)		mr[12]									
7. Georgian	✓	✓			✓						
8. Greek									✓		
9. Hindi	✓	✓	✓								
10. Inga											
11. Irish	✓	✓									
12. Khasi	✓	✓			✓			✓			
13. Klamath					✓						
14. Macedonian	✓	✓							✓		
15. Manx		mr									
16. Pashto	✓	✓			✓				✓		
17. Polish	✓	✓					✓		✓		
18. Romanian	✓	✓									
19. Russian	✓	✓	✓						✓		
20. Serbian	✓	✓							✓		
21. Slovak	✓	✓				✓			✓		
22. Slovenian	✓	✓	✓						✓		
23. Sorbian (l)	✓	✓					✓				
24. Sorbian (u)	✓	✓							✓		
25. Taba	✓								✓	✓	
26. Tsou									✓		
27. Ukranian	✓	✓	✓								
28. Yiddish	✓	✓									

[12] In both Scots Gaelic and Manx, present day *mr* originates in *mn* (Harbert p.c).

APPENDIX III: Table of clusters by the features [voice].[13,14]

Language	[–v][–v]	[+v][+v]	[–v][+v]	[+v][–v]	Language family
1. Aguatatec	✓		✓		Mayan
2. Aleut	✓				Eskimo-Aleut
3. Amuesha	✓		✓	✓	Arawakan
4. Belarusian	✓	✓			IE
5. Bilaan	✓		✓	✓	Austronesian
6. Biloxi	✓		✓		Siouan
7. Breton	✓				IE
8. Bulgarian	✓	✓			IE
9. Cambodian	✓		✓		Mon-Khmer
10. Camsa	✓		✓		Language Isolate
11. Chatino	✓		✓[15]		Oto-Manguean
12. Cornish	✓				IE
13. Czech	✓	✓			IE
14. Dutch	✓				IE
15. Frisian	✓				IE
16. Georgian	✓	✓	✓		Kartvelian
17. Greek	✓	✓			IE
18. Hebrew (Modern)	✓	✓	✓	✓	Semitic
19. Hindi	✓				IE
20. Hixkaryana	✓				Carib
21. Hungarian	✓				Uralic(Fino-Ugric)
22. Inga	✓				Quechua
23. Khasi	✓	✓	✓	✓	Austronesian
24. Kutenai	✓				Language Isolate
25. Lithuanian	✓	✓			IE
26. Macedonian	✓	✓			IE
27. Manx	✓				IE
28. Norwegian	✓				IE
29. Pashto	✓	✓			IE
30. Polish	✓				IE
31. Romani	✓				IE
32. Romanian	✓	✓			IE
33. Russian	✓	✓			IE
34. Serbian	✓	✓			IE

[13] In Danish (Hansen 1967), Gaelic (both Scots and Irish) (Green 1997, Harbert p.c), German (Iverson and Salmons 1995, Jessen 2001 among others, see chapter 1 for more sources), Icelandic (Rögnvaldsson 1993) and Klamath (Blevins p.c) stops are claimed to be distinguished by aspiration and not by voicing. Therefore, none of these languages were included in the chart in appendix III.
[14] In Hungarian *v* behaves as a sonorant (Barkai and Horvath 1978) and therefore clusters with *v* as second member were not included.
[15] Chatino may or may not have [–v][+v], not clear from data in source.

APPENDIX III (Continued).

Language	[−v][−v]	[+v][+v]	[−v][+v]	[+v][−v]	Language family
35. Seri	✓				Hokan
36. Slovak	✓	✓			IE
37. Slovenian	✓	✓			IE
38. Sorbian (low)	✓	✓			IE
39. Sorbian (upp)	✓	✓			IE
40. Swedish	✓				IE
41. Taba	✓				Austronesian
42. Totonac	✓				Totonacan
43. Tsou	✓	✓[16]	✓	✓	Austronesian
44. Ukrainian	✓	✓			IE
45. Yiddish	✓	✓	✓		IE
46. Zapotec(Isthmus)	✓				Oto-Manguean

[16] One [+v][+v] cluster *zv*, which may or may not be an obstruent [+v][+v] cluster, depending on the status of *v* in Tsou and whether it behaves as a sonorant. If *v* is a sonorant, then Tsou would be classified as a type 5 language.

APPENDIX IV: Cluster type distribution

OS	OO	SS	SO
62/62	53/62	31/62	18/62
100%	85%	50%	29%

[−v][−v]	[+v][+v]	[−v][+v]	[+v][−v]
46/46	20/46	12/46	5/46
100%	43%	26%	11%

CHAPTER 3

VOICING IN MODERN HEBREW OBSTRUENT CLUSTERS

3.0. Introduction

In chapter 2 Modern Hebrew was classified as a type 2 language in terms of the feature [sonorant], and as a type 4 language in terms of the feature [voice]. That is to say, Modern Hebrew word initial onsets admit only OS and OO clusters but allow all voicing combinations in obstruent clusters, [−v][−v], [+v][+v], [−v][+v] and [+v][−v]. Thus, onset clusters are more restricted in terms of sonority than voicing. Sonorant clusters SS and sonority reversals SO are strictly banned and when such clusters arise as a result of morphological processes, they are broken by vowel epenthesis, as will be shown in sections 3.3 and 3.4. However, when clusters of varying voicing arise due to morphological processes such as plural suffixation or diminutive reduplication (Kreitman 2003), they are not broken up by vowel epenthesis. This clearly shows that [+v][−v] clusters are not treated in the same way as SO clusters, contrary to claims by Lombardi (1991) and Morelli (1999).

I begin this chapter with an overview of possible obstruent onset clusters in Modern Hebrew provided in section 3.2. In section 3.3 I discuss morphological processes whereby clusters arise systematically, in particular, plural affixation in the Segolate noun class and diminutive reduplication. In section 3.4 I address the phonological representation of clusters and discuss the claim that the phonological representation of [+v][−v] clusters is the same as that of clusters with falling sonority. I discuss and reject proposals to treat the voiced C1 as an appendix and show that such a solution is not suitable for [+v][−v] clusters.

3.1. Obstruent onset clusters

Of particular interest in this work is the distribution of voicing in obstruent clusters. But in order to fully address the issue of voicing in obstruent clusters, we also need to address how manner and place of articulation affect the range of possible clusters.

Table 3.1: Obstruent clusters in Modern Hebrew

Second member →														
				Voiceless						**Voiced**				
				Stops			**Fricatives**			**Stops**			**Fricatives**	
				p	**t**	**k**	**f**	**s**	**x**	**b**	**d**	**g**	**v**	**z**
First member ↑	**Voiceless**	**Stops**	**p**		pt	pk		ps	px		pd	pg		pz
			t			tk	tf		tx			tg	tv	tz
			k		kt		kf	ks	kx		kd		kv	kz
		Fric	**f**		*ft*									
			s	sp	st	sk	sf		sx	sb	sd	sg	sv	
			x		*xt*									
	Voiced	**Stops**	**b**		bt	bk		bs	bx		bd	bg		bz
			d			dk	df		dx			dg	dv	
			g				gf	gs	gx		gd		gv	gz
		Fri	**v**											
			z			zk			zx	zb	zd	zg	zv	

Table (3.1) lists obstruent onset clusters classified by voicing and manner of articulation. For a full range of biconsonental onset clusters in Modern Hebrew, including both obstruent and sonorant consonants, see table (1.3) in chapter 1. While there are gaps in table (3.1), we find evidence for all combinations of manner and voicing in obstruent clusters. Starting with manner of articulation, all four logically possible manner combinations occur: stop-stop, fricative-fricative, stop-fricative and fricative-stop. Moreover, all logically possible voicing combinations also occur: [–v][–v], [+v][+v], [–v][+v] and [+v][–v]. In other words, voicing combinations are in no way restricted. There are as many [+v][–v] clusters (12) as there are [+v][+v] clusters (12), and almost as many [+v][–v] (12) as there are [–v][+v] clusters (13).

Table (3.2) lists monomorphemic forms which exemplify all possible voicing combinations, cross-referenced with the four possible combinations of manner of articulation.

Table 3.2: Examples of obstruent clusters in Modern Hebrew

		Voiceless		**Voiced**	
		stop	**fricative**	**stop**	**fricative**
Voiceless	**stop**	***pk***_uda_ ‘command’	***ps***_ila_ ‘strike’	***bg***_ida_ ‘treason’	***kv***_ar_ ‘already’
	fric	***sk***_ira_ ‘review’	***sf***_og_ ‘sponge’	***zg***_ugit_ ‘glass’	***zv***_axim_ ‘offering’
Voiced	**stop**	***dk***_ira_ ‘stab’	***dx***_isa_ ‘compression’	***dg***_ima_ ‘sample’	***bz***_iza_ ‘looting’
	fric	***zk***_ena_ ‘old lady’	***zx***_ut_ ‘right’	***zd***_oni_ ‘evil’	***zv***_uv_ ‘fly’

Note that table (3.1) has a number of gaps. I attribute these gaps to the restrictions on the range of possible combinations of place features within onset obstruent clusters. I first discuss restrictions on clusters in which the two members have the same place of articulation. The strongest restriction on place is that two labial consonants may not occur within a cluster. Note that no such cases are listed in table (3.1). Coronals and velars are much less restricted. Still, the only occurring coronal-coronal and velar-velar clusters are those whose members differ in manner of articulation. Clusters whose members have the same manner and place of articulation, such as **td*, **dt*, **sz* and **zs*, including geminates such as **tt* or **dd*, do not occur. Clusters of two coronals that differ in manner of articulation occur frequently, as in *zd*, *st* and *sd*. However, the cluster **zt* does not occur, although its members differ in manner of articulation.

Velars are restricted in a similar way. That is, clusters whose members have the same place and manner of articulation, such as **gk* or **kg*, are not permitted. Clusters whose members differ in both place and manner of articulation, such as *kx* and *gx*, do occur.

Next, I turn to restrictions on clusters with members that differ in place of articulation. I begin by exploring clusters whose first member is a labial consonant. A labial stop as a first member may combine both with coronal stops, as in *pt*, *pd*, *bt*, *bd*, and with coronal fricatives, as in *ps*, *pz*, *bs, bz*. Labial stops combine fairly freely with velars as well. Both the combinations of labial stop-velar stop, as in *pk*, *pg*, *bk*, *bg*, and labial stop-velar fricative, as in *px*, *bx*, are attested. Clusters with a labial fricative as a first member are much more restricted. The only cluster with an initial labial fricative followed by a coronal is *ft* informs like *ftax* 'open (command),' which is an innovation and is still in free variation with *ptax*. Labial fricatives do not combine with velars at all, regardless of their manner of articulation.

We now turn to clusters whose first member is a coronal consonant, starting with coronal-labial clusters. Coronal stops may only combine with labial fricatives as in *tf*, *tv*, *df* and *dv*; there are no coronal stop – labial stop clusters. Clusters with a coronal fricative as their first member are not restricted in the same manner. The voiceless coronal fricative *s* may combine both with labial stops, as in *sp*, *sb*, and with labial fricatives, as in *sf*, *sv*. The clusters *sp* and *sb* occur only in borrowed words or foreign names, as in *sport*, *spa*, *sbaro* 'name of a restaurant', while *sf* and *sv* clusters occur in native Modern Hebrew words, as in *sfira* 'a count', *svara* 'conjecture' and in borrowed words such as *sfera* 'sphere'. The voiced coronal fricative may only combine with voiced labials to yield *zb* and *zv*. Coronal-velar clusters are less restricted than coronal-labial clusters. Coronals combine with all velars regardless of manner of articulation, yielding stop-stop, stop-fricative, fricative-stop, and fricative-fricative clusters, as in *tk*, *tg*, *tx*, *dk*, *dg*, *dx*, *sk*, *sg*, *sx*, *zk*, *zg* and *zx*.

Finally, we focus on clusters whose first member is a velar consonant. Velar stops freely combine with coronals. All stop-stop combinations occur, with the exception of **gt*, as in *kt*, *kd*, *gd*; and all stop-fricative combinations are attested clusters, as in *ks*, *kz*, *gs* and *gz*. However, the only case of a velar fricative forming a cluster with a coronal is *xt* as in *xtav* 'writing (n.).' There are no other cases of clusters beginning with a velar fricative. Velar stops may also combine with labial fricatives, as in *kf*, *kv*, *gf* and *gv*, but do not combine with labial stops. However, velar fricatives do not form clusters with either labial stops or labial fricatives.

We may conclude that restrictions on place of articulation in clusters are crucially responsible for the asymmetries in table (3.1). Relevant for the present discussion is that voicing in clusters is not restricted. While there are a number of gaps in the occurring clusters in table (3.1), they can be accounted for in terms of restrictions on possible place and manner combinations. Crucially, the four voicing combinations, which are cross-referenced with the two manners of articulation, as shown in table (3.2), are all empirically attested.

3.2. How clusters arise

In Modern Hebrew onset clusters occur in: (i) simplex forms and (ii) as a result of morphological processes. Clusters in morphologically simplex (mono-morphemic) forms are exemplified in table (3.2).

Clusters may arise from several morphological processes, most notably, from affixation. One such affixation process, plural affixation in the noun class known as Segolates, will be discussed in section 3.3.1. Plural affixation causes vowel syncope, resulting in an onset cluster. Another affixation processes is diminutive reduplication where an infix brings about vowel syncope, and, as a result, an onset cluster.[1] This process will be detailed in section 3.3.2, and is also addressed in Kreitman (2003). Although the clusters discussed in sections 3.3.1 and 3.3.2 occur in morphologically complex forms, they are not separated by morpheme boundaries, as will become evident in section 3.4, and are therefore considered clusters in light of the definition of clusters in chapter 2, section 2.1.

3.2.1. Segolates

Segolates are a class of nouns which are characterized by their vocalic patterns: both vowels of the bisyllabic form are *e* as in ***k**e**l**e**v*** 'dog' (Segol is the Hebrew name for the vowel *e*). Segolate nouns stand out because of their distinct stress pattern, which is penultimate rather than ultimate as in most other nouns.[2] Historically most Segolates had the stem pattern CaCC, CiCC and CoCC (Bolozky 1978) and an additional vowel *e* was epenthesized to break the coda cluster (for example */bokr/* → [*bˈoker*] 'morning'), resulting in bisyllabic forms. Forms such as *boker* 'morning' and *godel* 'size', which historically belonged to the Segolate class because of their stress pattern, do not belong to this class synchronically because their vocalic pattern is not *e,e* as is the case in the majority of segolate nouns. The stem patterns containing vowels such as *a*, *i* and *o* were historically included in the Segolate noun class because of their initial stress pattern. Biblical Hebrew inherited these forms from Proto-Semitic in which Segolate nouns had a similar shape (Green 2004). In Modern Hebrew, despite the stress pattern, these nouns are not classified as Segolates. In sum, Modern Hebrew

[1] There are other morphological processes that result in clusters such as feminine suffix in adjectives *šaxor* → *šxora* 'black', but only plural affixation in segolates and diminutive reduplication are discussed in this work.
[2] Modern Hebrew is an iambic language (Bat-El 1989, Ussishkin 2000). However, there are claims (Bat-El 2004, Becker 2003, Graf and Ussishkin 2003) that Modern Hebrew exhibits some trochaic patterns.

segolates are characterized by the vowel *e* in both syllables, as in *kelev* ‘dog’ and *melex* ‘king’.

Segolate nouns differ from other nouns mainly in their stress pattern, which is penultimate rather than ultimate. Stress shifts to the final syllable when a plural suffix is added. To account for the peculiar stress pattern, segolates have traditionally been analyzed as having a CVCC underlying form, which accounted for their stress pattern, (Bolozky 1978, Green 2004, McCarthy 1998). Table (3.3) reflects the challenges posed by the segolate noun group. The leftmost column shows the surface form, the middle column lists underlying forms proposed by McCarthy (1998), and the rightmost column lists underlying forms proposed by Green (2004). The underlying representations proposed by both McCarthy (1998) and Green (2004), although different, are opaquely related to their surface forms.

Table 3.3: Opacity in Segolate noun forms in Modern Hebrew

Surface form	Underlying form (McCarthy, 1998)	Underlying form (Green, 2004)
melex	melk	malk
sefer		sepr
deše/dɛšɛ	dešʔ	dašʔ

Both Green’s (2004) and McCarthy’s (1998) accounts of Segolates assume that only the first vowel is underlying, while the second vowel is epenthesized. Green (2004) also has to explain the vowel variation between what he assumes to be the underlying vowel *a* and the surface vowel *e* which, he claims, results from vowel assimilation (for a detailed account see Green 2004). Stress is assigned to the underlying vowel, which surfaces as the penultimate vowel after V2 is epenthesized.

These accounts are convenient for explaining the penultimate stress pattern in Biblical Hebrew, but fail to account for the full range of Modern Hebrew facts.

I follow Bat-El (1989) in assuming that the underlying representation of Segolate nouns in Modern Hebrew is CV_1CV_2C. The most common pattern is the one where both V1 and V2 are *e*. When a plural suffix is attached, the first vowel, V1, is syncopated and a cluster is formed at the beginning of the word as in (1) (the vowel alternation from *e* to *a* is discussed in Bat – El (1989) and is outside the scope of this work):

1)

Singular	**Plural**	**Meaning**		
semel	*smalim*	‘sign’		
perek	*prakim*	‘chapter’	OS clusters	
delet	*dlatot*	‘door’		
kelev	*klavim*	‘dog’		
sefer	*sfarim*	‘book’	[–v][–v]	
kesem	*ksamim*	‘magic’		
gezer	*gzarim*	‘carrot’	[+v][+v]	
gever	*gvarim*	‘man’		OO clusters
peger	*pgarim*	‘carrion’	[–v][+v]	
kever	*kvarim*	‘grave’		
gefen	*gfanim*	‘vine’		
dekel	*dkalim*	‘palm’	[+v][–v]	
gešer	*gešarim*	‘bridge’		

There are cases where the clusters that arise as a result of the process of affixation are impermissible. These are SS or SO sequences or certain OO impermissible sequences. In those cases, the clusters are broken by a vowel insertion, as in (2). The impermissible clusters broken by epenthesis are SS or SO clusters in (2a), and OO clusters in (2b) (on impermissible OO clusters, see section 3.2).

In (2a) the *e* in the singular is the underlying vowel while the *e* in the plural form is an epenthetic *e*, whose sole purpose is to break up the cluster. In (2b), the *e* in the singular is again the underlying vowel, but the vowel epenthesized in the plural form is *a*, rather than *e*. The quality of the epenthetic vowel is determined by the nature of the immediately preceding consonant. If the preceding consonant is a dorsal (a velar, a uvular or a glottal obstruent) the epenthetic vowel is *a*, in all other cases it is *e*. In Kreitman (2003) I claimed that the velar *x* requires a low vowel *a* following it to break up a cluster as in *xamuts* 'sour' ~ *x**a**matsmats* 'sourish', which is distinct from *yarok* 'green' ~ *y**e**rakrak* 'greenish',[3] where there is no back consonant (velar or uvular) to cause vowel lowering and thus, the epenthetic vowel is *e*.

2)

	Singular	Plural	Meaning	Ill form	Impermissible cluster
(a)	*nexed*	*nexadim*	'grandson'	**nxadim*	**nx* – SO
	nexes	*nexasim*	'property'	**nxasim*	**nx* – SO
	meser	*mesarim*	'message'	**msarim*	**ms* – SO
	remez	*remazim*	'clue'	**rmazim*	**rm* – SS
(b)	*xeder*	*xadarim*	'room'	**xdarim*	**xd* – OO
	xevel	*xavalim*	'rope'	**xvalim*	**xv* – OO
	ʔesev	*ʔasavim*	'weed'	**ʔsavim*	**ʔs* – OO

The question that still remains to be addressed is the motivation for vowel syncope when a plural suffix is added. Ussishkin's (2000) arguments for bisyllabicity in the verbal system can be extended to the nominal system as well. I argue, following Ussishkin (2000), that bisyllabicity is a strong property of Modern Hebrew not just in the verbal system but in the nominal system as well. Although it is more easily

[3] The choice of vowel is closely related to guttural lowering (Bolozky 1980).

violated in the nominal system, bisyllabicity is nonetheless a strong property of the language, which dominates constraints against deletion. The demand to maintain a bisyllabic prosodic word is so strong that when a plural suffix is added, the vowel of the first syllable is syncopated to maintain a bisyllabic prosodic word.

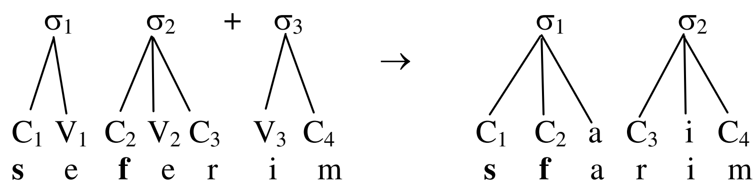


Figure 3.1: Formation of a bisyllabic prosodic word as a result of plural affixation.

As illustrated in figure (3.1), once the singular form is suffixed by the plural suffix -*im*, resulting in an additional syllable, the vowel between the first and second consonant is syncopated. The result is that C1 and C2, which were separated in the input form by V1, form a cluster in the output. Instead of having three syllables in the output, the bisyllabic optimal word size is observed by syncopating the input vowel. Thus, the output accommodates all four consonants in the input while maintaining the bisyllabicity of the prosodic word.

However, if deletion of a vowel results in an impermissible cluster, then a vowel is inserted to prevent the formation of an illicit cluster. As a result of epenthesis, the plural form is tri-syllabic, as illustrated in figure (3.2).

In this case, constraints on prosodic word size yield under the pressure to satisfy the more highly ranked constraints on syllable phonotactics. In figure (3.2) the input *nexed* ‘grandchild’ surfaces as *nexadim* ‘grandchildren’ instead of the bisyllabic but ill formed **nxadim*.[4]

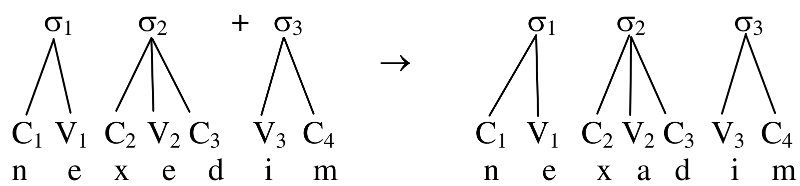


Figure 3.2: Trisyllabic plural forms in order to avoid impermissible clusters

It is also important to note that while the Segolate plural forms are morphologically complex, there is no morpheme boundary within the clusters that are formed. The clusters are formed word initially and the morpheme boundary appears word finally, as in *kelev* ‘dog’ + *im* (masculine plural) → *klav*+*im* ‘dogs’, making these derived clusters consistent with our definition of a cluster in chapter 2 section 2.1.

To summarize we can say that if and when impermissible clusters arise, they are broken up by vowel insertion. The quality of the vowel inserted depends on the immediate phonetic environment. The default epenthetic vowel is *e* unless a dorsal consonant precedes the vowel, in which case the vowel that breaks up the cluster is *a*. The obvious targets for cluster breakup are cases of sonority decline. Less obvious candidates for epenthesis, which contain a sonority rise, such as **xm*, also exist, although *gn*, *gm* as well as *kn* and *km* are perfectly well-formed clusters. This can be attributed to the fact that the cluster **xm* begins with the fricative *x*, which, for historical reasons is still quite rare and disfavored word initially.

[4] The vocalic alternation between *e* and *a* in the second syllable constitutes an opaque case in OT and is beyond the scope of this work.

Crucially, while impermissible obstruent clusters, as detailed in section 3.2, are broken up by epenthesis, as in (2), illustrated by figure (3.2), this is not the case with any of the voicing combinations, including [+v][–v] obstruent clusters, as illustrated in (1) by words like *dekel* → *dkalim* 'palm'.

3.2.2. Diminutive Reduplication

Diminutive Reduplication in Modern Hebrew exemplified in (3), results from interactions between reduplication specific constraints on the one hand (McCarthy and Prince 1995), and both general prosodic constraints on word size and constraints on syllable phonotactics, on the other, as analyzed in (Kreitman 2003).[5]

3)

Base – input	**Meaning**	**Reduplicated form – output**	**Meaning**
C_1V_1 C_2V_2 C_3		C_1C_2 a C_3 C_2 a C_3	
zakan	'beard'	*zkankan*	'small beard'
gever	'man'	*gvarvar*	'young man'
varod	'pink'	*vradrad*	'pinkish'
gezer	'carrot'	*gzarzar*	'small carrot'
kelev	'dog'	*klavlav*	'puppy/small dog'

The bisyllabicity of the reduplicated form is attributed to the general constraints on word size in Modern Hebrew. As in the case of segolates, discussed in section 3.3.1, the bisyllabic shape of the reduplicated form is captured by general constraints on word size in Modern Hebrew (Ussishkin 2000), which require that the phonological word be minimally and maximally bi-syllabic. These constraints which govern word size in Modern Hebrew are directly responsible for the size of the output of the

[5] For an alternative account see Graf (2002).

diminutive reduplication. It is due to both minimal and maximal restrictions on word size that the first syllable of the input to diminutive reduplication is not realized in the output. The reduplicant corresponds to the second syllable of the base with its original input vowel "overwritten" with the reduplicant's prespecified vowel. As shown in figure (3.3), *zakan* yields the reduplicant *kan*, which, when affixed to the base *zakan* forms a bisyllabic reduplicated form as in *zkankan*. In Kreitman (2003) I argue that the reduplicant is the first occurrence of *kan*, underlined in z*kankan* and that it is an infix as illustrated in figure (3.3) (for a full analysis see Kreitman (2003)):

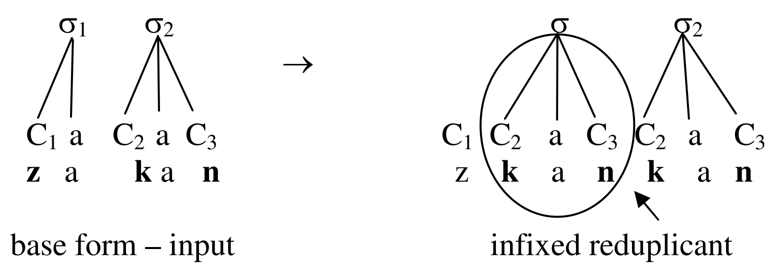


Figure 3.3: Syllabification of the input *zakan* and the reduplicated output *zkankan*.

Faithfulness relations between the base and the reduplicant are responsible for the vocalic overwriting of the base vowel with the vowel from the reduplicant. While the reduplication process alters the vocalic pattern of the input, it strongly demands preservation of the consonantal material. The first vowel of the base is lost in the reduplicated form due to pressures of constraints on word size. As a result, C1 and C2, which are separated by a vowel in the input, form a cluster in the reduplicated form (schematized in figure (3.3)).

However, as we have seen in the case of suffixation in Segolates, phonotactic constraints on syllable wellformedness override prosodic size constraints. Prosodic constraints on word size require that the input material fit into a bisyllabic "template" causing loss of phonological material. When impermissible clusters arise due to loss of segmental material, phonotactic constraints require a breakup of those clusters, and emerge as dominant. Impermissible clusters are broken up by vowel insertion as in (4), resulting in trisyllabic outputs.

4)

	Base – input	Meaning	Reduplicated form – output	Meaning
	$C_1V_1\ C_2V_2\ C_3$		C_1C_2 a C_3 C_2 a C_3	
(a)	*lavan*	'white'	*le.van.van*	'whit-ish'
	yarok	'green'	*ye.rak.rak*	'green-ish'
(b)	*xamuts*	'sour'	*xa.mats.mats*	'sour-ish'

In (4), the phonotactic violation is resolved by vowel insertion. The relevant constraints on cluster phonotactics dominate constraints which prohibit epenthesis. The vowel that breaks up the illicit cluster is *a* in the environment of dorsals, as in (4b), and the default vowel *e* elsewhere, as in (4a). Phonotactic constraints take precedence over constraints responsible for prosodic word size. The constraints on clusters that arise as a result of diminutive reduplication are the same as those imposed on clusters that arise as a result of segolate plural forms:

5) Impermissible clusters:

(a) Sonorant - Obstruent clusters

(b) Sonorant - Sonorant clusters

(c) Non-occurring Obstruent - Obstruent clusters, as detailed in section 3.2

As evident from (5), clusters that arise as a result of an affixation process, diminutive reduplication or plural suffixing, are subject to the general constraints on clusters in the language. However, [+v][-v] clusters are not broken up by epenthesis as was demonstrated in (1) and illustrated by figure (3.3).

3.3. Onset clusters or appendices?

In this chapter we saw two types of consonantal sequences, those that may not form an onset cluster, and are broken up by vowel epenthesis, and those that may form an onset cluster. Clusters broken up by vowel epenthesis are those that violate phonotactic constraints on onsets. First, such clusters include sonorant clusters SS and sonority reversed clusters SO, those impermissible in a type 2 language like Modern Hebrew, which allows only OS and OO combinations of the feature [sonorant]. Second, such clusters include OS and impermissible OO clusters, as detailed in section 3.2; for example the clusters **xm* and **xd*, which are an OS and OO clusters respectively, are still broken by vowel epenthesis. In chapter 2 I provided evidence that [+v][−v] clusters are "regular" clusters in Modern Hebrew. This is further confirmed by the evidence provided in this chapter: when such clusters arise by virtue of morphological processes, they are not broken by vowel epenthesis.

Lombardi (1991) relates sonority to voicing. For her (as well as for Levin 1985 and Steriade 1982 *among others*), voiced obstruents are more sonorous than voiceless obstruents. She further claims that [+v][−v] obstruent clusters are in fact clusters with falling sonority and therefore parallel SO clusters. In this section I will provide phonological arguments against the claim that [+v][−v] obstruent clusters have a special phonological status and consequently, a phonological representation that differs from the representation of other obstruent clusters. I will argue that [+v][−v] obstruent clusters are representationally of the OO type rather than of the SO type, as proposed by Lombardi (1991) and Morelli (1999). That is, I will argue that solutions that have been proposed in the literature to resolve the problem of sonority reversed clusters do not apply in the case of [+v][−v] obstruent clusters. The phonological representations that have been proposed for SO, such as the appendix, are not needed for [+v][−v] obstruent clustering.

In the discussion of onset clusters in Modern Hebrew presented in this chapter, as well as in chapter 1, I have shown that the following consonant sequences are disallowed in onset position in Modern Hebrew:

6)

1. Sonorant - Obstruent clusters
2. Sonorant - Sonorant clusters
3. Sequence of two labials.

If any sequence listed in (6) occurs in word initial position, it is broken by an epenthetic vowel. Clusters that are permissible, surface as onset sequences, with no interfering vowel. This is the case with [+v][−v] obstruent clusters. In what follows I will argue that [+v][−v] sequences form regular obstruent clusters, and need not be analyzed in terms of a word-initial appendices. I first outline arguments in favor of an appendix based analysis, and then give counterarguments why the appendix is not an appropriate phonological representation for [+v][−v] obstruent onset clusters in Modern Hebrew.

3.3.1. Appendix

To resolve counter examples to the Sonority Sequencing Principle (SSP), that is, to reconcile the existence of sonority reversed clusters, with the overwhelming observation that languages tend to prefer syllables which rise in sonority towards the peak and fall in sonority towards the margins, SO clusters are treated as special clusters with a special phonological representation. One solution that has been proposed in the literature is that of the appendix (for a comprehensive review of the appendix see Vaux 2004). When a sequence of two consonants of falling sonority is found in word initial onset position then the first member of the sequence is associated with a prosodic segment other than the syllable. Although, the prosodic level to which the first segment is attached has been subject for debate, proposals ranging from feet to Prosodic Word among others, I will only focus on the Prosodic Word as a target for association. An example of syllabification under this theoretical framework is given in figure (3.4a). However, if the two segments rise in sonority, then they are both associated with the same syllable and neither is attached to the Prosodic Word directly as in figure (3.4b), where we see an illustration of the English word *three*. The SSP predicts the internal organization of a syllable and its components but does not pertain to the internal organization of prosodic domains higher than the syllable. Segments outside the syllable domain are not subject to the SSP as illustrated in figure (3.4a), where we have an example of syllabification of the Russian word *rtɨ* 'mouths' illustrating the solution of the appendix. The word *rtɨ* is an example of a sonority reversed onset cluster where the first member of the cluster *r* is more sonorous than the second member of the cluster, *t*. In figure (3.4a), the segment C1, the first, less sonorous, member of the cluster, is directly linked to the Prosodic Word level, while C2, the obstruent *t*, is associated with the syllable. The fact that the sonorant *r* is attached to a higher level results in the sequence *rt* not being tautosyllabic and hence not a cluster (according to the definition in chapter 2 section 2.1.1).

The crux of the proposal is that in a sequence of two onset consonants, if a sonorant precedes an obstruent, the sonorant is not syllabified in the same syllable with the obstruent. Rather, the "extrametrical" consonants, which are attached at a different prosodic level, are not a part of the syllable but are attached directly to the

Prosodic Word (or another prosodic domain). In this manner the problem of the offending reversed sonority sequence is resolved as the SO sequence no longer forms an onset cluster.

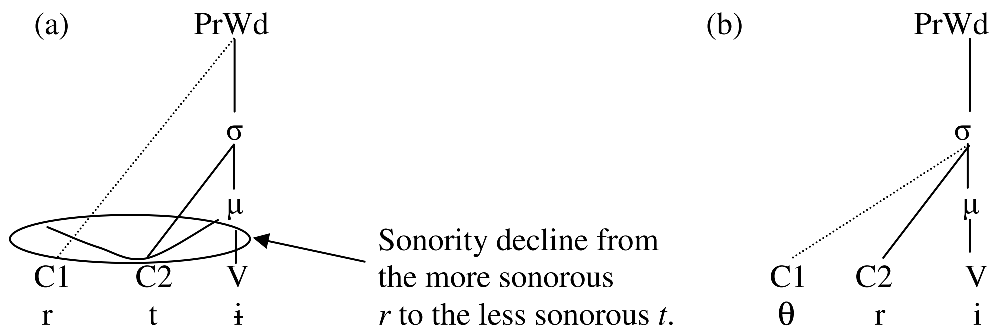


Figure 3.4: Examples of syllabification of falling (a) and rising (b) sonority.

If [+v][−v] clusters parallel clusters with sonority reversals, then they should be syllabified as in figure (3.4a). I will argue that [+v][−v] clusters do not have the phonological representation in figure (3.4a); rather, they have the phonological representation illustrated by figure (3.4b).

3.3.2. Arguing against the appendix

Recall that both SO and SS clusters are strictly forbidden in Modern Hebrew. When SO or SS sequences arise through morphological processes of plural suffixation or diminutive reduplication, they are broken by vowel epenthesis. If clusters with falling sonority form a sequence of an appended consonant and a consonant attached to the syllable, as in figure (3.4a), then there should not be a need for vowel epenthesis. That is, if the word initial onset sequence is not tautosyllabic, as would be the case if one of the segments was in the appendix, then there would be no need to break up the consonant sequence, and SO and SS sequences would be possible in word initial position in Modern Hebrew.

Let us assume, for the sake of argument, that diminutive reduplication has two different prosodic outputs, one for illicit clusters when they arise, and the other for all remaining cases. The two representations in figure (3.5) illustrate the two syllabifications. In figure (3.5a) we have a representation for all permissible clusters, while in figure (3.5b), we have a representations for impermissible clusters, say, those that fall in sonority.

In figure (3.5a) the output forms a "real" cluster, as defined in chapter 2 section 2.1. Outputs of the type depicted in figure (3.5a) are presumably reserved for forms such as *kelev* 'dog' → *klavlav* 'puppy', where the output form has an initial cluster with a sonority rise. Structures of the type depicted in figure (3.5b) arise when an illicit cluster is formed, for example, a cluster with falling sonority as in the word *levanvan*, where the initial cluster is **lv*. However, if the offending cluster had the prosodic structure of figure (3.5b), then the sequence is not tautosyallbic and thus, should not be broken by an epenthetic vowel. The fact that the cluster **lv* is broken up by epenthesis strongly suggests that figure (3.5b) is not a plausible output for the process of diminutive reduplication.

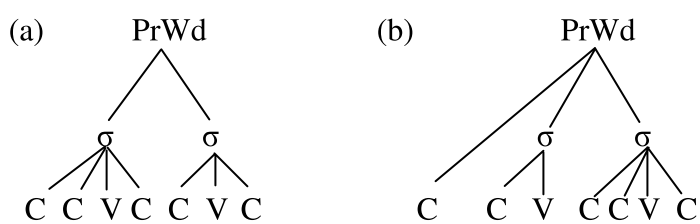


Figure 3.5: Possible phonological representations for initial consonants sequences.

Additional support for the claim that the phonological representation in figure (3.5a) is the only possible output form of diminutive reduplication comes from further facts about the syllabification of forms such as *levanvan*. In such cases the cluster is broken up by vowel insertion, as in *lavan* ‘white’ → *levanvan* ‘whitish’, when the word occurs in isolation. However, if the same word is preceded by a specifier such as the definite article *ha* ‘the’, then *l* can be resyllabified with the specifier, as in figure (3.6). When *l* becomes the coda of the previous syllable, the function word *ha*, then the sequence *lv* does occur. However, as shown in figure (3.6) the sequence *lv* is heterosyllabic, and as such does not form a cluster as defined in chapter 2 section 2.1. Because the sequence *lv* is not tautosyllabic, it is not an onset cluster of declining sonority, and no vowel epenthesis is necessary.

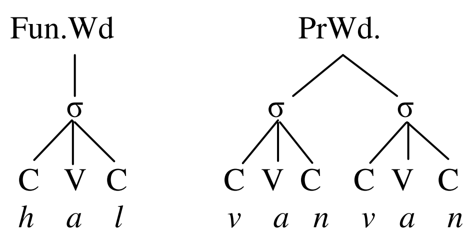


Figure 3.6: Resyllabification of the word *levanvan* with a preceding specifier.

Without resyllabification of the onset *l* as the coda of the specifier, the sequence **lv* cannot occur, as illustrated in the ill-formed representation in figure (3.7):

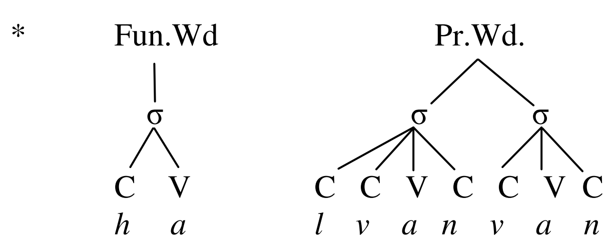


Figure 3.7: Illicit syllabification of specifier + the word *levanvan*.

Resyllabification is optional, however, and if it does not take effect, the output will be *ha.le.van.van*, with the sequence **lv* broken by an epenthetic vowel, which is distinct from the resyllabified version *hal.van.van* shown in figure (3.6).

However, one must take into account that this occurs *only* in cases of sequences with falling sonority and is completely blocked in cases where there is no falling sonority.

By contrast, there is no evidence that forms such as *kelev* 'dog' → *klavlav* 'puppy' are subject to resyllabification. As shown in figure (3.8), the syllabification of *ha+klavlav* is *ha.klav.lav*. There is no evidence for the syllabification *hak.lav.lav*, as in figure (3.9).

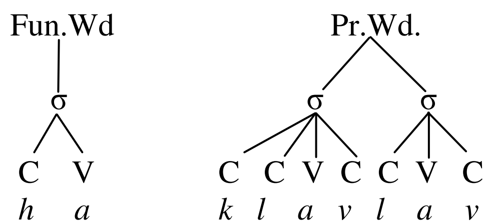


Figure 3.8: Licit syllabification of *haklavlav*.

*

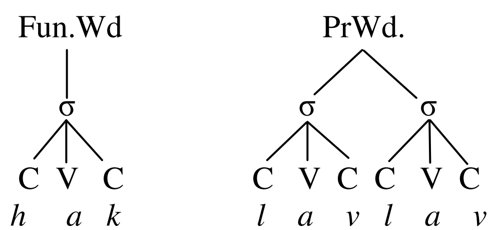


Figure 3.9: Illicit resyllabification of *haklavlav*.

In sum, when sonority reversals arise, and must be resolved, there are only two options to resolve an illicit SO clusters. One way to resolve sonority reversed clusters is to eliminate the cluster status of C1 and C2 by epenthesizing a vowel between the two segments as in figure (3.10a). Another way to resolve sonority reversals is to resyllabify the offending sonorant C1 with the previous specifier as in figure (3.10b).

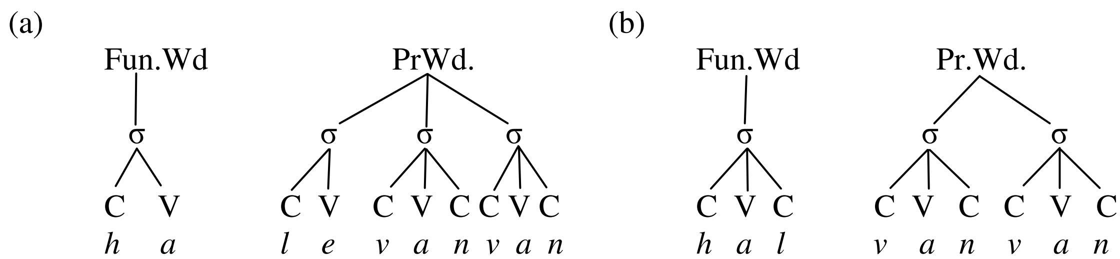


Figure 3.10: Possible phonological syllabification for SO initial clusters.

Proponents of the appendix solution might argue that the fact that the first consonant of the sequence, *l*, is able to detach from the Prosodic Word and adjoin the previous function word strengthens the argument for it being an appendix. An appendix easily detaches and resyllabifies. However, because the illicit SO onset cluster must be eliminated as it constitutes a cluster with falling sonority, it is either broken by vowel epenthesis or by resyllabification. The fact that an illicit SO sequence is not tolerated, suggests that all occurring onset sequences are tautosyllabic onset cluster. The set of occurring onset clusters in Modern Hebrew, a type 2 language regarding sonority clustering, excludes sonority reversed clusters.

3.3.3. [+v][−v] obstruent sequences are clusters as well

The syllabic status of [+v][−v] obstruent clusters is nondistinct from other permissible OO clusters. That is, a sequence of [+v][−v] obstruents is tautosyllabic and behaves as a cluster as defined in chapter 2 section 2.1. Figure (3.11a) illustrates the syllabification of *zkankan* ‘small beard’, which contains an initial *zk* sequence. In figure (3.11b) we see the only possible syllabification of *zkankan* when preceded by the definite article *ha.* The syllabification in (3.11c) is not a possible syllabification. The syllabification of this form is identical to the syllabification of OS clusters such as *kl* in the word *klavlav* ‘puppy’, in figure (3.8).

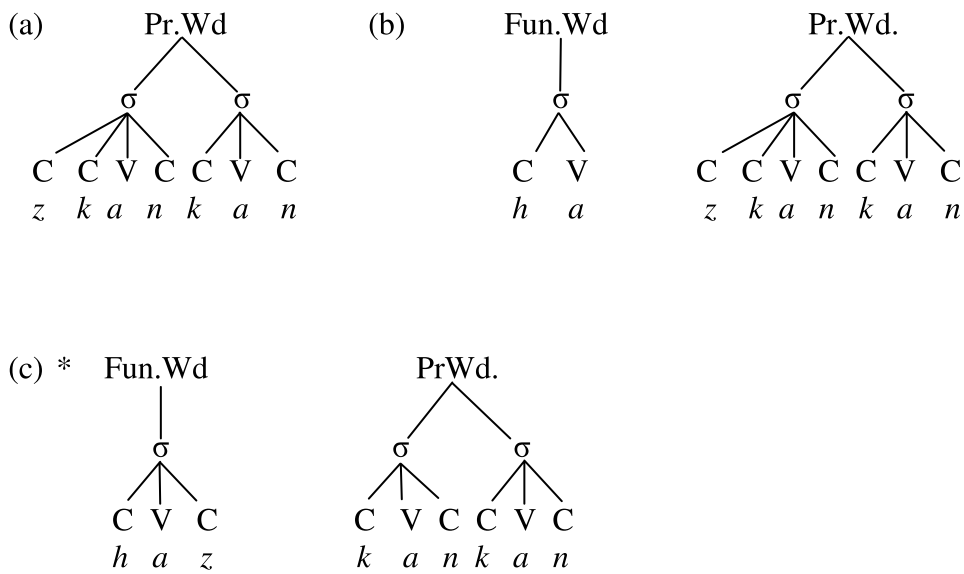


Figure 3.11: Syllabification of *zkankan* ‘small beard’.

The fact that forms like *zkankan* ‘small beard’ behave like *klavlav* ‘puppy’, strongly suggests that the syllabic status of [+v][−v] obstruent clusters does not differ from the syllabic status of SO and other OO clusters. By contrast, SS and SO clusters (as well as impermissible OO clusters, as detailed in section 3.2) are broken by vowel epenthesis. If [+v][−v] clusters behaved like clusters of falling sonority, they would be broken by vowel epenthesis, which they are clearly not. Moreover, if it were the case that the first member of a [+v][−v] cluster was an appendix, and attached at the Prosodic Word level, then it should have been easily resyllabified with the previous function word. As illustrated by figures (3.6) and (3.10), it is possible to resyllabify material as the coda of the function word *ha*. Since the first member of the onset cluster cannot be resyllabified as a coda of the preceding function word, it suggests that [+v][−v] obstruent sequences function as tautosyllabic sequences, that is, as onset clusters, and that the appendix is not necessary for their proper syllabic representation.

In sum, the fact that [+v][−v] obstruent clusters behave as OS and other OO clusters provides crucial evidence against the claim that these clusters pattern like SO clusters, as suggested by Lombardi (1991) and Morelli (1999).

3.4. Summary

In this chapter I presented a case study of Modern Hebrew obstruent clusters, with special emphasis on voicing in clusters. I have shown that all voicing onset sequences pattern alike, which argues against a special phonological status of [+v][−v] obstruent onset clusters. In arguing against the special phonological status of [+v][−v] obstruent onset clusters, I specifically argue against representing such clusters by invoking a syllable initial appendix. In sum, I argue that all permissible onset clusters in Modern Hebrew, including [+v][−v] obstruent clusters, are proper tautosyllabic clusters as defined in chapter 2. I have provided phonological evidence from two cases of prosodically governed morphological processes of affixation: plural suffixation in the Segolate noun class and diminutive reduplication affixation, to argue against the need for a syllable initial appendix in the representation of [+v][−v] obstruent onset clusters. In chapter 5 I will present phonetic data on voicing in onset clusters in Modern Hebrew, including the typologically rare [+v][−v] obstruent onset clusters.

CHAPTER 4

PHONETIC ASPECTS OF VOICING AND ITS REALIZATION IN OBSTRUENT CLUSTERS

4.0. Introduction

In this chapter I address the phonetic realization of voicing in onset clusters. I will consider three major aspects crucial to understanding voicing: physiology of voicing, its acoustics and the impact of phonological and phonetic environment on the realization of voicing. All these aspects of the phonetics of voicing will play a role in the analysis of the Modern Hebrew data on voicing in obstruent onset cluster to be presented in chapter 5.

The phonetic realization of voicing is heavily influenced by both the inherent nature of the segment and the immediately surrounding phonetic environment. In order to address the realization of voicing in obstruent clusters, we must first explore how voicing is produced physiologically and what are the physical conditions required for voicing to be realized.

In section 4.2 of this chapter I explore the physiology of voicing, that is, the muscular activity necessary for the production and maintenance of voicing. In section 4.3 I provide an explanation of how both voice production and voice maintenance are implemented in obstruents. I address the acoustic properties correlated with voicing gestures in section 4.4, and the perception of voicing cues in various phonetic environments in sections 4.5 and 4.6. I end this chapter with predictions about the theoretically possible realizations of voicing in clusters in Modern Hebrew, based on evidence regarding the realization of voicing in singleton word initial obstruents which will be empirically tested in chapter 5.

4.1. Voice Production

The muscles responsible for voicing distinctions are the Posterior Crico-Arytenoid muscle (PCA), which is responsible for voicelessness, and Interarytenoid (IA), Lateral Cricoarytenoid (LCA) and the Thyroarytenoid (TA) muscles, all of which are responsible for voicing. The vocalis muscle helps control the mass and stiffness of the vocal folds, which plays an important role in the realization of voicing. The vocal folds are attached to the arytenoid cartilages which are attached to the IA muscle. The IA muscle is responsible for approximating the arytenoid cartilages which are positioned appropriately for air flow to cause vibration of the vocal folds. The IA muscle and the PCA muscle work complementarily to each other. The PCA muscle activity is suppressed during voiced segments, but increases during voiceless segments. By contrast, the IA muscle is active during voiced segments, but its activity is decreased during voiceless ones. The LCA remains suppressed during both voiced and voiceless obstruents but is recruited when extra tension is needed and has therefore been grouped with the vocalis muscle (Hirose 1997, Löfqvist 1980, Stevens 1998). Abduction of the PCA muscle results in the spreading of the vocal folds, which are kept open allowing air to flow freely from the lungs through the vocal tract, the result being that the vocal folds do not vibrate. The lack of vibration of vocal folds results, of course, in voicelessness.

According to Hirose (1997) vibration of the vocal folds has four stages: (i) The adducting muscles, LCA, IA and TA, are activated, which brings the anterior ends of arytenoid cartilages together, approximating the vocal folds. The procedure is accompanied by suppression of the abductor muscle and occurs only when vocal fold vibration is initiated, as opposed to involuntary and passive vibration, to be discussed later in this section. (ii) Air is forced through the vocal tract from the lungs causing the vocal folds to get “sucked” together by Bernoulli’s aerodynamic law coupled with the elasticity of the tissues. (iii) When the vocal folds are stuck together they create a closure which causes a cessation of air flow from the lungs and results in a rise in sub-glottal pressure. (iv) When the sub-glottal pressure is greater than the supra-glottal pressure, the folds are blown apart and the air escapes into the supra-glottal space.

Once the air escapes into the oral cavity, the trans-glottal pressure falls and the vocal folds return to their adducted position. This cycle repeats itself as long as the range of required pressure differences between the sub-glottal and supra-glottal cavities is maintained. Or to quote Stevens (1998): "Vocal fold vibration is, of course, possible only if a transglottal pressure is maintained, so that a positive airflow occurs through the glottis from the lungs." However, initiation of voicing, or bringing the vocal folds from a non-vibrating state to a vibrating state on the one hand, and maintaining vocal fold vibration on the other, require different physiological mechanisms. Below, I first address voicing initiation and then voicing maintenance.

Initiating vibration of the vocal folds requires an increased amount of sub-glottal pressure and voluntary activation of the adducting muscles. The speaker must actively adduct the appropriate muscles to bring the vocal folds together to allow the cyclic process of vocal fold vibration to begin. Once vocal fold vibration is initiated, vibration must be maintained for several periods.

In order to maintain vocal fold vibration, pressure difference must exist between the supra-glottal cavity and the sub-glottal area. The sub-glottal pressure must be greater than the supra-glottal pressure in order to blow apart the vocal cords, which are initially "stuck together". During the production of vowels and sonorants, which are generally voiced,[1] the sub-glottal pressure is greater than the supra-glottal pressure because the mouth is kept sufficiently open during the production of these segments to cause supra-glottal pressure to remain lower than the sub-glottal pressure. Under these

[1] Excluding voiceless sonorants and voiceless vowels in such language as Icelandic and Japanese.

conditions it is quite easy to maintain vibration of the vocal folds with no need for any additional laryngeal maneuvers. The large aperture and the low pressure in the supra-glottal area make maintaining pressure differences quite natural. In this case vocal fold vibration is passive on the speaker's part and occurs naturally due to proper physical and physiological conditions. While maintenance of vocal fold vibration in sonorants is quite natural, in obstruents it is more complex due to the rising supra-glottal pressure which results from the occlusion necessary for obstruents. The production of obstruents involves creating a complete closure for stops or a very narrow constriction for fricatives, resulting in gradually increasing supra-glottal pressure. Because of the rise in supra-glottal pressure in obstruents, vocal fold vibration is maintained by employing additional strategies that create sub-glottal pressure higher than the supra-glottal pressure. In section 4.2 I address the physiology of voicing in obstruents in detail.

4.2. Voicing in obstruents

As already noted, the narrow constriction in the vocal tract during the production of obstruents causes an increase of pressure behind the occlusion, which in turn causes a decrease of trans-glottal pressure. The pressure in the vocal tract rises gradually until it is almost equalized with sub-glottal pressure. When transglottal pressure decreases significantly, the vocal folds cannot maintain vibration. Due to the reduced transglottal pressure, the strength of vocal fold vibration is gradually reduced and if the supraglottal pressure equalizes with the sub-glottal pressure, vibration can cease altogether, even before the release of the constriction. This is true for both stops and fricatives. However, because of the different degrees of constriction for stops and fricatives, pressure equalizing is slightly different. In 4.3.1 I detail the physiology of voicing in stops and in fricatives, and discuss strategies for resolving the difficulty in maintaining vocal fold vibration in obstruents. In 4.2.2 I discuss voicelessness in these sounds.

4.2.1. Voicing in stops and fricatives

Stop consonants are produced in three stages: (i) creating a complete occlusion in the vocal tract (ii) maintaining the occlusion and impeding airflow causing pressure to build up, and (iii) releasing the closure and the pressure build up. During the closure there is no air flowing through the vocal tract and the intra-oral pressure increases. As a result, the trans-glottal pressure decreases fairly quickly below the threshold needed to maintain voicing and, with no special adjustments, glottal pulses can cease fairly quickly after closure, or even before the release. Once the oral occlusion is released, the transglottal pressure increases once again due to the drop in pressure in intra-oral cavity and voicing may resume. Thus, voicing in stops is less favorable because it is more difficult to produce, a property which is inherent to stops:

> [A]ccording to the myoelastic-aerodynamic theory of phonation…the vocal folds will oscillate only when there exists an adequate pressure drop and airflow across them. During stops, no air exits the mouth or nose, so that this condition is not obviously met. That observation suggests that voiced stops might be more difficult to produce, and thereby less 'natural; than their voiceless counterparts. (Westbury and Keating 1986:146)

Fricatives are formed by a narrowing of the articulators in the vocal tract without ever reaching complete closure, thus allowing a constant flow of air through the vocal tract. Because the narrowing in the vocal tract is so small, it causes the air flowing from the lungs through the vocal tract to come out in a turbulent noise, perceived as frication noise. However, because there is no complete occlusion in the vocal tract and airflow is continuous, the pressure differences between the sub-glottal and supra-glottal cavities take longer to equalize. Therefore, vocal fold vibration can be maintained for a longer period of time and is less subject to decay of voicing.

In sum, in order to maintain the pressure differences necessary for the vibration of the vocal folds in obstruents, special adjustments must be employed. One such adjustment is to diminish the volume of the sub-glottal area, and simultaneously increase the size of the supra-glottal area. This can be achieved by lowering the larynx. When the larynx is lowered, it slides over the convex surface of the spine resulting in a downward tilt of the cricoid cartilage. A byproduct of the sliding of the larynx over the convex surface is the tilting of the cricoid cartilage which results in a shortening and thus slacking and thickening of the vocal folds (Honda et al. 1993, Stevens 1998).[2] A consequence of the slacking of the vocal folds is that vocal folds can vibrate with reduced sub-glottal pressure, which may in turn facilitate maintaining vocal fold vibration for longer stretches of time during the production of obstruents. Thus, larynx lowering may be a good strategy for expanding the vocal tract and maintaining vocal fold vibration during the production of obstruents.

Other possible adjustments for expanding the vocal tract for the purpose of maintaining vocal fold vibration in obstruents include raising the soft palate, advancing the tongue root, or pulling the tongue dorsum and blade down toward the floor of the mouth (Stevens 1998, Westbury 1983). This last movement of the tongue root may also contribute to enlargement of the pharynx, which results in an enlargement of the supra-glottal cavity and thus aids in maintaining vocal fold vibration. As stated by Westbury (1983:1324), "It should be possible to affect voicing duration in stops by inhibiting or recruiting certain of the muscles bounding the vocal tract – e.g., the superior, middle and inferior pharyngeal constrictors, which bound the

---

[2] See similar effect on *f*0 in Honda *et al*. 1999.

pharynx posteriorly and laterally; or the buccinator and risorius which line the cheeks." (see also Westbury and Keating 1986)

4.2.2. Voicelessness in stops and fricatives

Initiation of voicelessness, just like initiation of voicing, involves a voluntary constriction of the relevant muscles. In order to maintain voicelessness, the vocal folds are actively kept apart through the entire duration of the voiceless segment. This active gesture has often been linked to the phonetic feature [spread glottis]. However, while maintaining voicelessness requires an active laryngeal gesture through the entire duration of the segment, maintaining voicing in obstruents involves an initial activation of the laryngeal muscles and additional laryngeal adjustments. That is, voicing maintenance in voiced obstruents requires more laryngeal maneuvers than maintaining voicelessness in obstruents. In sonorants, however, voicing may be maintained passively for a long time.

It is important to note that this active type of "voicelessness" should be distinguished from cases of failure to initiate voicing due to insufficient pressure differences or failure to activate the muscles responsible for voicing. Both active "voicelessness" (when the vocal folds are actively kept apart) and failure to produce voicing result in what appears as voicelessness on the spectrograms. However, the two sources of voicelessness, (i) active voicelessness with a voiceless phonological target, and (ii) failure to produce voicing (despite a phonologically voiced target), should be kept separate both in phonetics and in phonology.

Now that we have presented the physiological mechanisms responsible for voicing, crucial to our analysis are the acoustic properties correlated with these physiological mechanisms. In the next section I outline the acoustic correlates of voicing and the role they play in the perception of voicing.

4.3. Acoustic correlates of voicing

Acoustically voicing is characterized by a strong presence of low frequency spectral energy, represented as a "voice bar" along the bottom of the spectrogram. The periodicity which characterizes the voiced segments and is superimposed on the noise source is lacking in voiceless segments (Jackobson, Fant and Halle 1952, Stevens *et al.* 1986). For stop consonants, any vocal fold vibration which precedes the burst is perceived as voicing and for fricatives it is vibration of vocal folds during the production of the segment that cues voicing. Although vocal fold vibration is the most salient cue for voicing in obstruents, there are additional secondary cues, which may also cue voicing. Other cues for voicing in obstruents include duration of closure (for stops), low F1 values in transitions (Lisker 1975), a sharp drop in $f0$, transitions between segments, amplitude of stop burst, and the duration of the preceding vowel (Chen 1970, Lehiste 1970 *among others*). The presence of one or more of these secondary cues can cause listeners to perceive a segment as voiced even in the absence of a voice bar. For example Stevens and Blumstein (1981) have shown that lowering F1 can change listeners' perception of voicing, and cause listeners to hear voicing even in the absence of a voice bar. On the other hand, a sharp drop in $f0$ can shift listeners' judgments towards voicelessness. Formant transitions of velars tend to require longer VOT than other stops to be heard as voiceless (Abramson and Lisker 1973). Higher amplitude of aspiration tends to correlate with "voicelessness" while longer duration of a preceding vowel may result in perceiving the segment as voiced (Chen 1970, Lehiste 1970).

Secondary cues are highly sensitive to the immediate phonetic environment and are therefore much less reliable. Cues such as F1 and $f0$ transitions can contribute to the perception of voicing but are easily influenced by vowel quality and therefore their effectiveness is limited. Consequently, F1 and $f0$ transitions, just like other secondary cues, are not reliable cues for voicing.

On the other hand, the primary cue for voicing, vocal fold vibration, is compatible with any phonetic environment. In the next section I outline the phonetic cues associated with voicing and their perception in various environments. I address the problems of perception of voicing in various phonetic contexts and the influence of the immediate phonetic environment on various voicing cues.

4.4. Cues for voice perception

Auditory cues are perceived best when they are in a phonetic environment which emphasizes their unique distinctive phonetic characteristics. This is most apparent when auditory cues occur in an environment that is as acoustically distinct from the target cues as possible. For example, the frication noise of a stop burst is perceived well before a vowel because it usually follows a period of silence, which is the result of the closure portion of the stop preceding the release. During the closure portion no air escapes the vocal tract, resulting in a stretch of silence. The high frequency noise of the release burst is as different as can be from the silence that precedes it. However, if the stop is followed by a fricative, which contains frication noise in similar frequencies to those exploited by the release burst, the salience of the release burst of the stop loses its auditory advantage and can become very nearly assimilated in the frication noise of the following fricative. Therefore, acoustically, the most advantageous environment for stops is before a vowel because the vowel provides a "cushion" for the full realization of various acoustic cues of the stop burst. Aspiration before a vowel is salient because it exploits a different range of frequencies than those needed for the vowel. The aspiration noise is less likely to be auditorily "lost" before a vowel than before a fricative. Proper phonetic environment can also attenuate other perceptual cues such as F1 transition and $f0$ which can provide additional acoustic information about phonetic properties of segments.

The main acoustic cue for voicing is vocal fold vibration as discussed in previous sections of this chapter. It is, however, not the only cue for voicing and other, secondary cues, listed in section 4.3, also exist. Vocal fold vibration is the most stable acoustic cue and can occur in all phonetic environments, unlike secondary cues which are much more sensitive to their immediate phonetic environment. Wheeler (2005) (following Steriade 1997) divides voicing cues into three types: cues that are internal to the segment, or internal cues; transition cues following the segment, or offset cues; and transitional cues going into the segment, or onset cues. Clearly, offset cues are conditioned by a following segment whereas onset cues are sensitive to the preceding segment. Wheeler (2005:4), following Steriade (1997), divides these cues as in Table (4.1), (where S stands for sonorant, O for obstruent, and # for word boundary):

Table 4.1: Voicing cues in various word positions.

	Internal cues		**Offset cues**		**Onset cues**		
Context	**closure voicing**	**closure duration**	**S_1 duration**	**F1 values in S_1**	**burst duration/ amplitdue**	**VOT**	**$f0$ and F1 transitions in S_2**
1. S_1 _ S_2	✓	✓	✓	✓	✓	✓	✓
2. O _ S_2	✓	✓			✓	✓	✓
3. # _ S_2	✓				✓	✓	✓
4. S_1 _ #	✓	✓	✓	✓	(✓)		
5. S_1 _ O	✓	✓	✓	✓			
6. O _ O	✓	✓					
7. O _ #	✓	✓					
8. # _ O	✓				[✓]		

S_1 = preceding sonorant S_2 = following sonorant O = obstruent

Table (4.1) outlines clearly which cues are available in which context. For example, the cue "S_1 duration" is available only when obstruent voicing occurs after

S_1. The cue for F1 transition is only available when the obstruent is followed by a sonorant which can provide an environment for F1 to transition into. Thus, an obstruent in word initial prevocalic position, that is, in environment (3), has the following voicing onset cues: burst duration, amplitude, VOT and *f*0 and F1 transitions into the following sonorant as well as internal cues which include closure voicing. Since the focus of this work is voicing cues in word initial obstruent clusters we consider the relevant environment for this context, environments, (2) and (8) in table (4.1). For the first member of an obstruent cluster C1, there are very few cues for voicing, as reflected in environment (8). The cues available for C1 are closure voicing and burst amplitude for stops and consonant duration and voicing duration for fricatives. The cues available for voicing in C2, are reflected in environment (2) in table (4.1). They are closure voicing, closure duration, burst amplitude and aspiration, consonant duration, voicing duration (for fricatives), *f*0 and F1 transitions. Clearly, a second member of an obstruent cluster, which is next to a vowel, enjoys greater perceptual salience and a wider range of cues as it benefits from cues which are not available to the first member of the cluster. Moreover, C2 has the advantage of both onset cues and offset cues. That is, if one or more voicing cue is lost, voicing can still be recovered by relying on a range of secondary cues. C1 on the other hand, has only offset perceptual cues which means there is a much smaller range of cues to rely on and each cue becomes much more crucial to the perception of voicing. Losing one cue in C1 can influence the perception of voicing since there are fewer secondary cues that can be exploited to recover voicing. The first segment of the cluster is quite disadvantaged in the sense that there are more secondary cues signaling voicing in C2 than in C1.

However, the immediate phonetic environment influences not only the perception of voicing but also the production of voicing. In section 4.5 I address the production of voicing in various phonetic contexts, in order to understand why voicing cues may be lost in certain environments and not in others, and in order to explain which are favorable environments for the production of voicing.

4.5. The effect of phonetic environment on the production of voicing

The ease of vocal fold vibration is dependent on the immediate surrounding phonetic environment. Westbury and Keating (1986) point out two factors that play a role in the "naturalness" of voice realization, the inherent nature of the segment and the immediate phonetic environment in which that segment is realized. Keating and Westbury (1986) demonstrate that various positions in an utterance may be more or less favorable for voice production. An intervocalic position is the "simplest" position for realization of voicing distinction. "It is more likely for a stop to be voiced or voiceless in an articulatorily 'simple' vowel + stop + vowel string…" (Keating and Westbury 1986:149). Physiologically, the air from the lungs is already flowing during the production of the vowel and the mouth is open, thus supra-glottal pressure is kept low and there is no need to build up pressure to initiate vocal fold vibration. The sub-glottal pressure is derived from the elastic recoil of the lungs requiring no further adjustment to the sub-glottal pressure. No muscularly induced changes in the supra-glottal volume occur during the production of a vowel and usually the vocal folds are adducted for voicing. These conditions induce pressure changes above and below the glottis and therefore enable vocal folds to continue to vibrate with no additional effort on the speaker's part. In fact, in order to produce a voiceless segment intervocalically, additional laryngeal adjustments need to be made. Thus, inter-vocalic position is favorable for voicing in stops given the physiological state of the vocal tract and the laryngeal configuration that has already been initiated for the duration of the vowel.

4.5.1. Voicing in initial position

However, this is not the case with initial position. In order to produce voicing in utterance initial position, pressure must be built up first. According to Keating and Westbury (1986:153), "… stop voicing should be more likely utterance-medially than initially or finally, simply because that pressure difference in the former environment tends to be somewhat greater than in the latter environments." In fact, the physiological position of the lungs and the adducting muscles in utterance initial position, coupled with elastic recoil forces, provide a slow rise in sub-glottal pressure (Lisker 1963). In utterance initial position, initiating phonation requires moving the glottis from a respiratory state to a phonation state, which involves varying the pressure below and above the glottis (Keating and Westbury 1986, Sawashima and Hirose 1983). The transition of the articulators from a respiratory state to a vibrating state is greater in utterance initial position than in utterance medial or final position, since during respiration both abductor and adductor muscles are inactive, and the vocal folds are in an abducted state. The transition from a non-vibrating glottal state to a vibrating state in initial stops requires that the glottal aperture get quite small before the sub-glottal pressure increases enough to exceed that of the supra-glottal pressure (Kingston and Diehl 1994). Since it is difficult to achieve the pressure difference between the sub-glottal and supra-glottal cavities needed for vocal folds to start vibrating, the pressure differences do not always exceed the threshold for initiation of voicing prior to consonant release. This can potentially result in failure to vibrate the vocal folds utterance initially. In other words, in order to initiate voicing, as opposed to maintaining voicing, a larger difference between the sub-glottal and supra-glottal pressure is required.

If the vocal folds are already vibrating, maintaining their vibration does not entail a pressure build up but rather pressure maintenance (Jessen and Ringen 2002 *among* *others*). Thus, it is more advantageous for stops to be voiceless in utterance initial environment. Conversely, intervocalically it is "deactivating" voicing that is more difficult since physiologically the articulators are already "set up" for voice production (Kingston and Diehl 1994). The sub-glottal pressure is higher and the supra-glottal pressure is lower in intervocalic position, satisfying the basic requirements for voicing. Thus, utterance medially, unless vocal folds are "deactivated", they vibrate naturally and passively and are in fact maintaining vibration.

Due to the physical conditions outlined above, utterance initial position is a disadvantageous position for voicing since the glottal area is not predisposed to vibration. In this position vocal fold vibration must be actively initiated, which involves increasing the sub-glottal pressure usually by some of the laryngeal adjustments discussed in section 4.2. But environment alone is not responsible for the difficulty in voice production in obstruents. Additional difficulties that stem from the nature of obstruents also play a role in the "unnaturalness" of voicing in obstruents.

Thus, voiced obstruents in initial position are doubly unfavorable: first because voicing is not favorable in stops, and second, because it is unfavorable in initial position. This is true of stops to a greater extent than fricatives due to the differences in how voicing is realized in these segment types, as detailed in section 4.2. This results in voiced stops often "losing" their "voicing" when they are articulated in phonetic environments where voicing conditions are not met or are less favorable. In other words, the voicing property of voiced segments, i.e. vocal fold vibration, is not realized in certain phonetic environments. Voicing will be realized consistently only in those environments where suitable conditions for voicing are met.

However, lack of realization of voicing in disadvantageous phonetic environments does not mean that the speaker did not intend to produce voicing. Phonologically voiced obstruents in utterance initial position may lack voicing because of the unfavorable conditions in this environment. However, lack of voicing in stops utterance initially should not be attributed to lack of voicing gestures or activation of muscles responsible for voicing. Rather, it should be attributed to failure to reach the required pressure threshold to initiate vocal fold vibration. The less favorable the environment, the less likely it is for the speaker to realize voicing.

As we have seen, voicing in obstruents is less favorable due to the inherent properties of stops and fricatives. That combined with utterance initial position, which in itself is a disadvantageous environment for voicing, may result in failure to produce voicing. However, the more complex the immediate phonetic environment becomes, the more likely it is for voicing to fail. In section 4.5.2 we explore the production of voicing in utterance initial obstruent clusters, which is an even more complex phonetic environment for voicing, and therefore, even more unfavorable.

4.5.2. Voicing in clusters

Given an obstruent cluster, the question is what is the optimal voicing combination word initially and how is voicing implemented in these clusters. We begin by exploring the possibility of producing two separate laryngeal gestures in an obstruent sequence. We first address the question whether it is possible to have two different laryngeal gestures in an utterance initial obstruent cluster and then turn to the preferred order of these laryngeal gestures in onset clusters. This question is crucial for our understanding of the realization of voicing in word initial clusters.

It has been shown that it is possible to have two separate glottal gestures within a cluster. Löfqvist and Yoshioka (1980a,b), and Munhall and Löfqvis (1992), found that, in Swedish and Icelandic, sequences of two voiceless obstruents could be produced with either one or two glottal gestures. What determined the number of gestures was the nature of the segments in the sequence: voiceless stop + voiceless fricative or voiceless fricative + voiceless unaspirated stop "generally contained only one glottal articulatory gesture…On the other hand, a sequence of voiceless fricative + voiceless aspirated stop usually contained two separate laryngeal gestures" (Löfqvist and Yoshioka 1980a:792) Speech rate was also found to influence the number of gestures: "At slow rates two separate laryngeal movements were observed, while at fast rates only a single laryngeal movement was produced" (Munhall and Löfqvist 1992:121-122).

In other words, [–v][–v] clusters are straightforwardly realized as such. But the question arises how the remaining types of clusters [+v][+v], [–v][+v] and [+v][–v] are realized phonetically. Although [+v][+v] clusters have the same voicing specifications, a single voicing gesture that initiates voicing may not be sufficient to maintain voicing through the entire duration of the cluster. In [+v][+v] clusters, either an additional laryngeal gesture is required for the second segment or some laryngeal maneuvers such as larynx lowering discussed in section 4.2 may be needed to maintain voicing through the entire cluster. However, in [–v][+v] and [+v][–v] clusters, two separate laryngeal gestures have to be assumed.

I hypothesize that if two glottal gestures are possible within a sequence of two voiceless obstruents, it may well be possible to produce a wider range of laryngeal gestures in obstruent clusters. If more than one laryngeal gesture is physiologically possible in [–v][–v] clusters, then more than one laryngeal gesture should be physiologically possible in all other cluster types discussed here. In other words, I argue that all voicing combinations in word initial obstruent clusters are physiologically feasible.

In section 4.6 I address the logically possible voicing configuration and my predictions regarding their phonetic realization. These predictions will be compared and verified empirically with the actual production data from Modern Hebrew to be presented in chapter 5.

4.6. What we expect: Predictions

4.6.1. [−v][−v] clusters

In this section, based on the physical, physiological and perceptual conditions outlined thus far, I outline the logically possible phonetic realizations of the various voicing combinations with various cluster types discussed throughout this work.

Phonetically and physiologically the most “natural” and easiest sequence word initially is [−v][−v] since it requires no laryngeal adjustments and the vocal folds are left open for air to flow freely from the lungs out of the vocal tract. In the case of [−v][−v] we should expect little variation in the realization of this sequence as it is the most natural and phonetically simple sequence to realize. In figure (4.1) I provide a schema which illustrates the production of [−v][−v] clusters. The horizontal axis represents the vocal fold and the vertical line provides a boundary between C1 and C2. A straight line represents non-vibrating vocal folds and a wavy line represented vibrating vocal folds:

No vibration (voicelessness): ______ vocal folds vibrating (voicing): /\/\/\/\/

Vocal fold vibration:

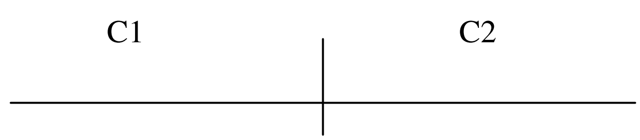


Figure 4.1: Possible realizations of voicing in [−v][−v] clusters.

4.6.2. [+v][+v] clusters

Although [+v][+v] sequences are more common cross-linguistically than either [−v][+v] and [+v][−v] sequences, they are much more complex physiologically. As detailed in section 4.2, pressure increase in the supra-glottal cavity encourages vocal fold vibration to cease before reaching the offset of C1 and the onset of C2. The result is that at the onset of the production of C2 vocal folds are not vibrating. Therefore, in order to produce a voiced C2 in a [+v][+v] cluster, after the offset of C1 where voicing is depressed, vocal fold vibration must be initiated again. Initiating vocal fold vibration twice is more complicated than having a voiceless following segment, which requires no laryngeal adjustments; a fact which makes a [+v][−v] cluster theoretically physiologically easier to realize than [+v][+v] clusters. Thus, we may conclude that in order to maintain vocal fold vibration through the entire [+v][+v] cluster, laryngeal adjustments are mandatory. Laryngeal adjustments can be in the form larynx lowering or other adjustments that will aid in maintaining vocal fold vibration, as discussed in section 4.2, or an additional muscular activity which will initiate vocal fold vibration again, after vocal fold vibration subsides due to pressure equalization. Only by either re-initiating vocal fold vibration during the production of C2 or by using any of the techniques outlined in section 4.2, such as larynx lowering or cheek puffing, voicing may be realized in C2. Without any additional adjustments, we can expect voicing to not be realized in C2, or to quickly decay during the production of C2.

Another technique that can be used to ensure voice production during both C1 and C2 is to initiate voicing late in the production of C1. That ensures that supra-glottal pressure does not rise very quickly and does not equalize with the sub-glottal pressure before the offset of the first consonant, allowing certain vocal fold vibration initiated during the production of C1 to persist into C2. I call these cases delayed voicing, since voicing is realized not at the onset of C1 but rather with a delay.

In the extreme case of delayed voicing we may see a phonetically voiceless C1 and a phonetically voiced C2 or, both C1 and C2 may surface as phonetically voiceless (a very extreme case). Since both C1 and C2 are phonologically voiced, we do not expect either C1 or C2, (certainly not both) to emerge as voiceless as there is no trigger for voicelessness. If such a case does occur, it provides strong evidence that a voiced target may fail to be initiated in word initial clusters.

Production of [+v][+v] sequence may have the phonetic realizations schematized in figure (4.2 (a-d)). The first case, (4.2a) represents perfect alignment of all laryngeal activities with the onset and offset of both C1 and C2. The second case, (4.2b), represents delayed voicing initiation where voicing initiation begins during the production of C1 but not at the onset of the segment. There is an additional possible sub-case for this voicing combination not schematized in figure (4.2), where voicing may not necessarily persist through the entire duration of C2 but may, in fact, decay mid way through the segment. The third case schematized in (4.2c) represents delayed initiation of vocal fold vibration but unlike (4.2b) where the vocal folds begin to vibrate during the production of C1, in (4.2c) vocal fold vibration is postponed until the onset of C2, note that vocal fold vibration may begin at the onset of C2 or at any other point during the production of C2. The last case, (4.2d) is an extreme case of delayed voicing where voicing begins at the onset of the vocalic portion of the syllable.

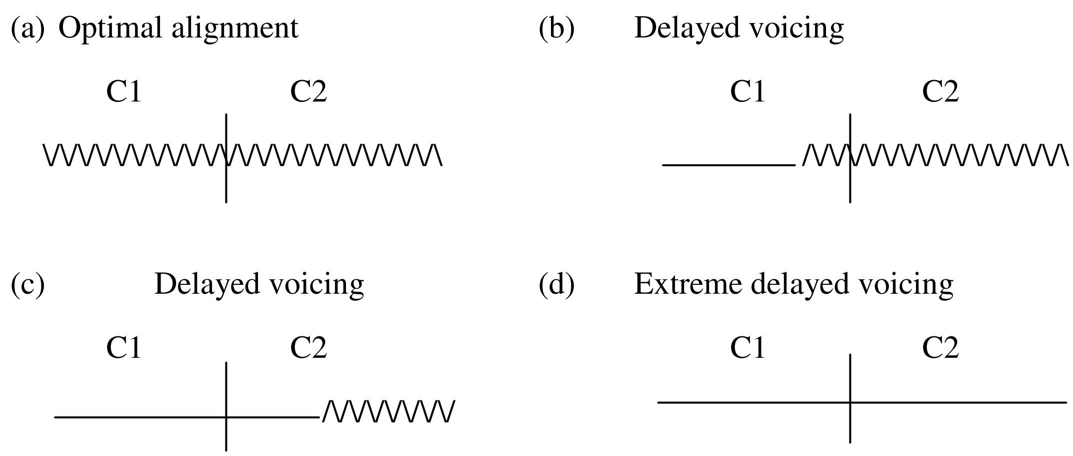


Figure 4.2: Possible realizations of voicing in [+v][+v] clusters.

4.6.3. [−v][+v] clusters

In the sequence [−v][+v] voicing increases gradually towards the nucleus of the syllable, from the voiceless C1 to the voiced C2. The first segment is voiceless while the second member of the sequence, the member which is adjacent to the vocalic nucleus of the syllable, is voiced. In [−v][+v] clusters vocal folds are open for air to flow freely during the production of the first segment but must close and begin vibrating during the production of the second segment. Sufficient pressure must be built up to initiate vocal fold vibration. Unless activation of muscular activity is perfectly timed to allow sufficient pressure build up to enable initiation of vocal fold vibration, vocal fold vibration may not be perfectly aligned with the beginning of C2 and may result in a delay of vocal fold vibration. This will result in vocal fold vibration starting mid-segment and possibly, in the extreme case, result in a voiceless second segment all together as schematized in figures (4.3b) and (4.3c) respectively. Conversely, if muscular activity begins earlier than the onset of the second segment, we may see vocal fold vibration begin during the production of C1, resulting in what is termed anticipatory voicing schematized in (4.3d). Anticipatory voicing is premature voicing which is initiated in anticipation of the voicing configuration of the following segment. Hypothetically, C1 and C2 can also both be voiced, this case is not illustrated in the scheme in figure (4.3).

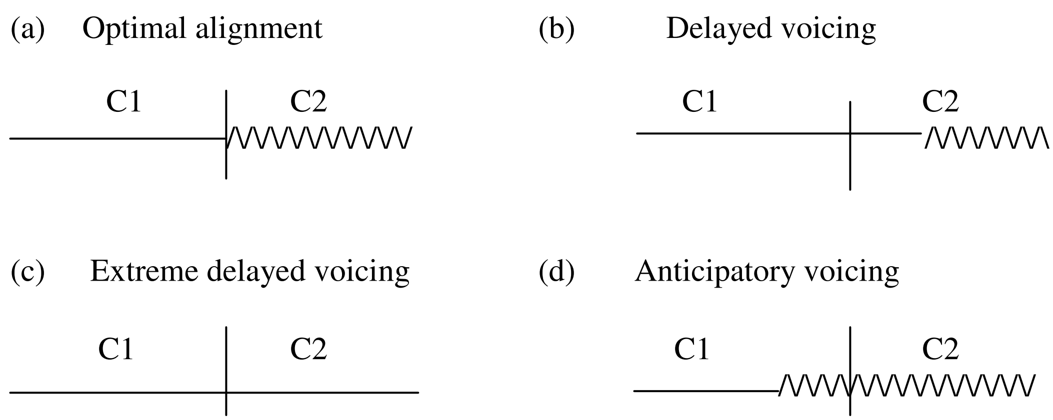


Figure 4.3: Possible realizations of voicing in [−v][+v] clusters.

4.6.4. [+v][−v] clusters

In a [+v][−v] sequence vocal fold vibration must be initiated for the first segment but then must be ceased for the second segment before resuming again for the production of the vocalic portion of the syllable. Perfect alignment of vocal fold vibration is schematized in figure (4.4a). However, cases where voicing onset and offset are misaligned with the initiation of C1 and C2 also occur. Most prominently is the case in which C1 is a stop in a [+v][−v] sequence; then we expect vocal fold vibration to be initiated prior to the burst and last through the closure portion of the stop, but vibration should cease with the release of the stop. Vocal fold vibration is initiated using muscular activation and after the initiation process of vocal fold vibrations, the vocal folds continue to vibrate independently. Initial activation of vocal folds is enough to set the vibrating motion of vocal folds and set proper conditions for independent vocal fold vibration. That is, the vocal folds will continue to vibrate as long as pressure differences are sufficient to allow them to vibrate. However, in the case of stops, the pressure continues to build up in the supra-glottal cavity as the occlusion in the vocal tract prevents air from flowing out of the vocal tract and therefore the trapped air causes a constant increase in supra-glottal pressure. The subglottal and supra-glottal pressures equalize fairly quickly, making conditions for maintaining vocal fold vibration unsuitable. Without additional muscular activity or some laryngeal adjustments such as larynx lowering, vocal fold vibration ceases even before the release of the burst. This end result leaves the larynx in an ideal position for the ensuing voiceless segment, as vocal folds cease to vibrate prior to the onset of C2. By the beginning of C2 vocal fold vibration has already ceased. Pressure equalizing between the subglottal and supraglottal cavities results in misalignments of vocal fold vibrations, that is, vocal folds stop vibrating during the production of C1 and do not vibrate during the production of C2, as schematized in (4.4b). In stops this is reflected in early termination of vocal fold vibration, considerably before the burst, and in fricatives voicing peters out before the end of the fricative. If voicing subsides before the end of the fricative, the result is a fricative which is only partially voiced at the onset and voiceless for the remainder of the fricative.

Conversely, if pressure does not drop sufficiently during the production of C1, voicing may persist into C2 and cease only during the production of C2, as schematized in (4.4c).

Another possible scenario we may encounter is one in which vocal fold vibration fails completely in C1 due to improper voicing conditions. That is, the speaker fails to build up sufficient pressure for vocal folds to initiate vibration and the pressure differences between the sub and supra glottal pressure fail to reach the threshold necessary for the vocal folds to begin vibrating. In this case vocal fold vibration fails all together and we may encounter a situation schematized in (4.4d), where vocal folds do not vibrate neither during C1 nor C2. We may conclude and say that all cases (4.4 (b-d)) represent misalignment of laryngeal activity with onset of segments. The only case of perfect alignment between segment onset and laryngeal activity is (4.4a). Lastly, the hypothetical case in which both C1 and C2 are voiced is not represented in the scheme in figure (4.4) but is logically possible.

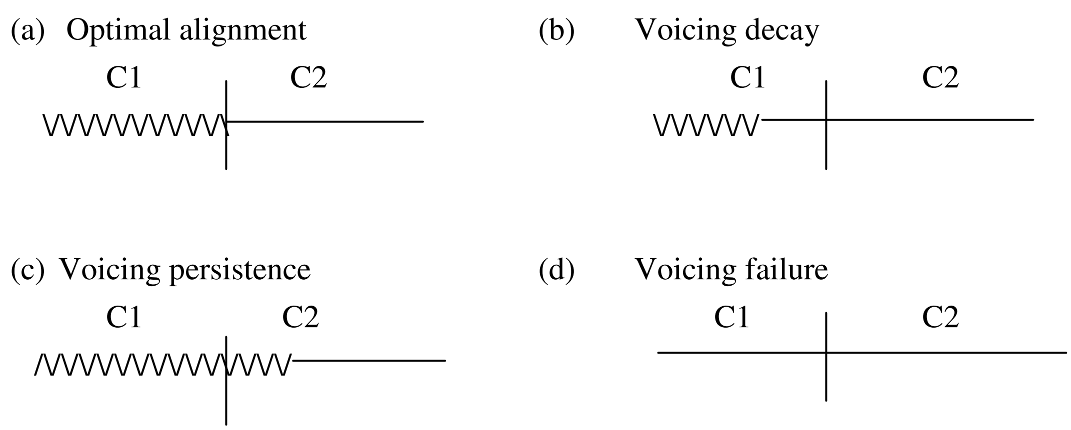


Figure 4.4: Possible realizations of voicing in [+v][–v] clusters.

4.6.5. Summary of predication

In all cases described and schematized in sections (4.6.1 – 4.6.4), none of the processes and the variations in realizations of the clusters are a result of phonological processes. All predictions are theoretically possible phonetic realizations of phonological targets based on physiological factors outlined in this chapter. Phonetic misalignment of laryngeal activity with onset and offset of segments are predicted based on factors such as word initial environment and inherent segmental properties. The misalignments are a direct result of the physical and physiological conditions required for voice productions. All predictions are theoretically predictable and will be tested empirically in chapter 5.

CHAPTER 5

PHONETIC EXPERIMENT

5.0. Introduction

The working assumption in this study is that if "devoicing" is a phonological or a categorical process then it is obligatory and should apply invariantly across the board. If, however, lack of voicing is a byproduct of phonetic difficulty (as has been argued in chapter 4), then voicing is expected to surface in some cases in word initial underlyingly voiced clusters, even if not in all of them. That is, even in voicing fails to be fully realized in some cases, a limited success rate is still expected. The realization of voicing in initial [+v] segments requires extra effort which is necessary to initiate vocal fold vibration. This is true for both [+v][+v] and [+v][−v] clusters, although [+v][−v] clusters have the additional complication of the time span available to the speaker to initiate voicing in the first member of the cluster before ceasing it for the second voiceless member. This environment is highly disadvantageous for the production of voicing for two fundamental reasons. First, initiation of voicing in utterance initial position is more difficult and less "natural" than in word medial position (Westbury and Keating 1986). Secondly, voicing must be initiated in utterance initial position right before a voiceless segment. The contradictory voicing specification of the two segments makes it physiologically difficult to produce voicing in this context. Given the difficulty in producing voicing in obstruent consonants, there is no reason to ever expect voicing in this environment. Therefore, in this specific context, we do not expect to find spontaneous or "passive" voicing. If any voicing occurs, it must be interpreted as an underlyingly voiced target otherwise voicing should never occur. In other words, unless the underlying target word initially is [+v], voicing in an utterance initial position and preceding a voiceless segment, is never expected to surface, since there is no trigger for voicing and the phonetic conditions are not conducive for its existence. “Passive voicing” can never occur in this context since the articulators are not arranged for voicing maintenance in this environment. Utterance initially the larynx must be adjusted to initiate voicing and laryngeal gestures for voicing initiation such as muscle contraction must be instigated. Therefore, without a phonological target which encourages initiation of such laryngeal gestures, voicing is never realized in this environment and is certainly not motivated phonetically. The rest of this chapter outlines a phonetic experiment conducted to further explore the phonetic nature of voicing in clusters and demonstrates the various phonetic realizations of voicing in all four clusters discussed in previous chapters: [−v][−v], [+v][+v], [−v][+v] and [+v][−v]. We will also show which of the theoretically possible phonetic realizations presented in chapter 4, section 4.6 are attested empircally.

This chapter includes both qualitative and quantitative results from the experiments and a discussion of the results follows each cluster type. A conclusion of all phonetic realizations will be presented at the end of the chapter.

5.1. Experiment Outline

5.1.1. Subjects

Six speakers were recorded, three males and three females, to ensure that meta-linguistic factors such as gender did not influence the results. Subjects were all native speakers of Modern Hebrew between the ages of 23 – 32. The subjects were mostly students at Cornell University or affiliated with the university (other than one subject recorded in Israel – subject #3). All subjects attended university and were either in the middle of their undergraduate or graduate education. All speakers spoke MH as a first language (one speaker, speaker 2, was bilingual with English his entire life). None of the subjects that participated in the study suffered from any speech impediments (one speaker, speaker 5, had a tonsillectomy, which to date has not been proven to influence voice production) nor showed any evidence of speech or hearing problems. The subjects were all naïve and none of the subjects had linguistic training of any sort.

Speaker gender breakdown is as follows: speaker 1, speaker 2 and speaker 3 are all male and speakers 4, 5 and 6 are all females.

5.1.2. Recording and analysis

Three subjects (speakers 1, 2 and 4) were recorded in a sound proof booth in the Cornell University Phonetic Laboratory using an ElectroVoice RE20 microphone and were recorded on a Panasonic SV-3200 DAT recorder. They were also simultaneously recorded into a Dell Inspiron 500m laptop computer. The other three subjects were recorded outside the laboratory (speakers 5 and 6 were recorded in their homes respectively and speaker 3 was recorded in his home in Israel). All subjects, but one, were recorded at a frequency of 44,000 Hz. at 16 bits. One subject (speaker #3), who was recorded in Israel, was recorded at a frequency of 22,050. Some background high frequency noise accompanies some of the recordings. This generally did not interfere with the analysis of the data. The data were analyzed on a Dell PC laptop computer, using PRAAT (2005) speech analysis program. Five tokens of each word were recorded for each speaker and four were labeled and analyzed.

5.1.3. Materials

The data collected consist of a list of words that the speakers were asked to read and are listed in the appendix of this chapter. Subjects were asked to read the words on the list at normal speech rate, pausing shortly between words. The reason for recording words in isolation and not in a frame sentence was to investigate the initiation of voicing in absolute initial position, as voicing initiation is physiologically distinct from voicing maintenance and it is the former and not the latter that is the focus of this work. Therefore, the target clusters were put in the context of absolute initial position.

The word list was composed of words beginning with obstruent clusters with the same voicing specifications and with contradictory voicing. Clusters with the same voicing specifications are used as a control group to compare with the clusters of varying voicing. Some of the phonemes were also recorded in isolation and not in clusters to compare realization of voicing in singleton forms and in clusters. A sample list is given in table (5.1). The word list given to speaker 3 was slightly different due to different recording conditions and timing restrictions (speaker 3 was recorded in Israel). Male speakers were recorded first. After the recordings of the male speakers were completed, the word list was extended. Therefore, female speakers generally have more tokens than the male speakers. This does not influence the overall results.

Table 5.1: Voicing combinations in word initial clusters in Modern Hebrew.

cluster type	cluster	word	cluster	word
(a) [-v][-v]	*kf*	***kf****arim* 'villages'	*tk*	***kt****amim* 'stains'
(b) [+v][+v]	*gv*	***gv****anim* 'shades'	*dg*	***dg****alim* 'flags'
(c) [-v][+v]	*kv*	***kv****arim* 'graves'	*kd*	***kd****amim* 'precedents'
(d) [+v][-v]	*gf*	***gf****anim* 'vines'	*dk*	***dk****alim* 'palms'

All words that were used in the experiments were real words. Nonsense words were not used to avoid meta-linguistic effects. Fortunately, examples exist for most cases and combinations of voicing in Modern Hebrew clusters.

Most words used were words in plural form to create minimal or close to minimal pairs for all forms. Although the words themselves are hetero-morphemic in the sense that they are plural forms with plural suffixes, no morpheme boundary is found within the target cluster. All words are bi-syllabic with stress falling on the second syllable, i.e. all target clusters are in unstressed syllables.[1] Although tokens with different vowels were recorded, results are reported only for those clusters that were followed by the vowel *a*. Tokens in which the target clusters were followed by the vowels *u* or *i* are not reported here because vowel quality seems to influence voicing realization. Forms with the vowel *a* were chosen because the vowel *a* has a minimal effect on production of voicing (Westbury 1983). However, this point is beyond the scope of this work and merits further investigation.

The word list consists not only of clusters with varying voicing but also with varying manner of articulation. For each voicing combination all possible manner of articulation combinations were also tested, to note whether manner of articulation influences realization of voicing in any way. There are four logical possibilities for obstruents (stops and fricatives) to combine as in (1):

1)

(a) Stop + Stop
(b) Stop + Fricative
(c) Fricative + Fricative
(d) Fricative + Stop

Crossing the four voicing combination possibilities with the various combinations of manner of articulation, yields sixteen possibilities for voicing and manner of articulation to combine as in (2):

[1] The vowel following the cluster, the unstressed vowel, is greatly reduced in comparison to the stressed vowel in the second syllable. It is also important to note that the singular forms of the plural forms tested, have a vowel intervening between C1 and C2.

2)

(a) [−voice][−voice]

i.	$[-\text{voice}]_{[\text{stop}]}[-\text{voice}]_{[\text{stop}]}$	=	$[-\text{v}]_{\text{stop}}[-\text{v}]_{\text{stop}}$
ii.	$[-\text{voice}]_{[\text{stop}]}[-\text{voice}]_{[\text{fricative}]}$	=	$[-\text{v}]_{\text{stop}}[-\text{v}]_{\text{fricative}}$
iii.	$[-\text{voice}]_{[\text{fricative}]}[-\text{voice}]_{[\text{stop}]}$	=	$[-\text{v}]_{\text{fricative}}[-\text{v}]_{\text{stop}}$
iv.	$[-\text{voice}]_{[\text{fricative}]}[-\text{voice}]_{[\text{fricative}]}$	=	$[-\text{v}]_{\text{fricative}}[-\text{v}]_{\text{fricative}}$

(b) [+voice][+voice]

v.	$[+\text{voice}]_{[\text{stop}]}[+\text{voice}]_{[\text{stop}]}$	=	$[+\text{v}]_{\text{stop}}[+\text{v}]_{\text{stop}}$
vi.	$[+\text{voice}]_{[\text{stop}]}[+\text{voice}]_{[\text{fricative}]}$	=	$[+\text{v}]_{\text{stop}}[+\text{v}]_{\text{fricative}}$
vii.	$[+\text{voice}]_{[\text{fricative}]}[+\text{voice}]_{[\text{stop}]}$	=	$[+\text{v}]_{\text{fricative}}[+\text{v}]_{\text{stop}}$
viii.	$[+\text{voice}]_{[\text{fricative}]}[+\text{voice}]_{[\text{fricative}]}$	=	$[+\text{v}]_{\text{fricative}}[+\text{v}]_{\text{fricative}}$

(c) [−voice][+voice]

ix.	$[-\text{voice}]_{[\text{stop}]}[+\text{voice}]_{[\text{stop}]}$	=	$[-\text{v}]_{\text{stop}}[+\text{v}]_{\text{stop}}$
x.	$[-\text{voice}]_{[\text{stop}]}[+\text{voice}]_{[\text{fricative}]}$	=	$[-\text{v}]_{\text{stop}}[+\text{v}]_{\text{fricative}}$
xi.	$[-\text{voice}]_{[\text{fricative}]}[+\text{voice}]_{[\text{stop}]}$	=	$[-\text{v}]_{\text{fricative}}[+\text{v}]_{\text{stop}}$
xii.	$[-\text{voice}]_{[\text{fricative}]}[+\text{voice}]_{[\text{fricative}]}$	=	$[-\text{v}]_{\text{fricative}}[+\text{v}]_{\text{fricative}}$

(d) [+voice][−voice]

xiii.	$[+\text{voice}]_{[\text{stop}]}[-\text{voice}]_{[\text{stop}]}$	=	$[+\text{v}]_{\text{stop}}[-\text{v}]_{\text{stop}}$
xiv.	$[+\text{voice}]_{[\text{stop}]}[-\text{voice}]_{[\text{fricative}]}$	=	$[+\text{v}]_{\text{stop}}[-\text{v}]_{\text{fricative}}$
xv.	$[+\text{voice}]_{[\text{fricative}]}[-\text{voice}]_{[\text{stop}]}$	=	$[+\text{v}]_{\text{fricative}}[-\text{v}]_{\text{stop}}$
xvi.	$[+\text{voice}]_{[\text{fricative}]}[-\text{voice}]_{[\text{fricative}]}$	=	$[+\text{v}]_{\text{fricative}}[-\text{v}]_{\text{fricative}}$

The combinations in (2) can also be summarized in table (5.2) where examples for each combination are provided:

Table 5.2: Voicing combinations combined with varying obstruents.

(a) [−v][−v]	SS – ***kt***_amim_ ‘stains’	SF – ***ks***_afim_ ‘money (pl)’	(b) [+v][+v]	SS – ***dg***_alim_ ‘flags’	SF – ***gv***_arim_ ‘men’
	FS – ***sk***_arim_ ‘surveys’	FF – ***sf***_arim_ ‘books’		FS[2] –	FF – ***zv***_axim_ ‘sacrifices’
(c) [−v][+v]	SS –***kd***_amim_ ‘precedents’	SF – ***kz***_avim_ ‘lies’	(d) [+v][−v]	SS – ***dk***_alim_ ‘palm trees’	SF – ***dx***_afim_ ‘urges’
	FS – ***sg***_anim_ ‘vice…’	FF – ***sv***_axim_ ‘bushes’		FS – ***zk***_anim_ ‘beards’	FF – ***zx***_arim_ ‘males’

[±v] = voice S= Stop F = Fricative

As can be seen in (2) and in table (5.2), sixteen logical possibilities for obstruent clusters with varying manner of articulation and with different voicing specifications within a cluster exist, and fifteen of them were tested in the experiment which will be detailed in the rest of the chapter.

The obstruents that were tested in the experiment included all voiceless stops and their voiced counterparts as listed below, and some voiceless fricatives and their voiced counterparts as detailed in (3) (some fricatives rarely appear in word initial position or as a first member of a cluster and are therefore not included in the study, as discussed in chapters 1 and 3).

[2] Although words with this cluster exist, there are no words that form minimal pairs with the words in the rest of the word list, therefore, no words with this specific initial cluster were tested.

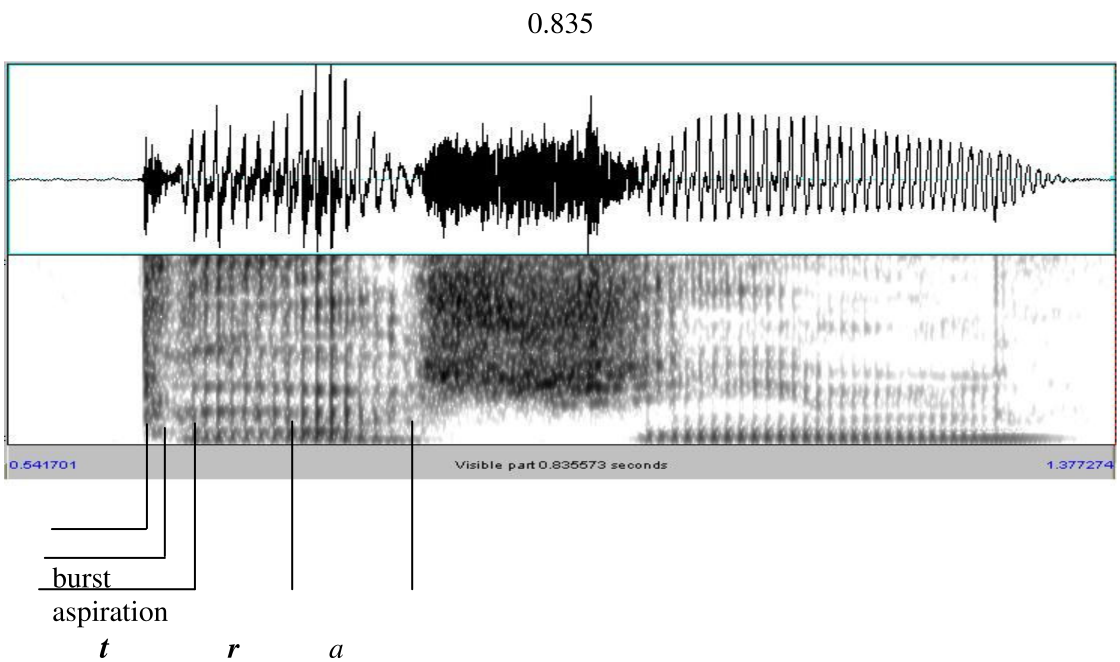
3)	[p] ~ [b],	[t] ~ [d],	[k] ~ [g]
	[f] ~ [v],	[s] ~ [z],	[x] ~ [–]

5.1.3.1. The missing voiced counterpart of [x]

Initially *r* was used as the voiced counterpart of *x*; *r* and *x* are used for convenience for what is assumed to be uvular segments: [x] being the uvular voiceless fricative and [r] being a uvular segment, as was discussed in chapter 3. There is some evidence from child language acquisition that *r* may be the voiced counterpart of *x*, based on confusion of the two phonemes. Moreover, based on evidence from Biblical Hebrew, it has been claimed that *x* is a voiceless velar fricative and *r* in MH is its voiced velar counterpart (Blanc 1964, Chayen 1972). Others claim that both *x* and *r* in MH are uvular (Berman 1997, Bolozky 1972, 1978). In other words, both the place and manner of articulation of these two segments are disputed.

While there is no debate that *x* is a fricative, the nature of *r* is not so straightforward. Bolozky (1972, 1978) claims that *r* is a uvular fricative while Berman (1997) claim that the MH *r* is a uvular trill. In either case, *r* behaves as a sonorant phonologically and therefore, cannot be used as the voiced counterpart of the uvular voiceless fricative *x* (for a more comprehensive study of *x* and *r* see Kreitman and Bolozky 2007).

For the purpose of the experiment outlined in this work, *r* was initially used as the voiced counterpart of the fricative *x* (based on the literature (Blanc 1964, Chayen 1972), however, it soon became apparent that this was not the correct treatment of *r*. From acoustic evidence gathered, *r* resembles a sonorant in most cases, and behaves as a sonorant phonologically, as can be seen in spectrogram (5.1), and therefore was treated as such for the remainder of the experiment. As can be seen in the spectrogram of the word *trashim* 'rocks' in spectroram (5.1), the slightly aspirated *t* precedes the *r*, which is clearly very “vowel like”. The number above the spectrogram denotes the duration of the visible portion of the spectrogram.



Spectrogram 5.1: Speaker – 1 – ***tr****ashim* ‘rocks’

Based on the phonetic and phonological data collected it appears that *r* behaves like a sonorant rather than a fricative. Therefore, for the purpose of this study, *r* was excluded from the set of obstruents, leaving the phoneme *x* with no voiced counterpart.

5.1.4. Measurements

Several acoustic correlates associated with voicing have been measured as well as additional acoustic attributes that are not known to be correlated with voicing at this point, to test whether they are correlated with voicing or not. The following were measured: burst duration, release duration (which includes the aspiration or the voiced release portion), duration of voicing,[3] duration of fricative, duration of vowel following the cluster and pitch (*f0*) at onset of vowel,[4] for each segment. Amplitude was also measured but could not be used as a cue for distinguishing voicing in C1 and therefore was not included. All data were analyzed using a Hanning window with a 5ms window size. The data are presented in tables, a separate table for each consonant, C1 and C2.

In spectrogram (5.2) is a sample spectrogram with segmenting and labeling marked to illustrate how spectrograms were labeled. Points were used for labeling and the measurements were obtained between various points. The number following the point indicates whether the measurements pertain to C1 or to C2; vb1 is beginning of voicing in C1 and vb2 is beginning of voicing in C2. Points and measurements were obtained as listed in (4):

4)

Burst duration was measured between bb = burst begin and be = burst end,

Duration of **aspiration** was measured between rc = release continue and re = release end,

Closure duration was measured between cb = closure begin and ce = closure end,

Closure voicing was measured between vb = voicing begin and ve = voicing end,

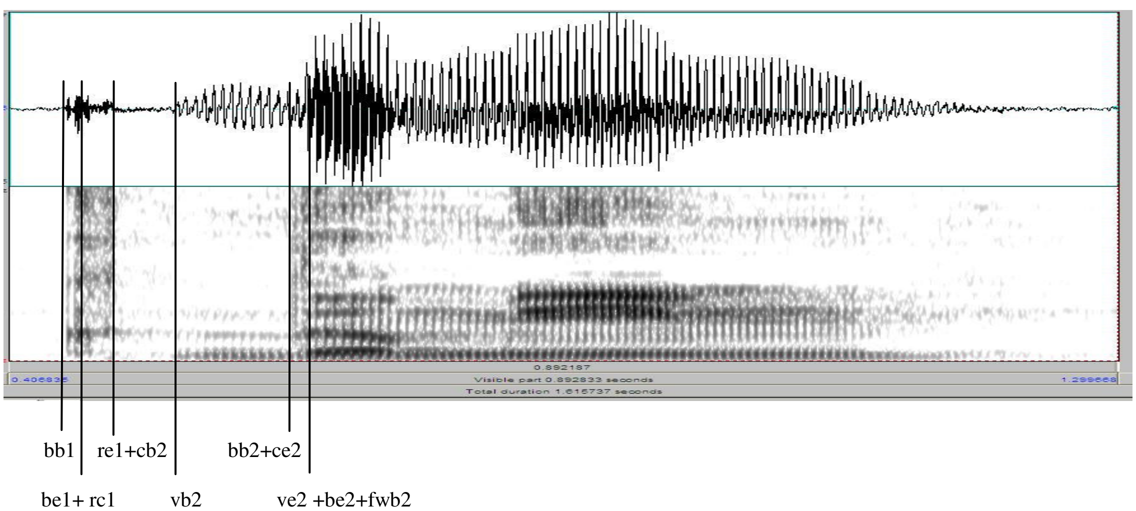
Fundamental frequency was measured at the beginning of the vowel following the cluster, at the point marked as fwb = following vowel begin. Fundamental frequency (*f*0) was retrieved using the PRAAT default settings.

---

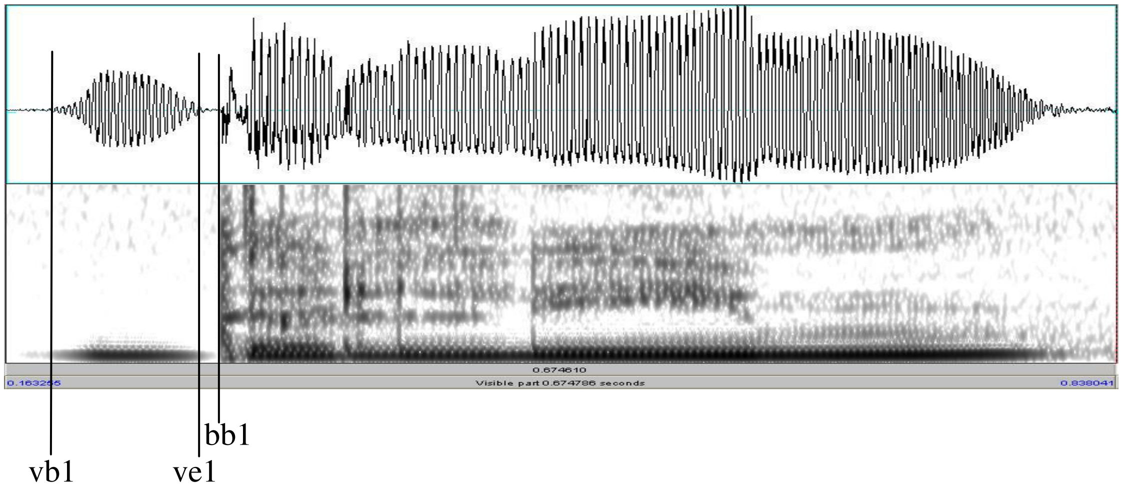
[3] For stop members voicing duration was measured as closure voicing, for fricatives this was measured as portion of the segment that was voiced. In absolute initial position closure voicing was measured from initiation of vocal fold vibration (the beginning of the voice bar).

[4] First formant in the vowel following C2 of the cluster was also measured but did not prove to be a reliable cue for voicing in C1 and was therefore not included in the results.

It is important to mention that when voicing petered out prior to the burst, it was marked as such on the spectrogram and the measurement of duration of closure voicing was taken only for the voiced portion, as demonstrated in spectrogram (5.3). This method is slightly different from the measuring method employed by Lisker and Abramson (1964), who measured negative VOT from the beginning of vocal fold vibration to the onset of the burst. They do not address cases where voicing peters out prior to the burst and contains a portion of silence prior to the burst as in spectrogram (5.2). Therefore, in this work, voicing which precedes the burst is referred to as closure voicing and not negative VOT. Only the voiced portion of the closure is measured and duration of voicing refers only to the voiced portion of the closure, not including the portion without vocal fold vibration immediately preceding the burst.



Spectrogram 5.2: Sample labeled spectrogram $[-v]_{stop}[+v]_{stop}$ *kdamim* ‘precedents’.



Spectrogram 5.3: Voicing diminishes before burst (word *dalim* ‘poor’).

5.1.5. Results

Each of the sixteen possible clusters will be reported on separately. First, we begin by indicating the expected results for voicing. In figure (5.1) I list the expected voicing realizations for each cluster in an optimal case of perfect alignment between onset of voicing and segment onset. Vibrating vocal folds are depicted by a curvy line and spread glottis (non vibrating vocal folds) is indicated by a smooth line.

a. [−v][−v]
b. [+v][+v]
c. [−v][+v]
d. [+v][−v]

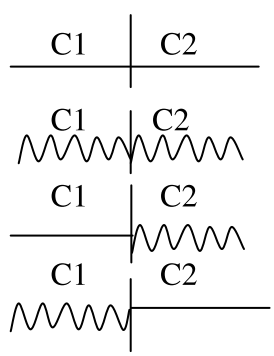


Figure 5.1: Optimal voicing configuration in initial clusters.

If voicing is perfectly aligned then when two voiceless segments follow each other there should be no voicing in either segment, as illustrated by figure (5.1a). When two voiced segments follow each other we expect voicing through both segments, as in figure (5.1b). If a voiceless segment is followed by a voiced segment then we expect the first segment to be voiceless (i.e. no vibration of the vocal folds) and the second segment to have voicing as in figure (5.1c). When a voiced segment is followed by a voiceless segment and if these clusters are phonetically realized as [+v][−v], then in the optimal case, one in which voicing is perfectly aligned, we expect voicing in the first segment but not in the second. However, since physiologically some clusters are more difficult to produce than others, as detailed in chapter 4, we may predict some misalignment for each cluster, as was discussed in chapter 4. In this chapter we will explore which theoretically phonetically predictable realizations outlined in chapter 4 section 4.6 are actually empirically attested.

The rest of the chapter will be organized as follows: I begin by discussing voicing in singletons before exploring the four possible voicing combinations listed in (2a-d) and illustrated in figure (5.1 (a-d)), which will be presented in separate sections. Then each of the possible manners of articulation combinations listed in (2 i-xvi) and in table (5.1) will be presented in a subsection illustrating all sixteen possible cluster combinations. Spectrograms will be presented for the various phonetic realizations of each possible cluster followed by quantitative data for each subsection. Then a discussion of the results will follow each section before a final conclusion and discussion of all the results.

5.1.6. Voicing in singletons

Modern Hebrew distinguishes between voiced and voiceless obstruents. Generally, the [+voice] feature is realized as closure voicing in most environments. In utterance initial position the duration of the closure voicing depends on place of articulation but some closure voicing is usually present in all voiced stop consonants. Laufer (1994, 1995, 1998) found the following VOT values for voiced stop consonants (although it is not clear whether any of these cases had voicing that tapered towards the burst or whether there were any periods of silence preceding the burst). The measurements are taken out of Laufer (1998) noting that he measured closure voicing preceding the burst using Lisker and Abramson's proposed measurement system: negative VOT – preceding the burst, positive VOT – following the burst. The average durations of VOT, presented in table (5.3), were as follows: $b = -95$ msc, $d = -102$ msc, $g = -81$ msc. As anticipated the velar stop has the shortest VOT but nonetheless all voiced stops have negative VOTs. For voiceless stops he found the generally anticipated short lag VOT: $p = 20$ msc, $t = 26$ msc, $k = 51$ msc. As can be seen from Laufer's results, the velar voiceless stop has a considerably longer VOT (bordering on long lag VOT).

Table 5.3: VOT values for stop consonants in MH (Laufer 1998)

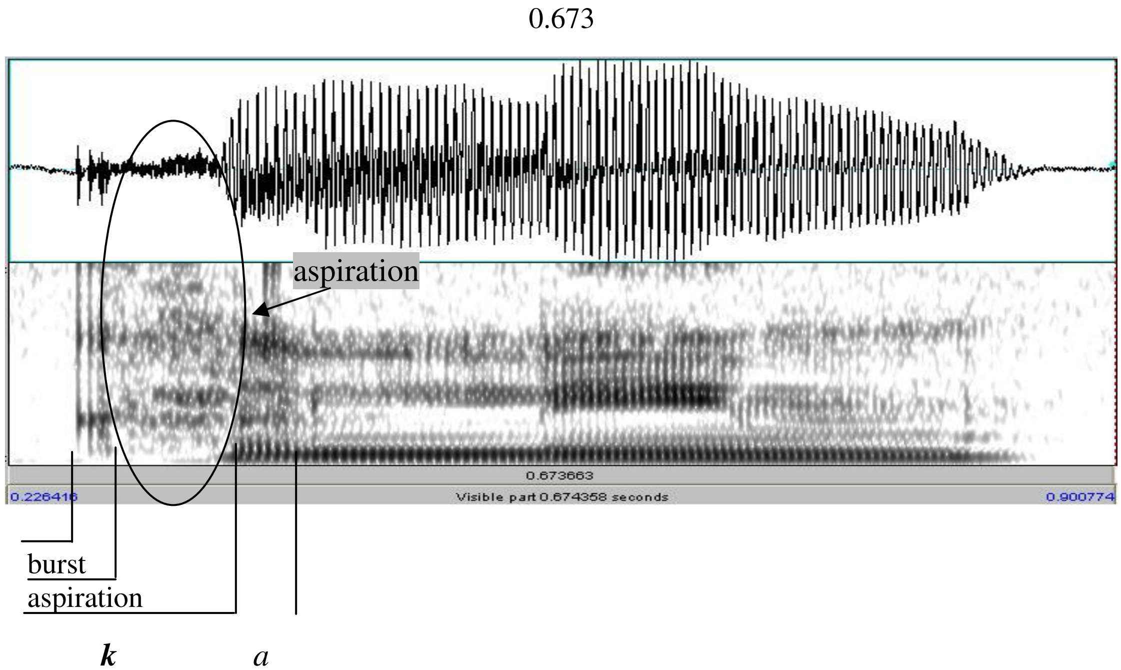
	b	**d**	**g**	**p**	**t**	**k**
VOT value	– 95 ms	– 102 ms	– 81 ms	+ 20 ms	+ 26 ms	+51 ms

Laufer found that Hebrew voicing is similar to that of Polish in the duration he calculated for closure voicing. However, all data collected by Laufer are for simplex onsets, there are no data available to date on the values of neither closure voicing nor positive VOTs in complex onsets. Aspiration is not categorically distinctive in Modern

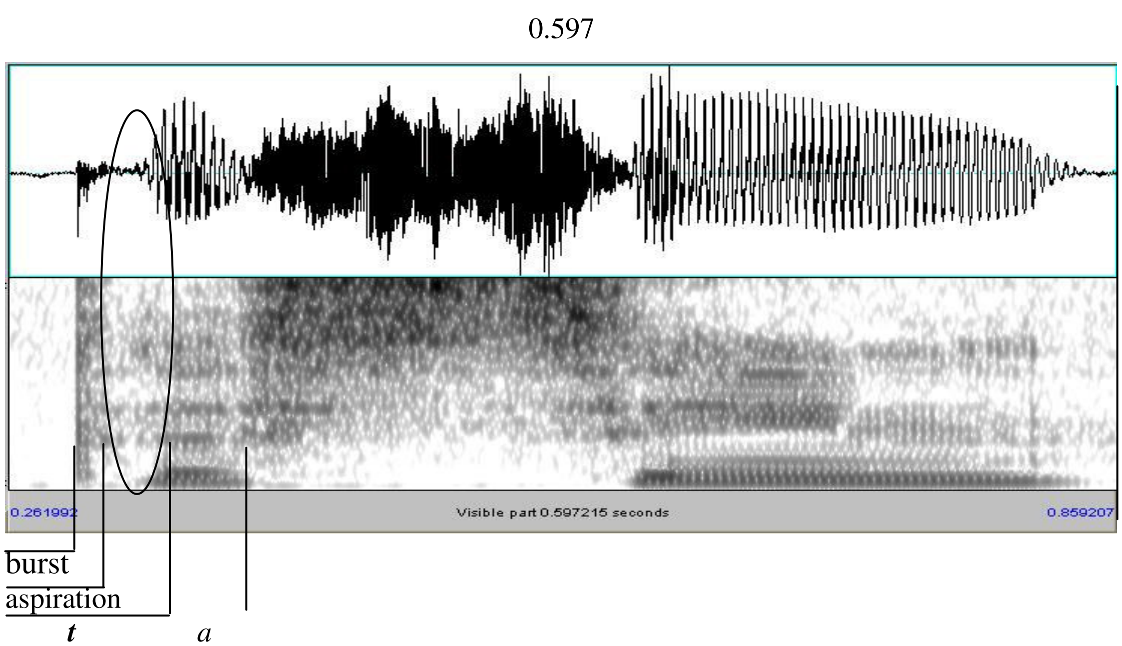
Hebrew, and therefore both short and long lag VOT's are associated with voiceless segments, while closure voicing is an indicator of voicing.

It is important to note that while these measurements provide support for the claim that voicing is realized as vocal fold vibration prior to the burst, the measurements which will be presented in the remainder of the chapter are based on the measuring method detailed in section 5.1.4. That is, measurements for closure voicing will be provided but the measured portion is only the portion of closure which actually contains vocal fold vibration, and does not include the portion preceding the burst where vocal fold vibration tapers off, if such exists. In this method I depart from Lisker and Abramson (1964) and Laufer (1998). The portion of silence is excluded from the measurements in this work, which result in a smaller portion of the closure being voiced, as will become apparent later in this chapter.

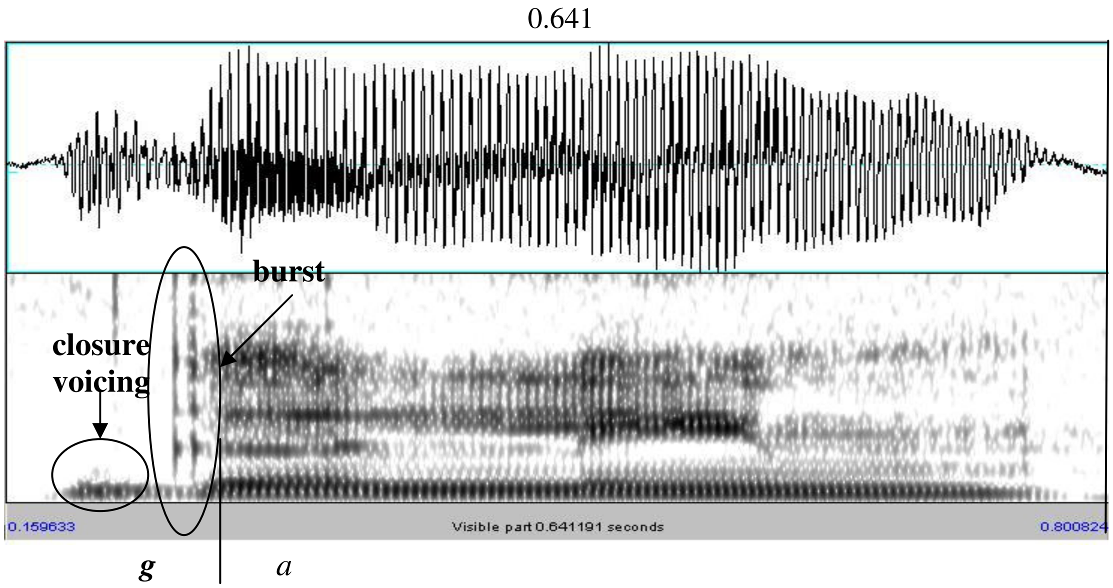
Spectrograms (5.4) and (5.5) are spectrograms of the voiceless segments *k* and *t* respectively and their voiced counterparts *g* and *d* are presented in spectrograms (5.6) and (5.7). The spectrograms corroborate Laufer's findings that velar stops have a greater aspirated portion than alveolar stops. Thus, it appears that place of articulation plays a significant role in the amount of aspiration that accompanies the stop. This is an important observation, since, as will become apparent through the rest of the chapter, aspiration varies quite greatly from speaker to speaker and there is great variation within individual speaker's data as well. That is, aspiration is entirely not categorical and can be realized in many different ways, either as short lag or as a long lag VOT. Additional factors such as place of articulation and rate of speech determine the amount of aspiration that is realized in a voiceless segment.



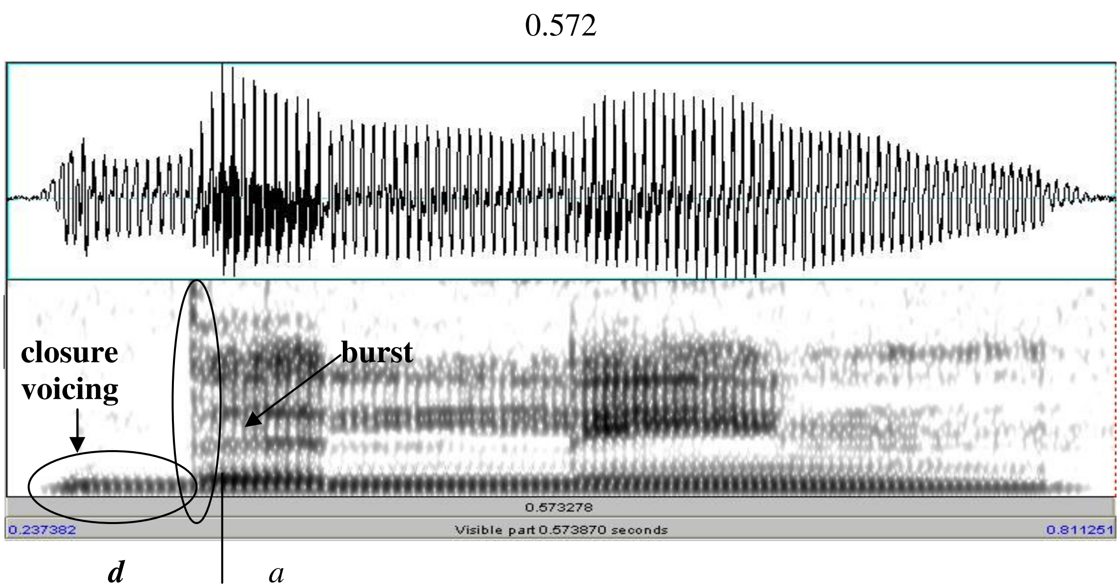
Spectrogram 5.4: Speaker – 4 – ***k****alim* ‘light (pl.)



Spectrogram 5.5: Speaker – 4 – ***t****asim* ‘flying (pres. pl.)’



Spectrogram 5.6: Speaker – 4 – ***g****alim* 'waves'



Spectrogram 5.7: Speaker – 4 – ***d****alim* 'poor (pl.)'

In both spectrograms, (5.7) and (5.8), closure voicing precedes the burst and it is the presence of this voiced portion prior to the burst which distinguishes voiced from voiceless segments. The duration of the voicing preceding the burst is not crucial, rather it is the presence or absence of the vibration of vocal folds that determines whether a segment is perceived as voiced or not. In both examples, there is no aspiration after the burst of the voiced segments but, as will be seen in this chapter, sometimes voiced segments are accompanied by a small portion of aspiration, especially in clusters.

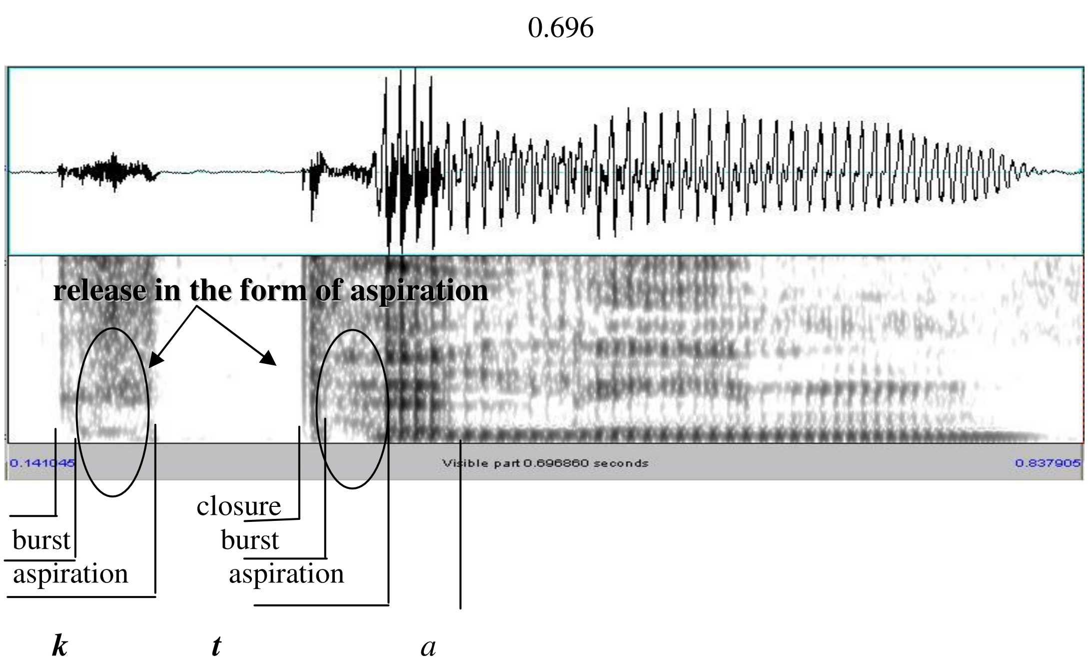
Voiced fricatives, which include only *v* and *z* in Modern Hebrew, as there is no velar voiced fricative, are usually realized with voicing. It has been argued that *v* is a sonorant in Hebrew (Barkai and Horvath 1978), as in other languages such as Hungarian where it behaves as a sonorant, based on the fact that it does not trigger voicing assimilation. However, it is undoubtedly influenced by voicing and participates in processes related to voicing such as voicing assimilation as in the word /savta/ → [safta] 'grandmother'. Since evidence in either direction is inconclusive and this issue is outside the scope of this work, I treat *v* as an obstruent fricative.

Now that we have detailed the method of measuring and established realization of voicing in singletons in Modern Hebrew, we will continue to test the predictions made in chapter 4 section 4.6 and explore voicing in clusters in Modern Hebrew. In each section I present the various voicing combinations and in each sub-section the varying manner combinations are addressed. Section 5.2 is dedicated to [–v][–v] clusters, section 5.3 to [+v][+v] clusters, section 5.4 to [–v][+v] clusters and finally section 5.5 will explore [+v][–v] clusters.

5.2. [−v][−v] clusters

5.2.1. (i) $[-v]_{stop}[-v]_{stop}$

In spectrogram (5.8) we see an example of two voiceless stops following one another. Evident from the spectrogram is the fact that both voiceless stops include a portion of aspiration. As expected, neither segment contains a voiced portion. This example represents the phonetic realization of all $[-v]_{stop}[-v]_{stop}$ clusters. None of the tokens contain any voiced portion and most tokens contain a certain portion of aspiration, although the duration of aspiration varies between short lag and long lag aspiration, as is expected considering the non-distinctiveness of aspiration. Measurements and quantitative data follow the spectrogram in section 5.2.1.1.



Spectrogram 5.8: Speaker – 1 – ***kt**alim* 'walls'

5.2.1.1. Quantitative results for $[-v]_{stop}[-v]_{stop}$

In this section I present measurements for $[-v]_{stop}[-v]_{stop}$ clusters. Data for each member of the cluster will be presented in a separate table. Table (5.4) presents data for C1. The data presented includes amount of aspiration and burst duration for each stop consonant separately. The column marked **asp**. denotes the amount of aspiration and **burst dur**. stands for burst duration.

Units of measurement are seconds with up to three digits after the decimal point, 0.100 is 100 msec. 0.010 are 10 msec. and 0.001 is 1 msec. The measurements were rounded up after the third digit (for example 0.0785 is 0.079 and 0.0783 is 0.078). Standard deviations are provided in parentheses after each measurement. Extreme outliers were excluded from the final measurements. If data were not available for various reasons n/a appears in the appropriate rubric as is the case for speaker 3, for whom there was no data for *p* due to a technical error.

Table 5.4: Voiceless stop as first member of a cluster.

C1	**p**		**t**		**k**	
	asp.	**burst dur.**	**asp.**	**burst dur.**	**asp.**	**burst dur.**
s. 1	0.030 (0.015)	0.005 (0.002)	0.033 (0.007)	0.009 (0.002)	0.033 (0.014)	0.021 (0.005)
s. 2	0.018 (0.004)	0.015 (0.004)	0.025 (0.010)	0.005 (0.002)	0.021 (0.004)	0.017 (0.004)
s. 3	n/a	n/a	0.025 (0.010)	0.010 (0.005)	0.024 (0.008)	0.013 (0.010)
s. 4	0.034 (0.007)	0.007 (0.006)	0.029 (0.006)	0.006 (0.001)	0.033 (0.012)	0.014 (0.005)
s. 5	0.044 (0.021)	0.009 (0.002)	0.029 (0.005)	0.009 (0.002)	0.030 (0.008)	0.012 (0.006)
s. 6	0.010 (0.003)	0.017 (0.011)	0.030 (0.006)	0.010 (0.004)	0.041 (0.004)	0.009 (0.005)

As can be seen in table (5.4), no voicing was evident in any of the segments, as expected for voiceless segments in word initial position, where there is no trigger for voicing. Moreover, a very large standard deviation is evident for both aspiration and burst duration. These vary not only across speakers but vary greatly within speakers. For most aspiration data presented in the table, a standard deviation of 20-50% is quite common, which is not unexpected considering aspiration is not categorical and does not contribute to categorical distinction. It has been mentioned that both short and long lag VOT mark voicelessness in stops in MH and the greatly varied standard deviations reflect this situation. Some stops exhibit short lag VOT and some stops have long lag VOT, a situation which results in greater standard deviations for aspiration measurements.

The inconsistencies in the duration of the bursts suggest that burst durations are also not distinctive. Bursts vary according to the place of articulation and rate of speech and it appears that when speakers use more careful speech bursts tend to be longer. Thus, they can not serve as a reliable cue for distinguishing voicing.

Next we present data for C2, which are presented in table (5.5) and include results for: duration of the closure for the second segment, which is essentially the closure between the two stops is presented, in a column entitled **closure dur.**, additional measurements include: aspiration, burst duration of the second stop, and results for *f0* at the onset of the vowel following the cluster. The fundamental frequency was not always tracked accurately. Where measurements were suspicious or absent, they were verified and where necessary, and possible, also obtained manually. Only clusters with a *t* or a *k* in C2 position were tested, therefore there is no data available for *p* in C2 position in $[-v]_{stop}[-v]_{stop}$ clusters.

Table 5.5: Voiceless stops as second member of a cluster.

C2	**t**			
	closure dur.	**asp.**	**burst dur.**	***f*0**
s. 1	0.102 (0.006)	0.036 (0.010)	0.012 (0.003)	133 (13.6)
s. 2	0.070 (0.003)	0.03 (0.006)	0.010 (0.003)	131
s. 3	0.079 (0.008)	0.017 (0.005)	0.008 (0.002)	151 (1.5)
s. 4	0.076 (0.007)	0.025 (0.004)	0.013 (0.004)	220 (8)
s. 5	0.098 (0.014)	0.020 (0.008)	0.009 (0.004)	256 (9)
s. 6	0.076 (0.012)	0.023 (0.006)	0.007 (0.005)	244 (3)
C2	**k**			
	closure dur.	**asp.**	**burst dur.**	***f*0**
s. 1	0.052 (0.008)	0.054 (0.013)	0.013 (0.006)	131 (6.4)
s. 2	0.051 (0.013)	0.03 (0.008)	0.02 (0.005)	143 (9)
s. 3	0.076 (0.004)	0.054 (0.015)	0.024 (0.010)	151 (4)
s. 4	0.061 (0.012)	0.043 (0.010)	0.020 (0.005)	225 (9)
s. 5	0.084 (0.006)	0.050 (0.010)	0.014 (0.005)	252 (21)
s. 6	0.075 (0.011)	0.047 (0.013)	0.014 (0.005)	238 (10)

Initially measurements for first formant, F1, were also included but they were later omitted for several reasons. First, the measurements were too imprecise and the first formant tracker was very unreliable in tracking the first formant. Since the measurements were obtained right after the second obstruent, the formant tracker was often unable to track measurements reliably due to the proximity to the obstruent. Secondly, after a deeper inspection of the results, F1, which was measured at the onset of the vowel following C2, did not seem to be influenced by the voicing specifications of the first segment, C1, and did not prove to be a cue for distinguishing voicing in C1. Since voicing in C1 is of crucial importance, especially for arguments made regarding [+v][–v] clusters, a reliable cue for voicing is necessary. Therefore, F1 measurements were not included in this work.

The great standard deviation is again evident for both aspiration and burst duration. When compared with the voiceless stops in C1 position, we can see that speakers vary on whether both aspiration and burst duration are shorter or longer in C2 than in C1, suggesting, once again, that these are not categorical and therefore speakers can afford to be sloppy in their realization.

Closure duration, however, exhibits a smaller standard deviation, not exceeding 10%, suggesting that it has a role in possibly cueing voiced and voiceless segments. Closure duration is a more reliable secondary cue for voicing. Note that closure duration also depends on place of articulation and is longer for alveolar stops than for velar stops. As will be seen later, closure duration is shorter for voiced segments.

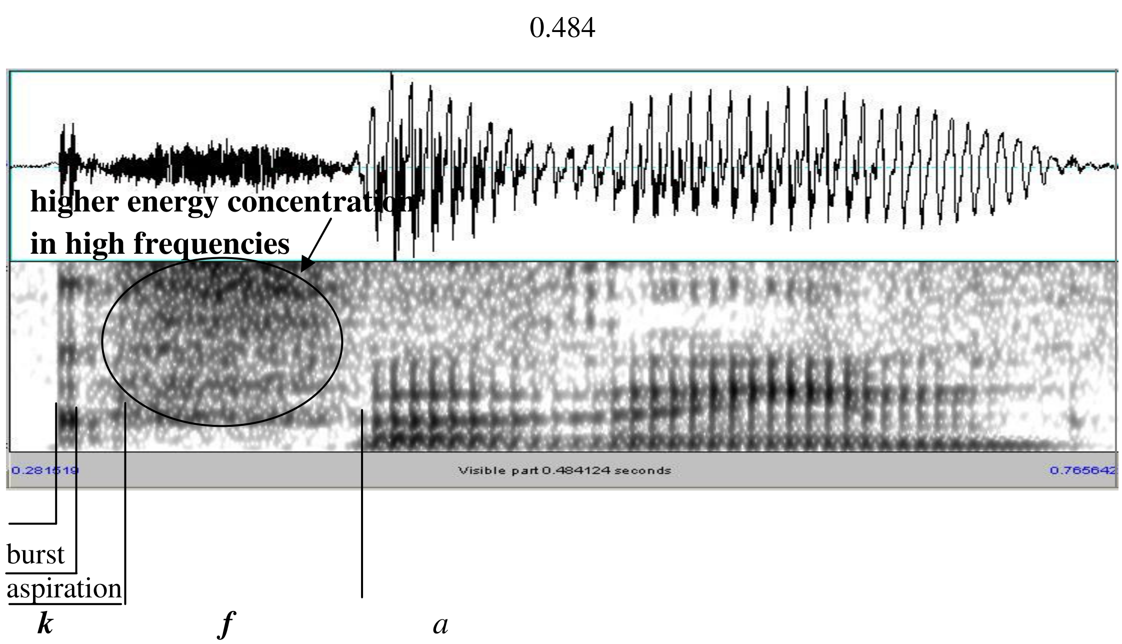
The fundamental frequency, ***f*0**, shows very little standard deviation and once again, as will be seen later, also contributes to cueing the distinction between voiced and voiceless segments, at least in C2.

We may summarize both tables and say that voicing never occurs during the production of two voiceless stops, as anticipated. Less expected is the amount of aspiration during the production of both consonants in the clusters, since it contradicts what has been found for English. Browman and Goldstein (1986) found that in English, a voiceless stop in a C2 position of a word initial cluster is less aspirated than a voiceless stop by itself. In our case, *k*, for example, in C2 position is considerably more aspirated than *k* in C1 position and when compared with Laufer's findings in table (5.3), *k* in C2 position seems to be as aspirated as *k* in singletons. However, both aspiration and burst duration depend on place of articulation and rate of speech and since they are not categorical, they result in large standard deviations and large variation in their realization. Speakers do not rely on duration of burst or the duration of aspiration as cues for categorical distinction between voiced and voiceless segments in Modern Hebrew.

5.2.2. (ii) $[-\text{v}]_{\text{stop}}[-\text{v}]_{\text{fricative}}$

We now turn to clusters where the first stop member is followed by a fricative, as exemplified in the spectrogram in spectrogram (5.9).

Evident in the spectrogram is the short duration of aspiration in *k* which is considerably shorter than in those cases where *k* is followed by a stop or even a vowel, as was previously seen in spectrograms (5.4) and (5.8). In this token the fricative is separated from the preceding stop by audition as well as stronger energy visible in high frequencies in the fricative but not in the aspiration portion of the stop. As expected, no voicing is evident in either C1 or C2, as will be transparent in the quantitative results which will be presented in the following section.



Spectrogram 5.9: Speaker – 2 – ***kf****arim* – 'villages'

5.2.2.1. Quantitative results $[-\text{v}]_{\text{stop}}[-\text{v}]_{\text{fricative}}$

Data for the first member of a $[-\text{v}]_{\text{stop}}[-\text{v}]_{\text{fricative}}$ cluster is presented in table (5.6). As mentioned in the previous section 5.2.2, and evident from table (5.6), aspiration is shorter when a stop C1 precedes a fricative C2 than when a stop C1 precedes a stop C2. There is no significant difference in the duration of the burst between $[-v]_{stop}[-v]_{fricative}$ clusters and $[-v]_{stop}[-v]_{stop}$ clusters although for some tokens burst duration is slightly longer while in others it is slightly shorter, which is consistent with the assumption that burst duration is not a reliable cue for voicing distinctions.

Table 5.6: Voiceless stops first member of a cluster.

C1	t		k	
	asp.	**burst dur.**	**asp.**	**burst dur.**
s. 1	0.020 (0.011)	0.008 (0.001)	0.013 (0.003)	0.017 (0.007)
s. 2	0.026 (0.007)	0.011 (0.007)	0.014 (0.009)	0.011 (0.007)
s. 3	0.015 (0.005)	0.014 (0.009)	0.02 (0.009)	0.014 (0.006)
s. 4	0.012 (0.002)	0.008 (0.004)	0.014 (0.005)	0.015 (0.007)
s. 5	0.021 (0.014)	0.007 (0.003)	0.020 (0.007)	0.014 (0.004)
s. 6	0.011 (0.007)	0.014 (0.007)	0.020 (0.009)	0.008 (0.005)

Next we turn to table (5.7) where data are presented for the second fricative member of a $[-v]_{stop}[-v]_{fricative}$ cluster. The duration of the fricative segment is provided in the column **fric. dur.** and the fundamental frequency is presented in the column ***f*0**. For speakers 1 and 3 no standard deviation is provided for ***f*0** for *f* and *s* because measurements were obtained for only one token. The standard deviation is very small for both fricative duration and fundamental frequency. The fricative *f* is the longest of all fricatives with *s* following and *x* being the shortest fricative. The small standard deviation suggests that fricative duration contributes to the perception of voicing distinction. This will become evident later in the chapter where we will see that voiced fricatives are shorter than voiceless ones.

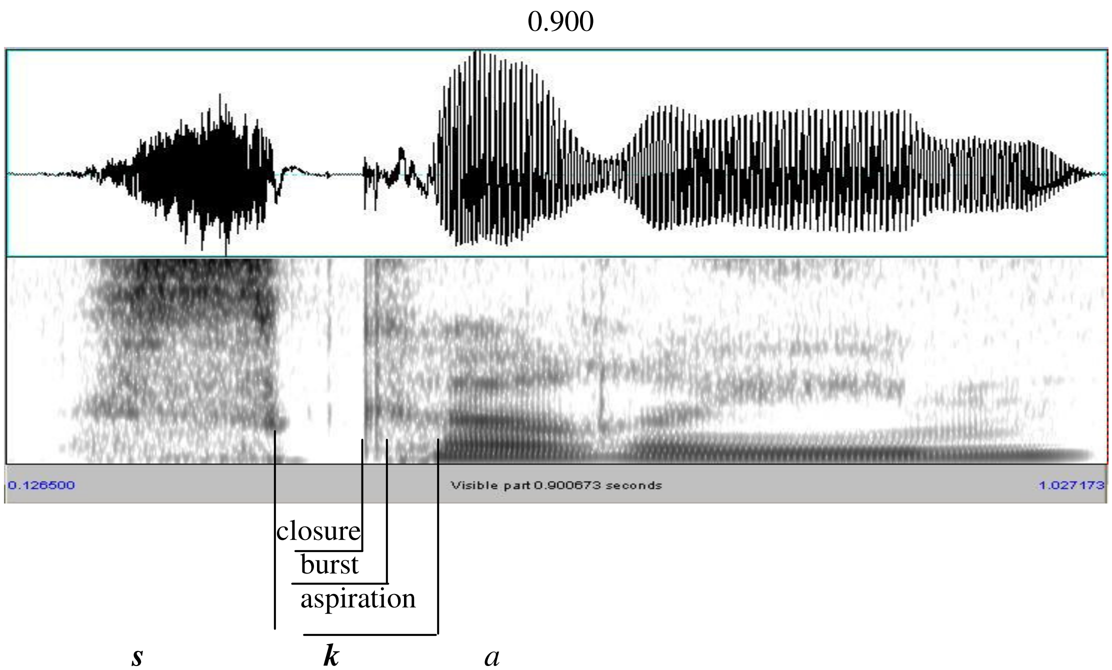
Once again the results are quite straight-forward and not surprising. There is no voicing evident in either one of the segments. Of interest here is the amount of aspiration in a stop preceding a fricative, which is much smaller than the amount of aspiration in a stop preceding another stop. This may be an effect of the proximity of the aspirated portion of the stop to the following fricative. In order to avoid confusion between the frication noise of the fricative and the aspirated portion of the stop, the aspirated portion is kept to a minimum. Moreover, it has been claimed all along that aspiration is not distinctive and therefore its presence is not crucial. Speakers can afford to be sloppy with its realization with little influence on the actual perception of the segment C1 as a voiceless segment.

Table 5.7: Voiceless fricatives as second member of a cluster.

C2	**f**	
	fric. Dur.	***f*0**
s. 1	0.164 (0.013)	147
s. 2	0.130 (0.014)	139 (3)
s. 3	0.149 (0.022)	147.5 (0.7)
s. 4	0.174 (0.022)	245 (7)
s. 5	0.174 (0.026)	289 (10)
s. 6	0.151 (0.029)	242 (10)
C2	**s**	
	fric. Dur.	***f*0**
s. 1	0.140 (0.017)	*
s. 2	0.121 (0.015)	142 (0.54)
s. 3	n/a	n/a
s. 4	0.137 (0.008)	235 (20)
s. 5	0.143 (0.007)	242 (21)
s. 6	0.145 (0.016)	246 (15)
C2	**x**	
	fric. Dur.	***f*0**
s. 1	0.120 (0.12)	127 (6)
s. 2	0.100 (0.007)	133 (7)
s. 3	0.128 (0.024)	148
s. 4	0.149 (0.007)	223 (14)
s. 5	0.133 (0.009)	257 (15)
s. 6	0.141 (0.008)	237 (10)

5.2.3. (iii) $[-\mathrm{v}]_{\mathrm{fricative}}[-\mathrm{v}]_{\mathrm{stop}}$

A voiceless fricative followed by a voiceless stop is a highly restricted cluster in MH. For historical reasons fricatives are rare in word initial position. Alveolar fricatives, *s* and *z*, however, are less restricted as first members of a cluster, as they occur frequently in borrowed words (see chapters 1 and 3 for more detailed discussion). Therefore, the only example of a voiceless fricative as a first member of a cluster and preceding a stop is that of *s*.[5] Quantitative data is presented in the following section.



Spectrogram 5.10: Speaker – 6 – ***sk****arim* ‘surveys’

5.2.3.1. Quantitative results for $[-\mathrm{v}]_{\mathrm{fricative}}[-\mathrm{v}]_{\mathrm{stop}}$

Table (5.8) lists data for the first member of the cluster, *s*. The only measurement of interest is the duration of the fricative, which is longer when *s* in C1 position is longer than in C2 position (as evident from table (5.7) in section 5.2.2.1). Also note the small

[5] Excluding forms like *xtiv* ‘writing’ and *ftax* ‘open (imp.)’ which are still marginal.

standard deviations, constistant with the claims that duration of a fricative is an important cue for voicing, a claim which will become increasingly apparent when we visit the production of voiced fricatives.

Table 5.8: Voiceless fricative as a first member of a cluster.

C1	s
	fric. dur.
s. 1	0.164 (0.012)
s. 2	0.162 (0.020)
s. 3	0.166 (0.024)
s. 4	0.166 (0.034)
s. 5	0.183 (0.020)
s. 6	0.154 (0.011)

In table (5.9) are measurements for *k* in C2 position. Once again, in table (5.9), aspiration and duration show a great standard deviation ranging from 10-50% and once again showing that these two cues are not reliable for distinguishing voicing.

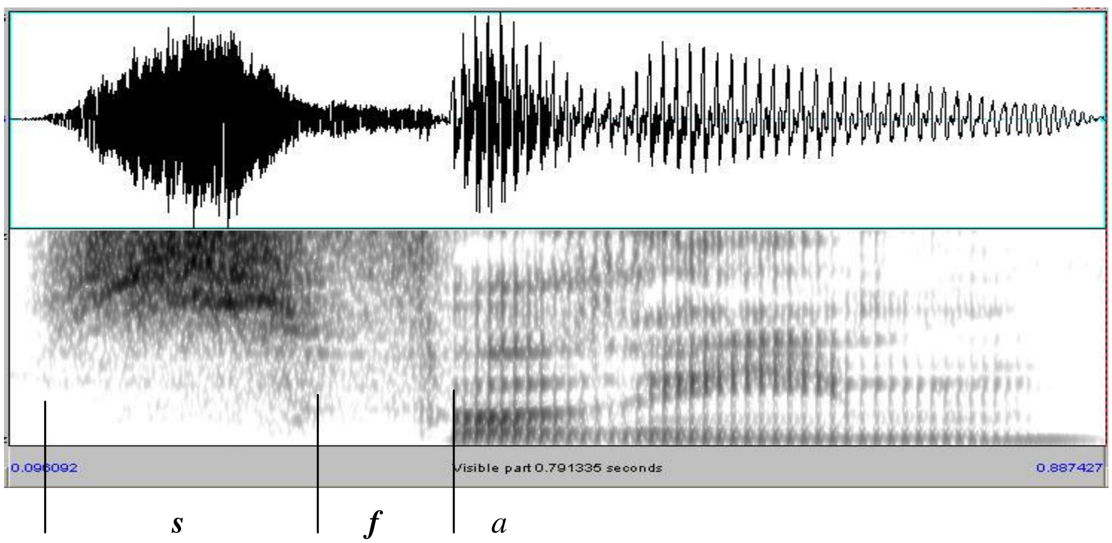
Table 5.9: Voiceless stop as a second member of a cluster.

C2	k				
	closure dur.	**asp.**	**Burst dur.**	***f*0**	**F1**
s. 1	0.070 (0.007)	0.40 (0.009)	0.015 (0.004)	135 (5.98)	496 (246)
s. 2	0.053 (0.011)	0.023 (0.005)	0.017 (0.006)	137 (7.7)	*
s. 3	0.057 (0.006)	0.026 (0.011)	0.027 (0.004)	157 (9.2)	445 (16)
s. 4	0.069 (0.005)	0.042 (0.016)	0.024 (0.014)	210 (9.28)	366 (252)
s. 5	0.077 (0.007)	0.058 (0.019)	0.016 (0.007)	259 (16.5)	474 (97)
s. 6	0.077 (0.009)	0.040 (0.004)	0.012 (0.006)	249 (4.8)	436 (64)

5.2.4. (iv) $[-v]_{fricative}[-v]_{fricative}$

A sample spectrogram for a voiceless fricative followed by another voiceless fricative is presented in spectrogram (5.11). The boundary between *s* and *f* are clearly visible in the spectrogram, with greater energy levels visible in higher frequencies during the production of *s* but not during the production of *f*. Quantitative results immediately follow the spectrogram in section 5.3.4.1.

0.791



Spectrogram 5.11: Speaker –1 – ***sf****arim* 'books'.

5.2.4.1. Quantitative results for $[-\mathrm{v}]_{\mathbf{fricative}}[-\mathrm{v}]_{\mathbf{fricative}}$

Table (5.10) provides data for the first member of the cluster *s*. Once again, the only relevant measurement is the duration of the fricative. Table (5.10) is immediately followed by table (5.11) which provides information about the second fricative member of a $[-\mathrm{v}]_{\mathrm{fricative}}[-\mathrm{v}]_{\mathrm{fricative}}$ cluster.

Table 5.10: Voiceless fricative as a first member of a cluster.

C1	**s**
	fric. dur.
s. 1	0.163 (0.02)
s. 2	0.147 (0.012)
s. 3	n/a [6]
s. 4	0.147 (0.026)
s. 5	0.172 (0.034)
s. 6	0.163 (0.009)

Table 5.11: Voiceless fricative as a second member of a cluster.

C2	**f**	
	fric. dur.	***f*0**
s. 1	0.101 (0.016)	151
s. 2	0.100 (0.011)	134 (2.9)
s. 3	n/a	n/a
s. 4	0.107 (0.009)	230 (4.29)
s. 5	0.11 (0.013)	237
s. 6	0.107 (0.006)	234 (12.3)
C2	**x**	
	fric. dur.	***f*0**
s. 1	n/a	n/a
s. 2	n/a	n/a
s. 3	n/a	n/a
s. 4	0.140 (0.013)	220 (11.25)
s. 5	0.135 (0.009)	239 (6.6)
s. 6	0.138 (0.006)	236 (5.17)

Both *x* and *f* in C2 position are shorter when C1 is a fricative than when C1 is a stop consonant. Generally *f* seems to be shorter than *x*. Data for fundamental frequency (***f*0**) were difficult to obtain and in many cases the default pitch tracker was unable to provide the data. Data for ***f*0** for speaker 1 and speaker 3 were obtained manually, at times not at the onset of the vowel but at the next closest possible point available. Data for *x* were not available for the male speakers because the cluster *sx*

[6] No data is available for speaker 3.

was not recorded for male speakers. Data for *x* in C2 position were available on the expanded list recorded for the female speakers.

5.2.5. Discussion of voiceless clusters

Results of the experiment outlined thus far show that in $[-\text{v}]_{\text{stop}}[-\text{v}]_{\text{stop}}$ clusters, both consonants are released and both consonants generally contain a portion of aspiration. This concurs with results found for Georgian by Chitoran (1998, 1999), where she shows that for Georgian stop sequences both stop segments are released. The large portion of aspiration that accompanies both segments counters what was found in English (Browman and Goldstein 1986). In English, when a stop occupies C2 position, it generally lacks that aspiration that it carries when it precedes a vowel. Clearly, this is not the case in Modern Hebrew, where both stop consonants contain a large aspirated release portion. The aspirated release is realized in stop consonants in both C1 and C2 positions.

Aspiration in stops in C1 position is significantly shorter when the stop is followed by a fricative. This is to be expected given the acoustic similarity between the aspirated portion of the stop release and the frication noise of the fricative.

In a $[-\text{v}]_{\text{fric}}[-\text{v}]_{\text{fric}}$ cluster, the second fricative is considerably shorter than the first. Both *x* and *f* are shorter when they follow a fricative C1 than a stop C1 and while *x* is marginally shorter, *f* is considerably shorter. There is little difference in the duration of a fricative C1, which is almost always *s* due to the fact that *f* and *x* are much more restricted and are less likely to occur in C1 position for historical reasons.

Voiceless clusters are realized as anticipated; no vocal fold vibration occurs in any of the tokens tested. Other than speaker variation in the duration of aspiration, duration of burst and fricative duration, there is very little variation in the realization of voiceless clusters. The variation in aspiration is more prominent when taking into account the large standard deviation that most speakers exhibit. This is also to be expected given the fact that aspiration is not phonologically distinctive in Modern Hebrew and therefore its presence to a greater or lesser degree is less crucial. Speakers can afford to be sloppy in the realization of this phonetic feature since it is not categorically distinctive. Other linguistic factors such as place of articulation, rate of speech and immediate phonetic environment play a significant role in the amount of aspiration we witness in the production of voiceless stops. In the chart in (5) I summarize the various realizations of an underlying /−v −v/ cluster. As can be seen from this chart, underlyingly /−v −v/ clusters are realized only as surface [−v][−v] clusters.

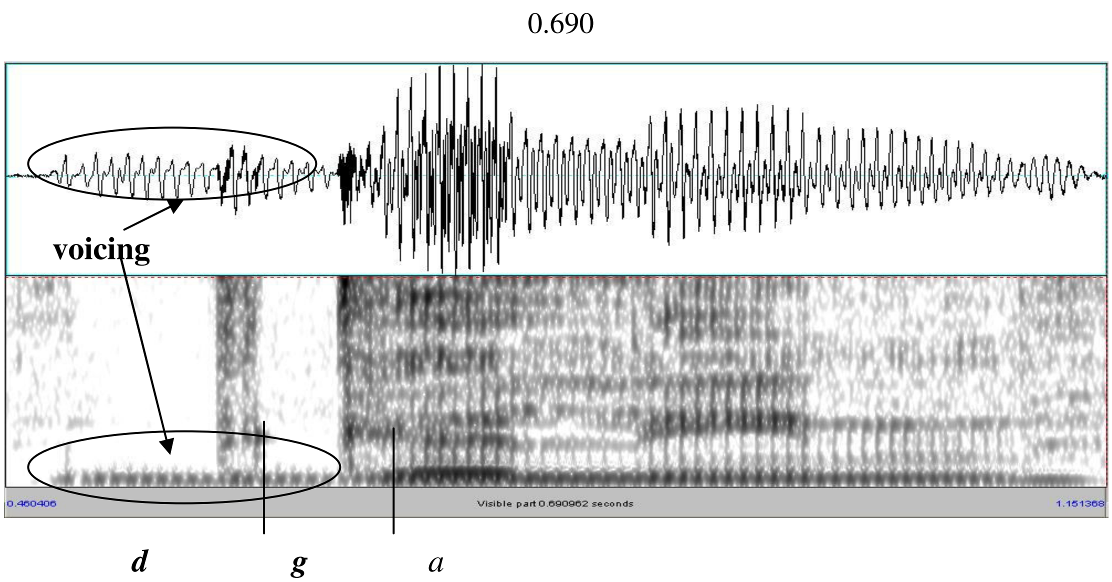
5)	Underlying target cluster		Phonetic realization
	/−v −v/	⟶	[−v][−v]
			*[+v][+v]
			*[−v][+v]
			*[+v][−v]

5.3. [+v][+v] clusters

5.3.1. (i) $[+v]_{stop}[+v]_{stop}$

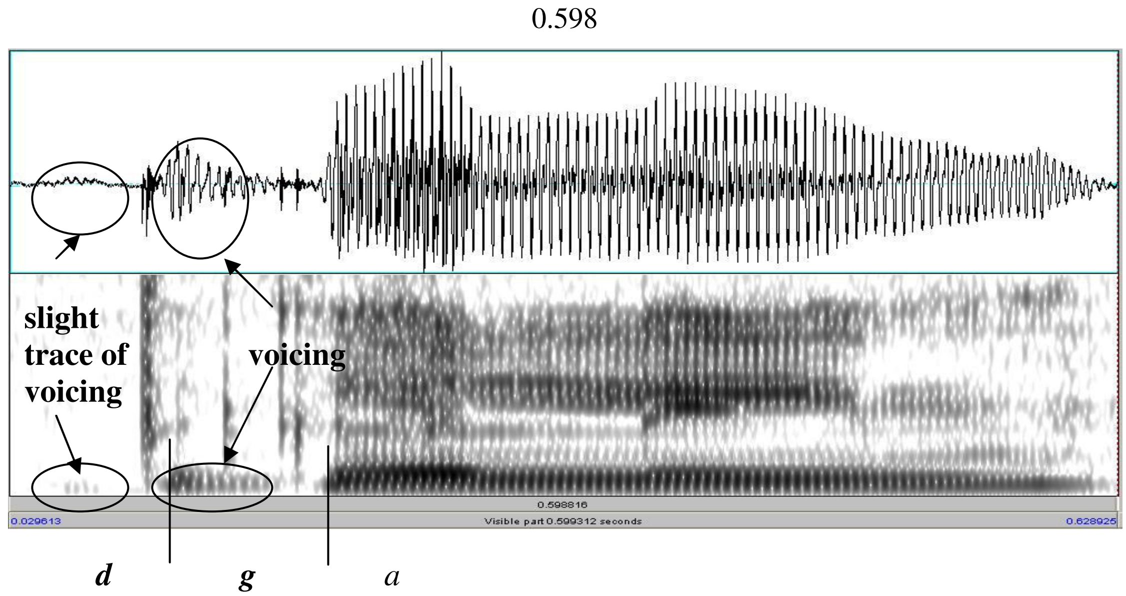
Although we expect both segments to be phonetically voiced in [+v][+v] clusters, we do observe great variation in the realization of $[+v]_{stop}[+v]_{stop}$ clusters. In stops we should expect to see closure voicing for both C1 and C2. However, as will be shown in the remainder of this section, this is not always the case.

Spectrogram (5.12) is a typical spectrogram for a sequence of voiced stops. Voicing is clearly evident in both segments. Both segments have closure voicing preceding the release and vibration of the vocal folds prior to release is quite visible in both spectrogram and waveform. Both segments are released with noticeable bursts.

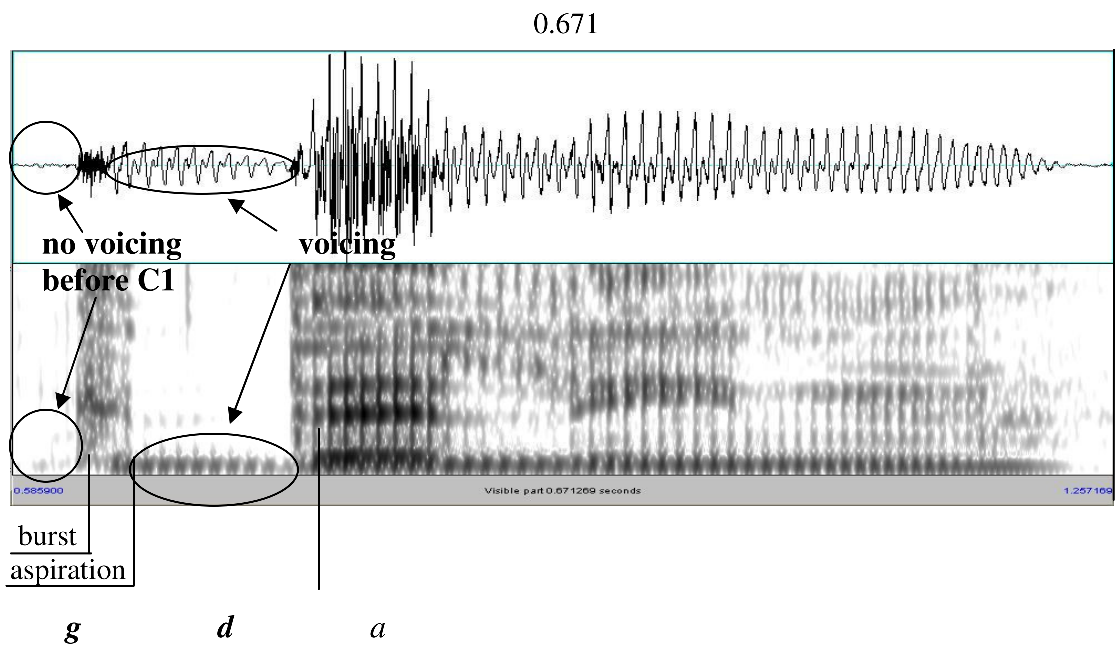


Spectrogram 5.12: Speaker – 1 – ***dg**alim* ‘flags’.

However, the voicing pattern exhibited in spectrogram (5.12) is not the only method for realizing voicing in [+v][+v] clusters. Another pattern for realizing [+v][+v] clusters is one where the first member of the cluster contains no voicing at all or very little voicing as exemplified by spectrograms (5.13) and (5.14), where it is evident that the first segment, although expected to have some closure voicing, contains very slight traces of voicing as in spectrogram (5.13) or no voicing at all, as in spectrogram (5.14). The slight traces of voicing in spectrogram (5.13) suggest that the phonological target of the segment is indeed voiced (/+v/) but is not fully realized phonetically. Even though the first voiced member of the cluster is followed by another voiced segment and there is no trigger for “voicelessness” in the first segment, the first member has virtually no voicing.



Spectrogram 5.13: Speaker – 4 – ***dg****alim* ‘flags’ – rep. no. 1

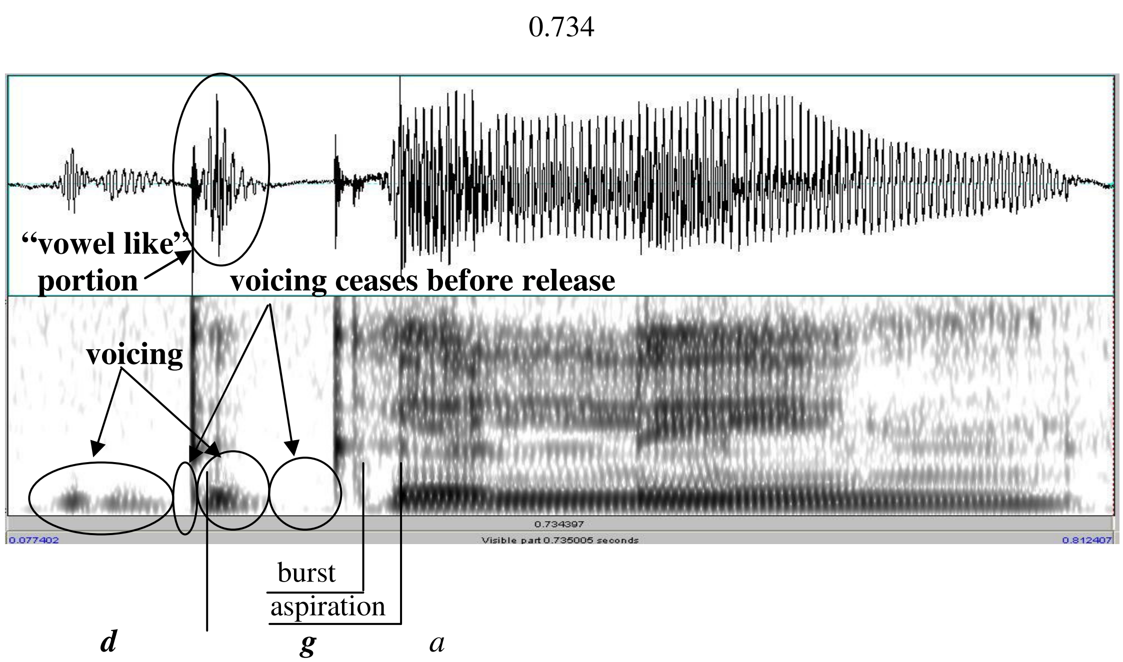


Spectrogram 5.14: Speaker – 1 – ***gd****alim* ‘sizes’.

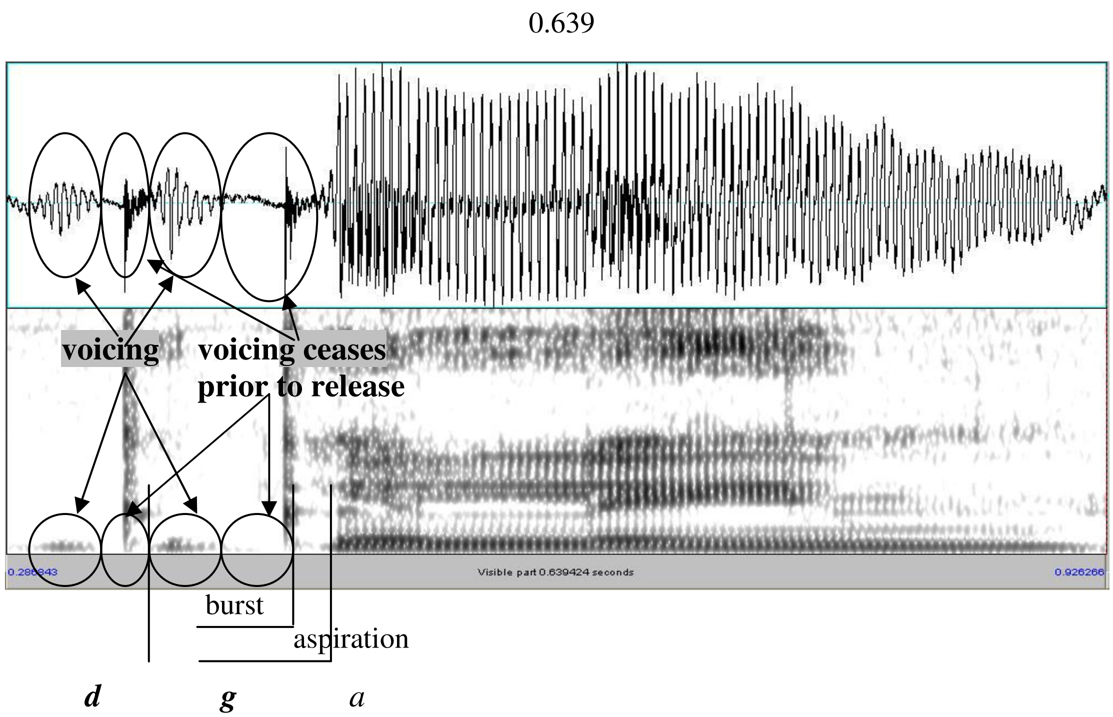
These realizations of [+v][+v] clusters suggest that voicing fails to be realized despite [+voice] being the underlying voicing specification of the segments. This pattern or realization of voiced segments occurs due to the physiological difficulty in producing voicing in absolute initial environment, a highly unfavorable environment for voicing, as discussed in chapter 4.

Failure to produce voicing in absolute initial position is quite rare in singletons, which suggests that the difficulty in producing voicing in both members of a [+v][+v] cluster, as well as maintaining voicing over the longer duration of the cluster, must be attributed to a difficulty in initiating and maintaining voicing specifically in clusters but not in singletons. This can account for this curious and unexpected pattern demonstrated by spectrograms (5.13) and (5.14).

The next pattern which demonstrates the difficulty in maintaining voicing in stops is a pattern which we predicted in chapter 4 section 4.6, where voicing diminishes before the burst release as can be seen in spectrograms (5.15) and (5.16). Note the "vowel like" release which will be addressed in section 5.3.5.1.



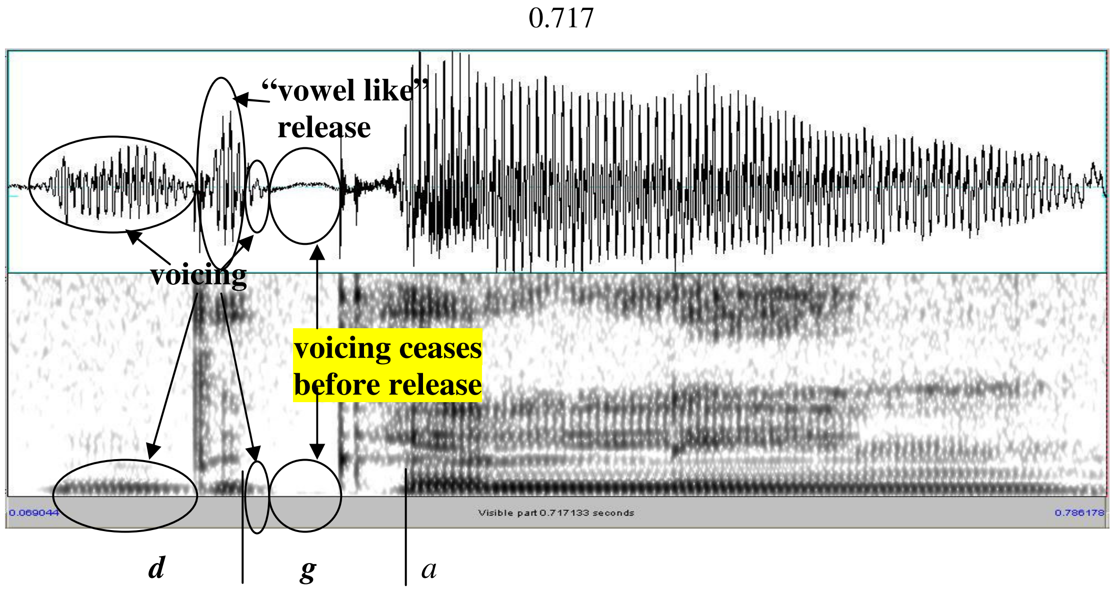
Spectrogram 5.15: Speaker – 4 – ***dg****alim* 'flags' – rep. no. 2.



Spectrogram 5.16: Speaker – 4 – ***dg****alim* ‘flags’ – rep. no. 4.

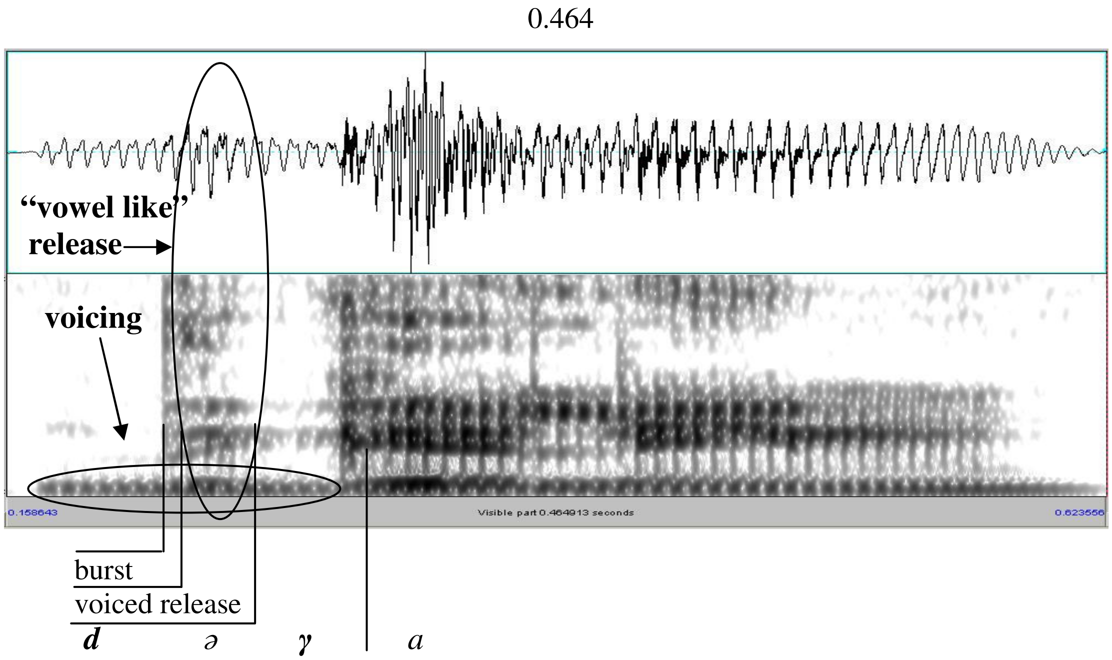
In spectrogram (5.15) and spectrogram (5.16), we can see that the voicing subsides before the release of both the first and the second voiced segments. This is not an uncommon pattern, which often arises due to an increase in the supra-glottal pressure and its equalization with the sub-glottal pressure (Kingston and Diehl 1994). When both sub-glottal and supra-glottal pressures equalize, vocal fold vibration cannot persist as vocal fold vibration is dependent on pressure differences. In these cases voicing is assumed to be initiated twice, with two separate gestures for voicing initiation.

However, voicing does not have to necessarily cease before both stop releases. In the example in spectrogram (5.17), voicing ceases before the release of the second stop but not before the release of the first.

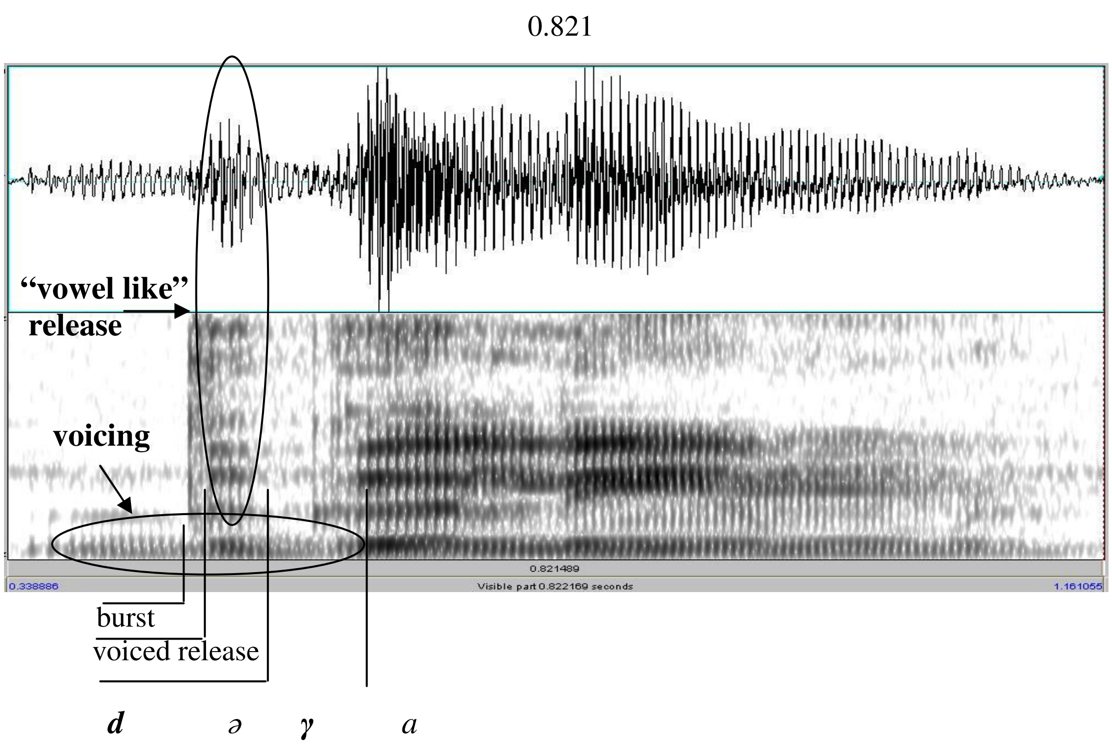


Spectrogram 5.17: Speaker – 4 – ***dg**alim* 'flags' – rep. no. 3

In spectrograms (5.15-5.17) there appears a "vowel-like" release after the first voiced segment, which is not a real vowel. This vowel like release appears in the context of [+v][+v] clusters or [−v][+v] clusters and functions as a transition between C1 and C2. It is not an epenthetic vowel but rather a vocalic portion caused by the transition of the articulators from one configuration to the next. It will be discussed further in section 5.3.5.1. In the example in spectrogram (5.18) the vowel like release, which is circled in the spectrogram and transcribed as a schwa under the spectrogram, is very prominent. Moreover, note that the *g* in both spectrogram (5.18) and (5.19) is produced with incomplete closure and seems to be almost lenited to *γ*.

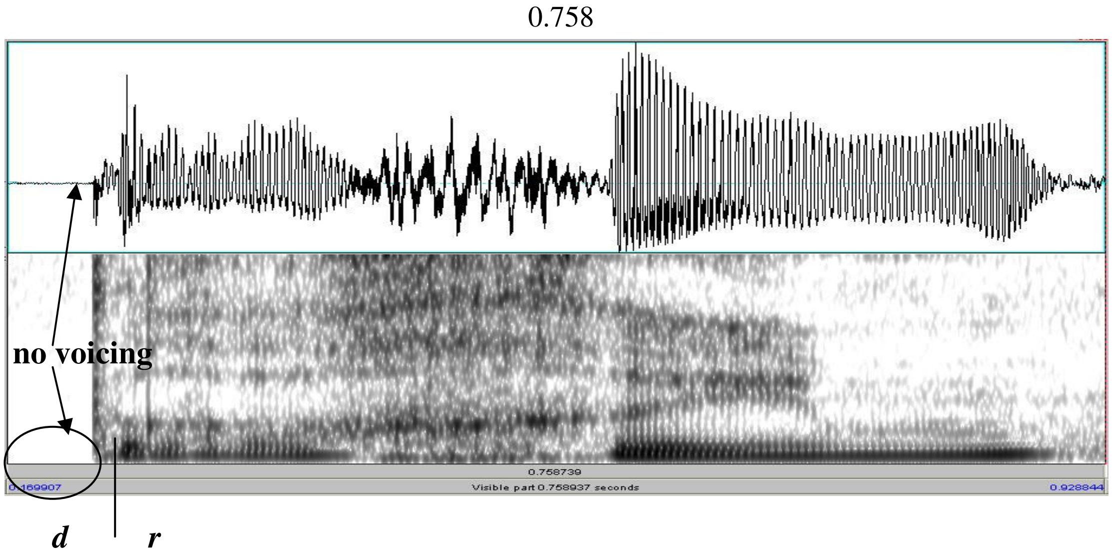


Spectrogram 5.18: Speaker – 2 – ***dg**alim* 'flags'.



Spectrogram 5.19: Speaker – 3 – ***dg**alim* 'flags'.

When a voiced C1 precedes a sonorant, it is not expected to surface as voiceless, since sonorants do not trigger voicelessness.[7] In light of this claim spectrogram (5.20) is puzzling, since C1 surfaces as voiceless, although it is followed by a sonorant and no trigger for voicelessness exists.

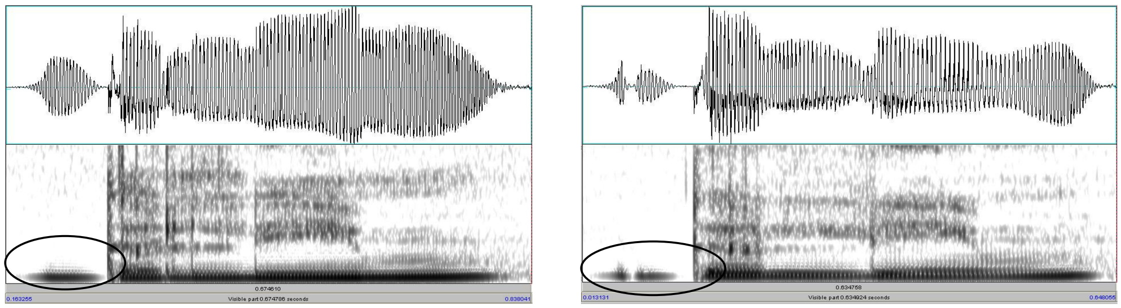


Spectrogram 5.20: Speaker – 5 – ***dr****axim* 'roads/paths/ways'.

Spectrogram (5.20) is a sample from speaker 5 who fails to voice voiced targets in word initial clusters more frequently than other speakers. This is not true for singletons for the same speaker. That is, voiced segments appearing before a vowel as in the two spectrograms in (5.21) always contain closure voicing, even for speaker 5. This further confirms the claim that initiating and maintaining voicing is physiologically more complex in clusters than in singletons.

[7] I do not take a stand on whether voicing is a uniary or binary feature. Assuming uniary features means that voicelessness is not a specified phonological target. If, however, we assume binary features then voicelessness is a target. However, this work does not offer conclusive evidence for either theory nor makes any claims regarding either theory.

0.674 0.634



d *d*

Spectrogram 5.21: Speaker – 5 – ***d****alim* ‘poor’ (2 tokens)

5.3.1.1. Quantitative results for $[+v]_{stop}[+v]_{stop}$

The spectrograms in section 5.3.1 illustrate the great variation in the realization of the $[+v]_{stop}[+v]_{stop}$ cluster type. Tables (5.12) and (5.13) will demonstrate the distribution of the patterns presented in section 5.3.1.

In table (5.12) data for C1 of a $[+v]_{stop}[+v]_{stop}$ cluster are presented. Closure voicing is presented in the column entitled **c.v.** and the number of tokens that contain closure voicing is listed in the column **# c.v.** For example, the number 4/5 for speaker 1 for the segment *b*, means that 4 out of 5 tokens were voiced and the other 1 token contains no voicing during closure. The next column is the duration of the burst, and once again, as in previous cases, here too the standard deviation can get quite large. The column **v.r**. stands for **v**oiced **r**elease which is “the vowel like portion” of the release. Notice that the voiced release never exceeds 30 milliseconds, which is too short to be a real vowel. The following column, **# v.r.** presents the number of tokens that contain a voiced release. For example, for speaker 1 for *b* 3/5 means that 3 tokens of the 5 recorded have a voiced release and 2 of the 5 tokens have no voiced release. The data suggest that voiced releases are more common for alveolars than for labials and velars. Evident from table (5.12) is the fact that for all voiced tokens, there is either a closure voicing or a voiced release. This is reflected in the column **no v.r. + no c.v**. However, the stop release does not replace closure voicing; that is, a stop segment can have both, closure voicing as well as a voiced release. By the same token, it is also possible for a segment to contain neither closure voicing nor a voiced release, although this pattern is somewhat less common, with only two such tokens occurring for speaker 4, one in *b* and one in *d*.

Table 5.12: Voiced stops as first member of a cluster.

C1	**b**						
	c.v	**# c.v**	**burst dur.**	**v. r**	**# v.r**	**no v.r. +no c.v**	**asp.**
s. 1	0.107 (0.035)	4/5	0.010 (0.004)	0.024 (0.002)	3/5	0/5	0.022 (0.005)
s. 2	0.066 (0.030)	4/4	0.010 (0.003)	0.024	1/4	0/4	No asp.
s. 3	n/a	n/a	n/a	n/a	n/a	n/a	n/a
s. 4	0.050 (0.003)	3/4	0.004 (0.003)	0.025	1/4	1/4	No asp.
s. 5	No voi	0/4	0.008 (0.001)	0.020 (0.008)	2/4	2/4	0.013 (0.003)
s. 6	0.056 (0.001)	3/4	0.008 (0.003)	0.02 (0.005)	3/4	0/4	0.009 (0.003)
C1	**d**						
	c.v	**# c.v**	**burst dur.**	**v. r**	**# v.r**	**no v.r. +no c.v**	**asp.**
s. 1	0.147 (0.032)	4/4	0.012 (0.002)	0.03 (0.01)	4/4	0/4	0.030 (0.01)
s. 2	0.078 (0.016)	3/3	0.018 (0.007)	0.017	1/3	0/3	0.009
s. 3	0.080 (0.031)	4/4	0.011 (0.004)	0.029 (0.011)	4/4	0/4	No asp.
s. 4	0.087 (0.008)	3/4	0.011 (0.004)	0.021 (0.002)	3/4	1/4	No asp.
s. 5	0.063	1/4	0.008 (0.002)	0.028 (0.009)	4/4	0/4	0.010 (0.003)
s. 6	0.065 (0.011)	3/4	0.011 (0.005)	0.027 (0.01)	4/4	0/4	0.014 (0.001)
C1	**g**						
	c.v	**# c.v**	**burst dur.**	**v. r**	**# v.r**	**no v.r. +no c.v**	**asp.**
s. 1	0.082 (0.058)	1/4	0.016 (0.007)	0.027 (0.009)	0/4	0/4	0.026 (0.008)
s. 2	0.084 (0.027)	1/4	0.012 (0.003)	0.027 (0.005)	0/4	0/4	No asp.
s. 3	0.080 (0.031)	0/4	0.011 (0.004)	0.029 (0.011)	1/4	0/4	No asp.
s. 4	0.061 (0.021)	1/4	0.006 (0.003)	0.029 (0.002)	0/4	0/4	0.012 (0.004)
s. 5	No voi.	0/3	0.008 (0.004)	0.026 (0.005)	1/3	0/3	0.01
s. 6	0.074 (0.030)	2/4	0.006 (0.001)	0.029 (0.008)	0/4	0/4	0.011 (0.004)

In the next table we explore the realization of C2 of a $[+v]_{stop}[+v]_{stop}$ cluster.

Table 5.13: Voiced stops as a second member of a cluster.

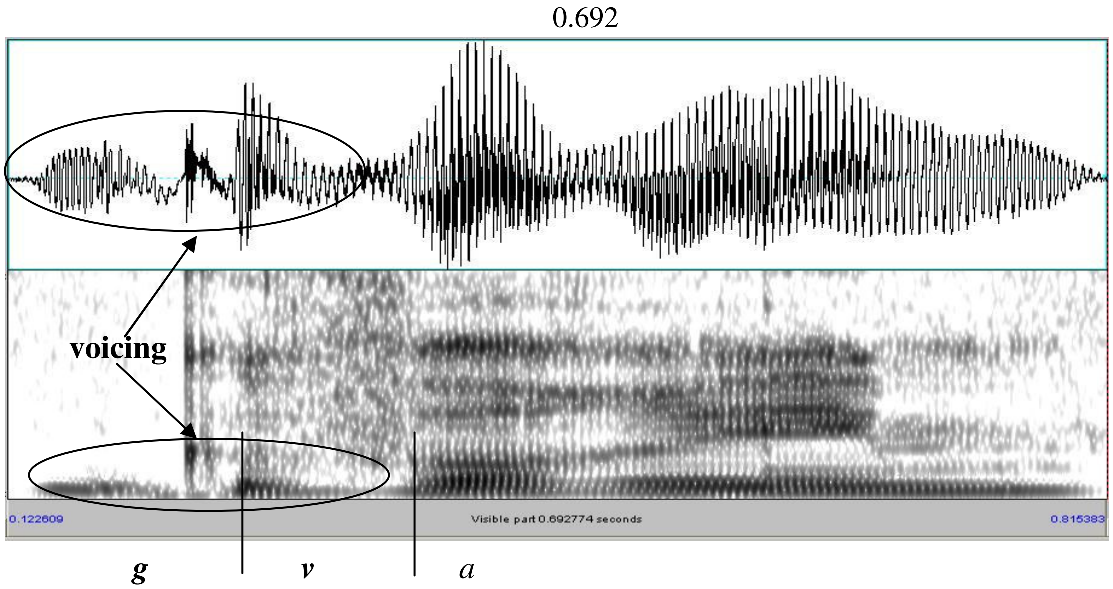
C2	**d**					
	c.v	**burst dur.**	**closure dur.**	**# c.v**	**asp.**	***f*0**
s. 1	0.106 (0.005)	0.012 (0.004)	0.09 (0.006)	4/4	0.009 (0.004)	127 (16)
s. 2	0.069 (0.018)	0.014 (0.002)	0.058 (0.014)	4/4	0.005	125 (2.4)
s. 3	0.085 (0.025)	0.011 (0.006)	0.069 (0.021)	4/4	0.011 (0.001)	134 (3)
s. 4	0.071 (0.012)	0.007 (0.001)	0.058 (0.023)	4/4	0.046	191 (13)
s. 5	0.079 (0.022)	0.006 (0.002)	0.093 (0.019)	3/3	0.011 (0.002)	229 (23)
s. 6	0.050 (0.004)	0.008 (0.003)	0.071 (0.004)	3/4	0.016 (0.006)	230 (10)
C2	**g**					
	c.v	**burst dur.**	**closure dur.**	**# c.v.**	**asp.**	***f*0**
s. 1	0.075 (0.028)	0.015 (0.005)	0.058 (0.008)	9/9	0.017 (0.008)	109 (5)
s. 2	0.065 (0.008)	0.013 (0.005)	0.050 (0.007)	7/7	0.007 (0.001)	128 (5)
s. 3	0.077 (0.013)	0.017 (0.004)	0.052 (0.008)	4/4	0.009 (0.003)	134 (3)
s. 4	0.027 (0.019)	0.014 (0.004)	0.066 (0.010)	8/8	0.015 (0.004)	219 (18)
s. 5	0.028 (0.009)	0.012 (0.006)	0.068 (0.012)	5/8	0.019 (0.007)	245 (7)
s. 6	0.040 (0.014)	0.006 (0.002)	0.066 (0.005)	7/8	0.019 (0.005)	250 (16)

In table (5.13) the first column represents closure voicing, which gives the average duration of voicing during the closure of the stop in C2 position. The next column represents the duration of the burst of the stop. The column **closure dur.** lists the duration of the stop closure C2 in $[+v]_{stop}[+v]_{stop}$ clusters. The number of tokens that contain voicing during closure follows under the column **# c.v.** For speaker 5 for *g* 5/8 in the column **# c.v.** indicates that 5 of 8 tokens were realized with closure voicing. The column **asp.** lists the aspiration following the release burst. It appears that aspiration following the burst, as expected, is considerably shorter after voiced segments than aspiration following voiceless segments. The last column lists the fundamental frequency, which is measured in the vowel following a voiced segment. Fundamental frequency is lower after a voiced segment than after a voiceless segment, as predicted and discussed in chapter 4.

Of interest in table (5.13) is the fact that for speakers 5 and 6 there are several tokens that are produced with no closure voicing. This is entirely unexpected for a second member of a $[+v]_{stop}[+v]_{stop}$ cluster, since there is no trigger for voicelessness in the second member of the cluster, which precedes a vowel. This case is considered a case of extreme delayed voicing and in all cases where C2 is missing closure voicing, C1 is missing voicing as well. There are no $[+v]_{stop}[+v]_{stop}$ tokens that are phonetically realized with voicing in C1 and no voicing in C2. However, a voiceless C2 is not at all a common pattern for [+v][+v] clusters, although it is more common in $[+v]_{stop}[+v]_{stop}$ clusters than in other combinations of manner of articulation, as will be seen in the remainder of the section. If fact, there is much less variation in the realization of voicing in other manner combinations as will be seen in the remainder of the chapter.

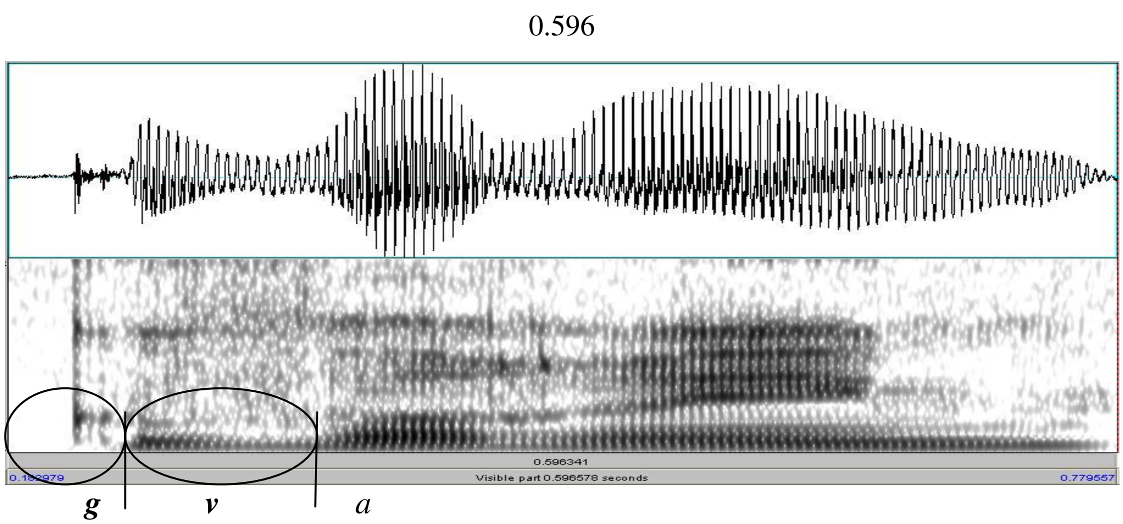
5.3.2. (ii) $[+v]_{stop}[+v]_{fricative}$

In the case of stops preceding a fricative, there is much less variation in the realization of voicing in both consonants. The following spectrogram is a representative example of the majority of cases. Voicing lasts through most of the cluster. Due to the lower number of voiced fricative phonemes in MH, the only fricatives that may occupy C2 position are either *v* or *z*, since there is no voiced uvular fricative.



Spectrogram 5.22: Speaker – 4 – ***gv****arim* ‘men’.

A less common pattern of voicing in stop-fricative voiced clusters is demonstrated in spectrogram (5.23), where C1 contains no closure voicing.



Spectrogram 5.23: Speaker – 4 – ***gv****arim* ‘men’.

In $[+v]_{stop}[+v]_{fricative}$ clusters, the pattern illustrated in spectrogram (5.23) is quite rare but does surface occasionally and resembles the pattern in $[+v]_{stop}[+v]_{stop}$ where the first voiced segment lacks voicing, demonstrated in spectrograms (5.13), (5.14) and in table (5.13). This pattern is much more common in female speakers than in male speakers as will become evident in the quantitative data in the following section.

5.3.2.1. Quantitative results for $[+v]_{stop}[+v]_{fricative}$

Table 5.14: Voiced stops as a first member of a cluster.

C1	**g**				
	c.v	**no c.v**	**burst dur.**	**v. r**	**asp.**
s. 1	0.096 (0.046)	0	0.021 (0.008)	0	No asp.
s. 2	0.100 (0.032)	0	0.011 (0.004)	0	No asp.
s. 3	0.096 (0.016)	0	0.014 (0.008)	0	0.013 (0.003)
s. 4	0.083 (0.009)	1/4	0.012 (0.008)	0	0.016 (0.006)
s. 5	0.050 (0.017)	2/4	0.009 (0.003)	0	0.02 (no s.d)
s. 6	0.054 (0.019)	1/4	0.010 (0.004)	0	0.017 (0.011)

As in previous tables the column **c.v.** lists the duration of the voice bar which precedes the release burst (**c**losure **v**oicing). The following column, **no c.v** enumerates the number of tokens which lack any voicing prior to the release burst. For speakers 1, 2 and 3 no tokens lack voicing, which means that all C1 tokens contain a voiced portion. For speakers 4 and 6 1 token of the 4 was missing the voiced portion and for speaker 5, half the tokens had no voicing. The next column is the duration of the release burst. The column **v.r** stands for **v**oiced **r**elease, which is the vowel like portion that we saw in the transition from one stop to the next in $[+v]_{stop}[+v]_{stop}$ clusters. As evident from table (5.14), there is no voiced transition from a stop C1 to a fricative C2 since voiced releases occur only between two stops and never occur before a fricative. Lastly, the aspiration is represented in the last column **asp**. Results concur with most predictions. Namely, most tokens contain a certain voiced portion, although there is great variation in the duration of closure voicing. This, in itself, is also expected, since as pointed out in section 5.1.6, it is only the presence or absence of voicing that is important for the perception of voicing, not the duration of voicing. More surprising are the tokens that lack voicing, since there is no trigger for voicelessness in [+v][+v] clusters. In the next table, we will explore data for voicing in a fricative C2.

Table 5.15: Voiced fricatives as a second member of a cluster.

C2	v			
	fric. dur.	**voice dur.**	**% fric voiced**	***f*0**
s. 1	0.142 (0.016)	0.142 (0.016)	100	103 (3)
s. 2	0.114 (0.003)	0.114 (0.003)	100	120 (3.4)
s. 3	0.109 (0.017)	0.109 (0.017)	100	132 (3)
s. 4	0.113 (0.014)	0.099 (0.023)	87	201 (7)
s. 5	0.154 (0.019)	0.135 (0.050)	88	224 (32)
s. 6	0.129 (0.026)	0.134 (0.029)	104	233 (4)

The first column, **fric. dur.,** lists the duration of the fricative. The next column **voice dur.** is the duration of the voiced portion of the fricative only. For all male speakers the voiced portion of the fricative is the same as the duration of the fricative, that is, for all male speakers all fricatives are completely voiced through their entire duration. This is not the case for speakers 4 and 5, initially it appears that based on the averages listed, fricatives are not fully voiced. But note the large standard deviations for both speakers 4 and 5 versus the considerably smaller standard deviations for speakers 1, 2 and 3. For both speakers 4 and 5 only 1 of the 4 fricatives is not fully voiced, causing the averages to appear skewed. Speaker 6 is the only one for whom the portion of the fricative which is voiced is more than 100%. This is due to one of the tokens, which is shorter than all other tokens, being produced with absolutely no voicing and all other 3 tokens being produced with 100% voicing. The fundamental frequency following a voiced segment is lower than the fundamental frequency of a voiceless segment, as is expected.

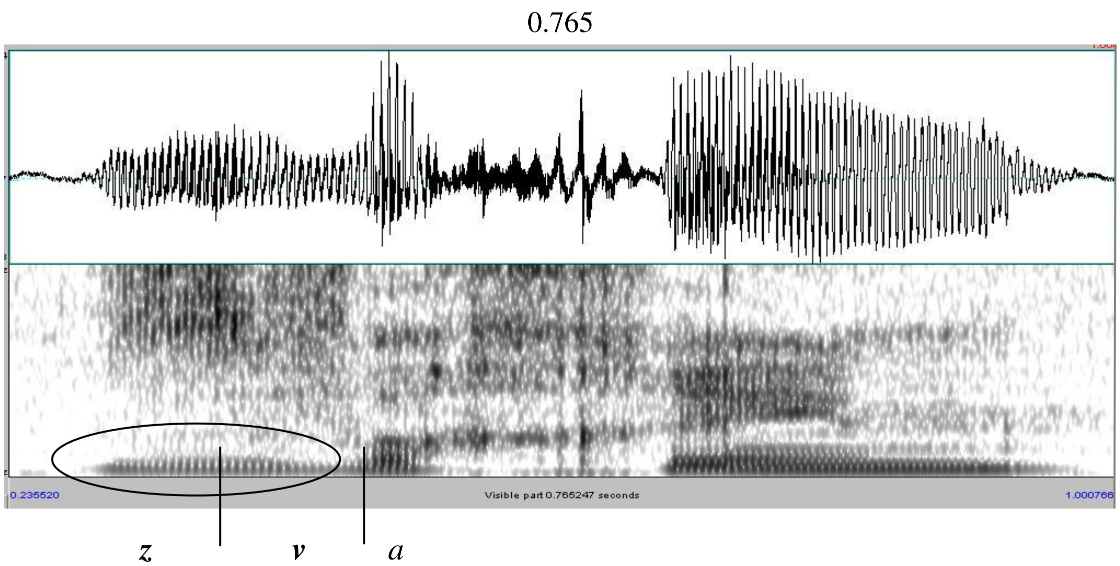
5.3.3. (iii) $[+v]_{fricative}[+v]_{stop}$

Generally, due to historical reasons, fricatives are rare in MH as first members of a cluster, and voiced fricatives are even more rare since there are only two of them (for a more detailed discussion see chapters 1 and 3). A further complication is the fact that stops are also rare in C2 position making the combination $[+v]_{fricative}[+v]_{stop}$ very uncommon. There are fewer clusters and fewer clustering possibilities for fricatives as first members of a cluster and stops as second members of a cluster. Since the number of words beginning with *v* is limited to several words, most of which have the same root: for example: *vradim* – 'roses', the only other possibility for words to begin with a voiced fricative is to begin with a *z*. There are few words that can exemplify $[+v]_{fric.}[+v]_{stop}$ cluster, like *zgugit* 'a glass,' however, since the author was unable to find minimal pairs consistent with the paradigm this cluster will not be analyzed in this work.

5.3.4. (iv) $[+v]_{fricative}[+v]_{fricative}$

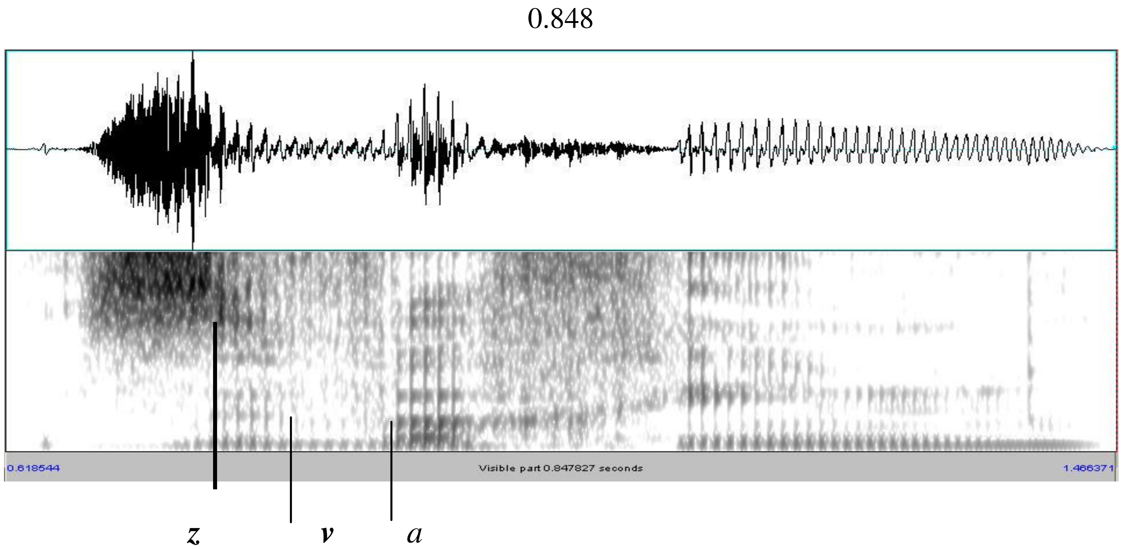
Due to the historical restrictions mentioned in chapters 1 and 3, the only $[+v]_{fricative}[+v]_{fricative}$ cluster that exists in Modern Hebrew is *zv*. Since *v* rarely appears in C1 position only *z* can appear in word initial position. Since *z* is the only possibility for a voiced fricative to occupy C1 position, the only option for a voiced fricative as the second member of a cluster is *v*.

There are three main strategies for producing this cluster. The first one is to produce voicing through the entire cluster with both fricatives fully voiced, as exemplified in spectrogram (5.24). The second strategy is by delaying voice production in the first segment and initiating voicing only in the middle of the first segment as exemplified by spectrogram (5.25). This pattern is not uncommon and both patterns are about equally common as will be evident from the quantitative data following the spectrograms.



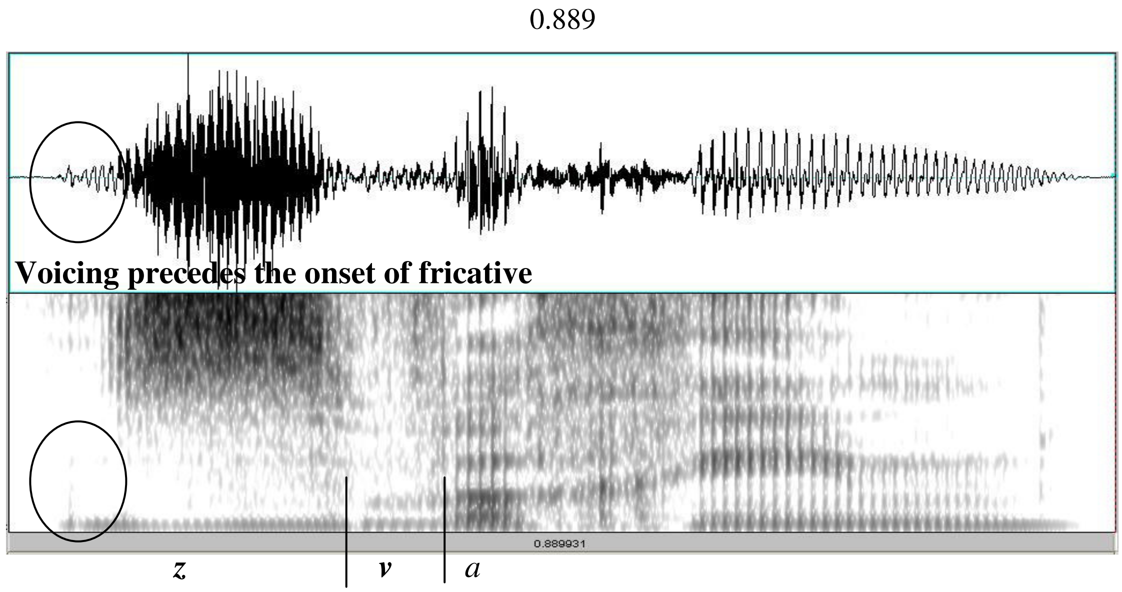
Spectrogram 5.24: Speaker – 4 – ***zv**axim* ‘sacrifices’.

Voicing can be seen to last through the entire cluster in spectrogram (5.24) and alignment between voicing initiation and consonant onset is optimal. Spectrogram (5.25) is an example of imperfect alignment between initiation of voicing and onset of segment. As emphasized in the spectrogram by a bolded line, voicing does not begin until the middle of the voiced fricative *z* and lasts all the way through the second member of the cluster. The first member of the cluster, the underlyingly voiced fricative, is only partially voiced phonetically. This delayed voicing initiation is not uncommon and is yet another example of realization of voicing in the unfavorable absolute initial position.



Spectrogram 5.25: Speaker – 1 – ***zv****axim* 'sacrifices'.

The last strategy involves initiating vocal fold vibration prior to the onset of the first segment, as exemplified by the spectrogram in spectrogram (5.26). This strategy is much more rare. I refer to this pattern as anticipatory voicing, where voicing begins prior to the initiation of frication. As can be seen in spectrogram (5.26), *v* is realized in such a manner that it resembles an approximant. Realizing the labial fricative as an approximant can be one of the strategies speakers employ for simplifying and realizing this segment in clusters.



Spectrogram 5.26: Speaker – 1 – *zvaxim* ‘sacrifices’.

5.3.4.1. Quantitative results for $[+v]_{fricative}[+v]_{fricative}$

In table (5.16) data is presented for the first segment of a $[+v]_{fricative}[+v]_{fricative}$ cluster. The first segment is a *z*, since, as has been mentioned in section 5.3.4, *z* is the only possible voiced fricative that can appear in C1 position.

Table 5.16: Voiced segments as second member of a cluster.

C1	z				
	voicing dur.	fric. dur.	tokens fully voiced	token vcd	% fric voiced
s. 1	0.140 (0.087)	0.160 (0.041)	1/4	4	87.5
s. 2	0.148 (0.022)	0.147 (0.015)	3/4	4	100
s. 3	n/a	n/a	n/a	n/a	n/a
s. 4	0.126 (0.035)	0.131 (0.028)	3/4	4	96
s. 5	0.078 (0.061)	0.163 (0.044)	1/4	4	48
s. 6	0.190 (0.034)	0.177 (0.032)	4/4	4	107[8]

[8] For this speaker only one token has vocal fold vibration preceding the initiation of frication, the rest of the tokens had 100% voicing through the duration of the fricative.

In table (5.16) voicing duration is listed in the column **voicing dur.** and is followed by **fric. dur.** which is the duration of the fricative. In the next column the number of the tokens that are fully voiced are listed in the column **tokens fully voiced** and the number of tokens that are voiced are listed in the next column **token vcd**. The next column **% fric. voiced**, gives the average percentage of the fricative which is voiced. Evident from the table is the fact that although all the tokens are voiced to some extent, few are fully voiced.

For speaker 2 in 2 of the 4 tokens voicing exceed 100%, meaning that voicing precedes the onset of the fricative. This pattern, exemplified by spectrogram (5.26), also exists for one token for speakers 1 and 6 (respectively). The most common pattern is for tokens to be fully voiced but delayed voicing is also quite common. Of the 20 tokens measured across 5 speakers 12 tokens are fully voiced, or more than fully voiced (when voicing precedes onset of fricative), and 8 tokens have delayed voicing.

Table 5.17: Voiced fricatives as a second member of a cluster.

C2	**v**			
	voice dur.	**fric. Dur.**	**% fric voiced**	***f*0**
s. 1	0.088 (0.023)	0.093 (0.021)	95	103 (4)
s. 2	0.084 (0.022)	0.084 (0.022)	100	120 (3)
s. 3	n/a	n/a	n/a	n/a
s. 4	0.067 (0.017)	0.067 (0.017)	100	192 (5.6)
s. 5	0.049 (0.014)	0.065 (0.009)	75	213 (2)
s. 6	0.097 (0.011)	0.097 (0.011)	100	212 (5)

Generally, as evident from table (5.17), *v* in C2 position is fully voiced. Only for speakers 1 and 5 it appears as though C2 is not fully voiced. The reason for this is that for speaker 1 one token is not fully voiced and for speaker 5 two tokens are not fully voiced. However, unlike in the case of C1 position, in C2 position voicing is not delayed, rather, it subsides before the offset of the fricative. Voicing is more consistent in C2 position than in C1 position and the majority of C2 tokens are fully voiced. The fundamental frequency, as expected, is lower after a voiced segment than a voiceless segment as evident from a comparison of *f0* in table (5.17) and *f0* in table (5.11).

5.3.5. [+v][+v] clusters summary and discussion

Articulation of two consecutive voiced obstruents in word initial position is physiologically difficult for two reasons. First, initiation of voicing in absolute initial position requires much more muscular effort; secondly, maintaining voicing in obstruents is more difficult due to increased oral pressure. Therefore, there is a much greater variation in the realization of the [+v][+v] cluster type than in the realizarion of [–v][–v] cluster type since the voiced clusters occur in a phonetically disadvantageous environment for realizing voicing (Keating and Westbury 1986). The phonologically voiced targets are not always realized as such phonetically. Speakers may choose various strategies to simplify the articulation of [+v][+v] clusters, as summarized in the scheme in (6).

6)	Underlying target		Phonetic realization
	Voiced stops	→	voiced stop (i.e. closure voicing)
		→	voiceless stop (no closure voicing)
	Voiced fricative	→	voiced fricatives
		→	voiced approximant
		→	voiceless fricatives

We begin our discussion with the first member of a [+v][+v] cluster The first member of a [+v]$_{stop}$[+v]$_{stop}$ onset cluster is generally realized with closure voicing, that is, vocal fold vibration usually precedes the burst. However, as voicing in utterance initial position is more complex, sometimes speakers fail to initiate vocal fold vibration and consequently fail to realize phonetically the phonologically voiced target. This results in a surface voiceless segment. Therefore, even in a voiced cluster, where we do not expect voiced segments to surface as voiceless, they do. The number of phonologically voiced C1 segments that are phonetically realized with no voicing, varies from speaker to speaker. Some speakers, like speaker 3, fully voice all C1 tokens, whereas speakers 1, 2, 4, 5 and 6 do not always realize voicing in C1. For speaker 1 15% of $[+v]_{stop}[+v]_{stop}$ tokens are realized with no voicing in C1 (2/13 tokens. For speaker 2 this number was smaller with only 9% (1/11) of all tokens realized with no voicing). For the female speakers these numbers were greater with speaker 4 realizing 25% (3/12) tokens with no voicing, speaker 5 was exceptional with 45% (7/11) of the tokens had no voicing in C1 in a $[+v]_{stop}[+v]_{stop}$ cluster and speaker 6 failed to realize voicing in 33% (4/12) of the tokens. The fact that this pattern repeats across most speakers, suggests that it is not speaker error and since voicelessness in this case is not the phonological target, we conclude that speakers fail to realize voicing in this context.

Lack of voicing in stops occurs in $[+v]_{stop}[+v]_{fricative}$ as well but only for the female speakers. It does not, however, occur in first members of a $[+v]_{fricative}[+v]_{stop}$ or $[+v]_{fricative}[+v]_{fricative}$ clusters. The fact that voicing fails in a stop C1 but not in a fricative C1 provides further evidence to the claim that surface voiceless stops in C1 position are a result of failure to produce voicing due to failure to reach the minimal threshold requires for initiating vocal fold vibration and that voicelessness is by no means a result of "devoicing" or voicing assimilation of any kind. Moreover, claims that surface voicelessness is a result of voicing assimilation cannot account for voicelessness in this particular environment since there is no trigger for voicelessness in the immediate phonetic environment to cause voicing assimilation. Furthermore, the difficulty in initiating vocal fold vibration in word initial position combined with the increased difficulty in maintaining voicing in stops, can easily account account for the fact that voicelessness occurs more frequently in stops than in fricatives.

Realization of voicing in the second member is also varied in [+v][+v] clusters. Voicing is generally realized in C2, though there are cases where it is not maintained through the entire duration of the segment. Rather, voicing decays before the release burst of C2, as exemplified in spectrograms (5.15), (5.16) and (5.17). Presumably this depends on the number of laryngeal gestures the speaker produces for the cluster. If only one gesture is produced, then voicing is partially maintained through the cluster, but the initial pressure build up is not sufficient to last through the entire duration of the cluster. If two laryngeal gestures are realized, then each of the segments might have a separate laryngeal gesture and voicing can be initiated for both C1 and C2. Recall that according to Löfqvist and Yoshioka (1980a, 1980b) and Munhall and Löfqvist (1992) two laryngeal gestures are possible for a sequence of two voiceless stops. We extended their findings to [+v][+v] clusters and hypothesize that two laryngeal gestures are also possible for voiced consonants and not just voiceless ones.

There are few cases where voicing is delayed so much that it does not begin until the vowel resulting in C2 being produced with no voicing at all. Other cases of incomplete voicing, especially in fricatives, include cases where voicing subsides towards the offset of the fricative. We never see cases where C1, the first member of the cluster is voiced but C2 is realized with no voicing.

To summarize the possible phonetic realizations of C1 are as follows:

7)	Underlying target cluster		Phonetic realization
	/+v +v/	→	[+v][+v]
		→	[−v][+v]
		→	([−v][−v]) rare
			*[+v][−v]

5.3.5.1. Voiced releases

Many of the $[+v]_{stop}[+v]_{stop}$ clusters contain a voiced release portion which was clearly visible in spectrograms (5.18-5.19). The voiced release portion is a vocalic release that resembles a schwa like vowel. However, this is not an underlying target vowel, but rather a vowel-like transition that is created as a result of the movement of articulators and continuous vibration of the vocal cords. The voiced release is by no means an underlying vowel; rather it is a *transitional vowel*, which is created as a result of the movement of the articulators from one position to the next and the tongue configuration during this movement. It is created during the release portion of the stop while vocal folds are still vibrating but before occlusion for the following stop consonant has been fully created. Since the voiced release is a transitional phase between one stop and another, it is found after C1 but not after C2. It is also absent from the transition of C1 in a $[+v]_{stop}[+v]_{fricative}$ cluster or from the transition of C1 in a $[+v]_{fricative}[+v]_{stop}$ or a $[+v]_{fricative}[+v]_{fricative}$ cluster. The voiced release does not replace closure voicing, and can appear in a segment together with closure voicing, that is a stop segment can have both a closure voicing as well as a voiced release.

The vowel like transition is similar to the transition found by Chitoran (1998, 1999) for Georgian: "… the C1 release in *gd* is followed by a short period of voicing, which looks like a vocalic portion". (1999: 103). This vocalic portion is created when the onset of C2 is misaligned and is delayed while the vocal cords keep vibrating during the transition of the articulators from their articulatory arrangement for C1 and their target articulatory arrangement for C2. This is illustrated in figure (5.2).

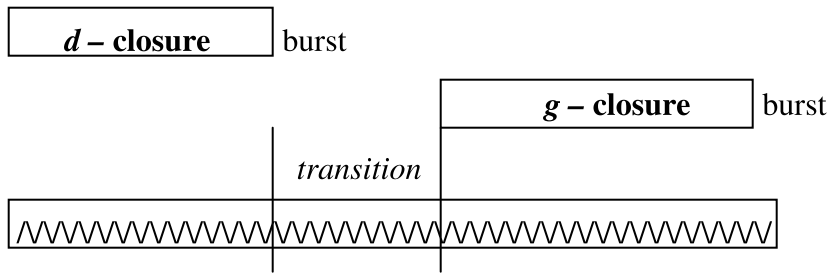


Figure 5.2: Illustration of vocal fold vibration in a $[+v]_{stop}[+v]_{stop}$ cluster.

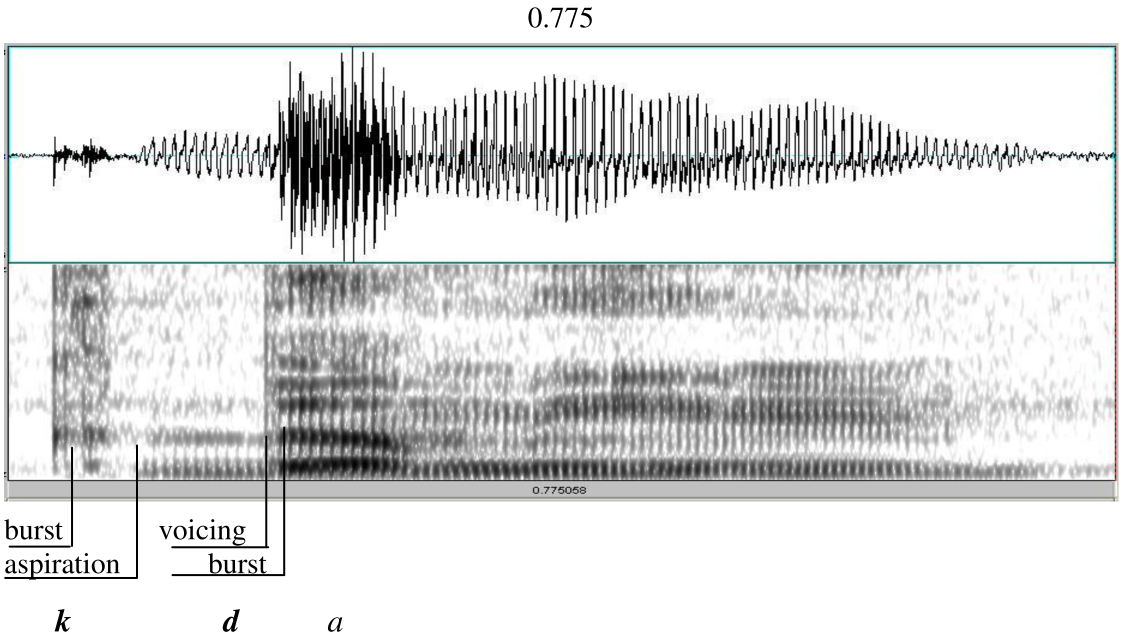
The vocalic portion is created during the transition of the articulators from one segment to the next. After the *d* has been fully released and before an occlusion is created in the velum for *g*, the vocal folds are still vibrating while the vocal tract is in an open position post release of the *d*; thus a vowel like portion results, which is not a phonological target but rather a transitional stage during the articulation of a bi-consonantal sequence. The transition, if it exists, never exceeds 30ms. for all speakers, suggesting that it is too short to be a vowel. From measurements of reduced, unstressed vowels in this experiment, the duration of reduced vowels in MH is 60ms., which is twice the length of the vocalic portion created between the two stop members of a [+v][+v] cluster. The duration of a reduced vowel in English is around 50ms. (Stevens 1998). The vocalic release is not unlike the release of an English final voiced consonant; if the final *d* in the word *bad* is released, a vocalic release resembling a schwa is evident in the spectrogram. However, this is clearly not an underlying target schwa but a release portion (Hertz, p.c).

Despite being transitional material, the voiced release may contribute to perception of voicing, although this must be kept separate from actual realization of voicing. The voiced release in the context of two voiced stops may be compared to aspirated release between two voiceless stops. Both are transitional material between two stop segments and both are a by product created while the articulators are transitioning from one segment to the next. But while vocal folds are fully spread and not vibrating during the transition of two voiceless segments, they do vibrate during the transition between two voiced segments. This may not be an entirely wrong comparison as it appears that the duration of voiced releases of voiced segments is roughly equivalent to the duration of aspiration of voiceless segments. Therefore, just as aspiration may be part of the release of a voiceless segment, similarly a voiced release is part of the release of a voiced segment.

5.4. [−v][+v] clusters

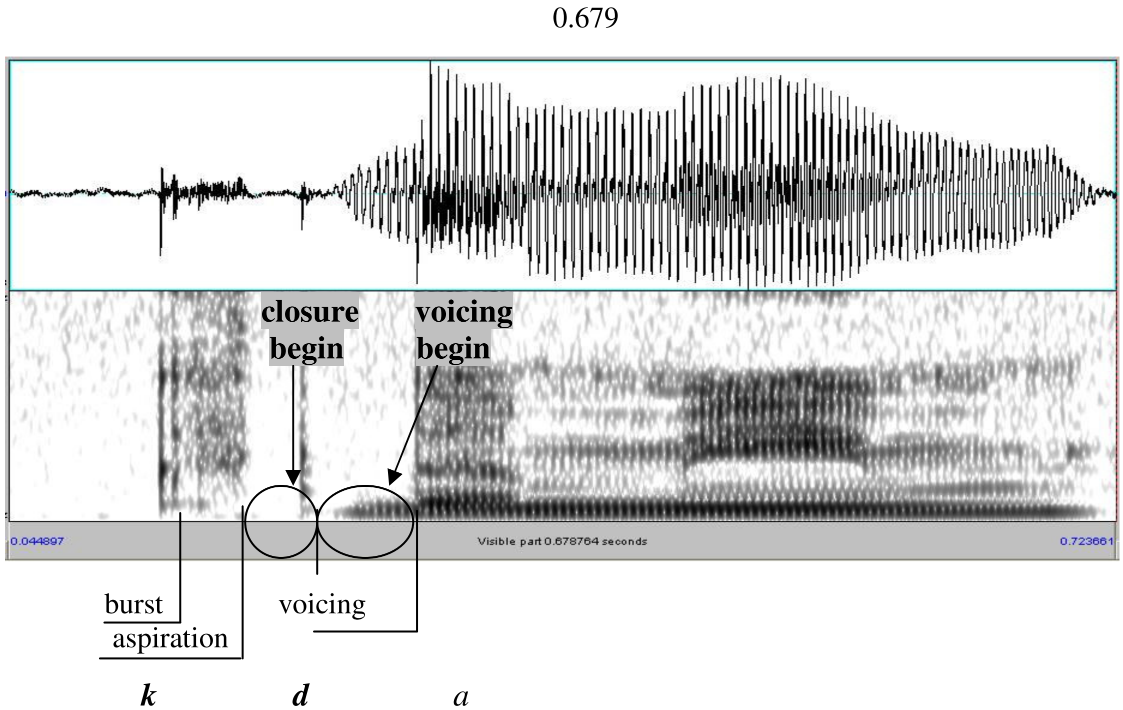
5.4.1. (i) $[-v]_{stop}[+v]_{stop}$

In sections 5.2 and 5.3 we examined clusters with the same voicing specifications in both members of the cluster. The third type of cluster we examine is a cluster with varying voicing specification, where C1 and C2 have different voicing specifications. We begin by exploring [–v][+v] clusters. Spectrogram (5.27) is an example of a voiceless segment followed by a voiced segment. In the spectrogram initiation of voicing in the second segment is aligned relatively well with the beginning of the closure for this segment. As expected, closure voicing is visible prior to the burst of C2. This pattern is the optimal realization of [−v][+v] cluster type.



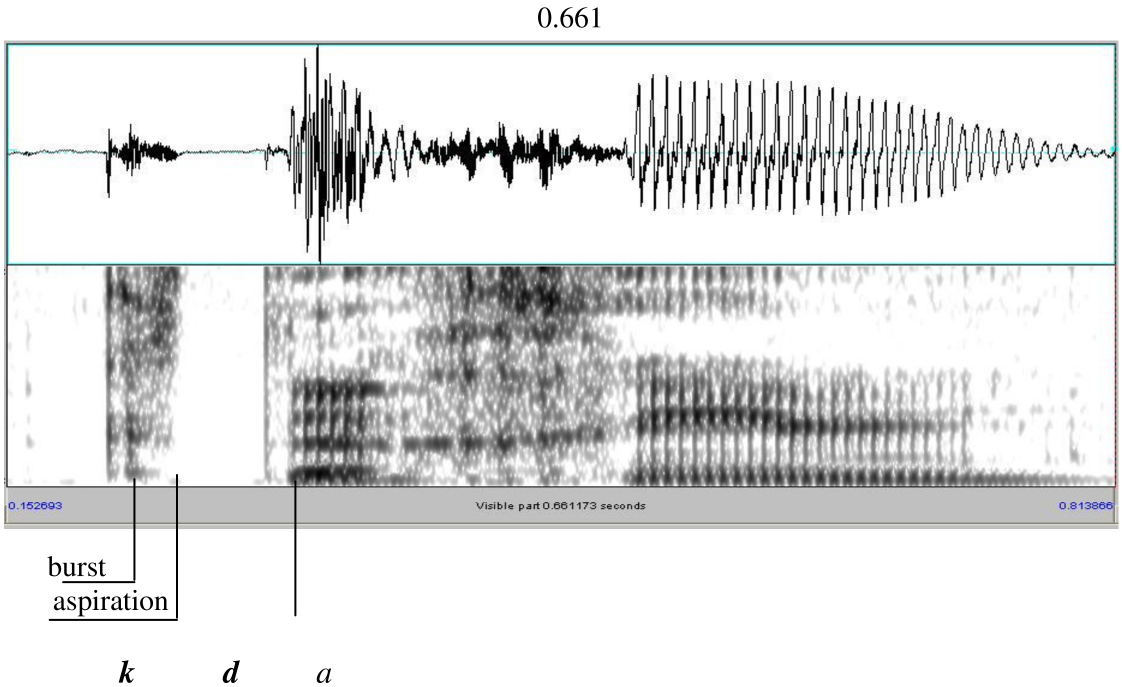
Spectrogram 5.27: Speaker – 3 – ***kd**amim* ‘precedent (plural)’.

However, the most common pattern for realizing voicing in $[-v]_{stop}[+v]_{stop}$ clusters is the pattern exemplified in spectrogram (5.28), where voicing of the second segment does not begin until the middle of the closure. Closure voicing still exists in these cases since voicing begins prior to the burst, however, there is a delay in the initiation of voicing. The voicing of *d* begins prior to the burst, in the middle of the closure. Clearly there is delayed voicing and the initiation of voicing is not aligned well with the initiation of closure. However, as long as closure voicing is present, even if it is partial, it is sufficient to perceive the segment as a voiced segment.



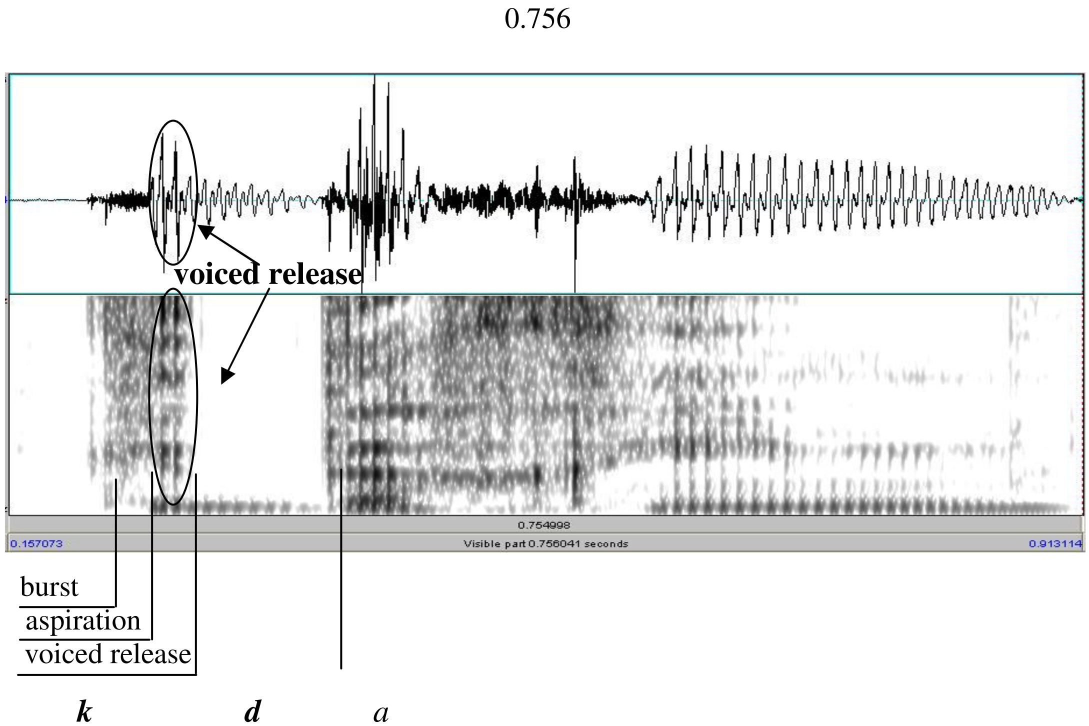
Spectrogram 5.28: Speaker – 4 – ***kd****amim* ‘precedence (plural)’.

A third variation for realizing [–v][+v] clusters is to produce both segments with no voicing at all, and with neither the first member of the cluster, nor the second member of the cluster, which is underlyingly voiced, containing any voiced portion, as illustrated in the example in spectrogram (5.29). Despite there being no voicing in the *d*, a voiced alveolar stop is still audible to the author. This must be attributed to the formant transition and a rapid rise in F1 when transitioning into the vowel (Lisker 1975), which contribute to the perception of voicing even when vocal fold vibration is not present.



Spectrogram 5.29: Speaker – 2 – ***kd**amim* ‘precedence (plural)’

One realization of voicing in a $[-v]_{stop}[+v]_{stop}$ cluster is reminiscent of the pattern of realization of $[+v]_{stop}[+v]_{stop}$ clusters, where a vocalic release is present. In this pattern the voiceless first member contains a visible voiced portion during the release, which was previously referred to as “vowel like” release. The portion circled in spectrogram (5.30) and marked “voiced release” is a vowel like transition from the first consonant to the second consonant. It exists in two cluster types $[+v]_{stop}[+v]_{stop}$ and in $[-v]_{stop}[+v]_{stop}$ and is speaker dependent. Some speakers exhibit this transition more frequently than others, as will become evident in table (5.18). The motivation for this voiced release, however, is different for $[-v]_{stop}[+v]_{stop}$ clusters and $[+v]_{stop}[+v]_{stop}$ clusters. While in $[+v]_{stop}[+v]_{stop}$ clusters the voicing is attributed to a voiced release of C1, in $[-v]_{stop}[+v]_{stop}$ clusters it is anticipatory voicing attributed to the anticipation of voicing in C2.



Spectrogram 5.30: Speaker – 1 – ***kd****amim* ‘precedence (pl.)’.

Although the transition in spectrogram (5.30) is a vowel like transition, it is not an epenthetic vowel. In the case of $[+v]_{stop}[+v]_{stop}$ clusters the vowel like transition can be thought of as a voiced release, or the voiced parllel of aspiration in voiceless segments. In the case of $[-v]_{stop}[+v]_{stop}$ clusters, it appears as part of the first, voiceless segment, and we do not expect to find voiced releases in voiceless segments. It is, in fact, part of the laryngeal adjustment in anticipation and preparation for producing voicing in the following voiced consonant. Much as in the case of $[+v]_{stop}[+v]_{stop}$ clusters, I argue that here too, the release portion is not a vowel since the duration of this voiced release is 30ms or less, which is too short to be a vowel. In Modern Hebrew, the duration of an unreduced vowel is around 120-140ms. and the duration of a reduced vowel is about 60ms. Thus, the voiced release, which averages 30ms across speakers and never exceeds 30ms, is too short to be a vowel. The distinction between the two cases, voiced release versus anticipatory voicing, is further supported by the nature of the release of the first stop. In spectrogram (5.30) we can clearly see that the stop is fully released with aspiration following the release. The vocalic portion follows the aspirated portion of the first stop. This is different from the voiced release in [+v][+v] cluster type, where the voiced release portion is adjacent to the actual release of the stop C1 and not separated by aspiration. Furthermore, the vocalic release in [–v][+v] cluster type is less common than the vocalic release in [+v][+v] cluster type, as is evident from the quantitative data presented in table (5.18).

5.4.1.1. Quantitative results for $[-v]_{stop}[+v]_{stop}$

Table 5.18: Voiceless stops as first member of a cluster.

C1	**p**					
	aspiration	**burst dur.**	**v. r**	**# token v.r**	**%**	**c.v**
s. 1	0.039 (0.006)	0.011 (0.003)	0.031 (0.003)	4/4	100	0
s. 2	0.006	0.016 (0.004)	0.024	1/4	25	0
s. 3	n/a	n/a	n/a	n/a	n/a	n/a
s. 4	0.016 (0.003)	0.008 (0.002)	0	0/4	0	0
s. 5	0.035 (0.005)	0.010 (0.001)	0	0/4	0	0
s. 6	0.015 (0.011)	0.014 (0.008)	0.026 (0.006)	3/4	75	0
	k					
	aspiration	**burst dur.**	**v. r**	**# token v.r**	**%**	**c.v**
s. 1	0.051 (0.006)	0.014 (0.006)	0.025 (0.002)	2/4	50	0
s. 2	0.022 (0.009)	0.024 (0.008)	0	0	0	0
s. 3	0.030 (0.012)	0.026 (0.012)	0	0/6	0	0
s. 4	0.038 (0.012)	0.01 (0.005)	0	0/7	0	0
s. 5	0.046 (0.012)	0.013 (0.007)	0.012 (no s.d)	2/7	0	0
s. 6	0.042 (0.007)	0.009 (0.0005)	0.028	1/4	25	0

In table (5.18), the presence of voiced releases is evident for speakers 1, 2, 5 and 6 but not for speakers 3 and 4. Despite the phonologically voiceless target for all segments, some of the tokens are realized with a voiced release. None of the tokens are produced with any closure voicing, as evident from the column **c.v**. in table (5.18), yet the presence of closure voicing may contribute to the perception of voicing in the first segment, despite the segments' voicelessness. Most tokens also contain an aspirated portion which is associated with voicelessness. The number of the tokens that contain a voiced release is enumerated in the column entitled **#token v.r** (number of tokens with a **v**oiced **r**elease). Not all speakers realize the voiced release a fact which lends further support to the claim that the vocalic portion is transitional. If the vocalic portion was an underlying target, we would expect speakers to be more consistent in its realization. Further support for the claim that in [–v][+v] clusters the voiced portion is associated with C2 and is anticipatory voicing, is gained from the observation that there are no cases where there is a vocalic transitional portion in C1 and C2 surfaces without voicing. To clarify, spectrogram (5.29) exemplified a case where both C1 and C2 were voiceless. C2 was realized with no closure voicing. Our claim was that this is an extreme case of delayed voicing or voicing failure and voicing is not realized until the onset of the vowel. However, a voiceless C2 never surfaces after a vocalic release. The vocalic release implies the realization of voicing in C2, as is evident in table (5.19).

For speakers 1 and 6 for *d* and for speakers 1 and 2 for *g*, voicing persists through the entire closure. Percentage of closure which is voiced always exceeds 100% implying that voicing lasts through the burst as well and may also persist through part of the release. Generally the second segment is voiced but sometimes voicing is delayed and is not initiated until the vowel, resulting in a voiceless C2. Once again, a voiceless C2, or lack of voicing, is more common with female speakers than male speakers.

An interesting anomaly is speaker 5, who, for the velar voiced stop has no voicing in any of the tokens. This is most likely related to the relatively greater difficulty in producing voicing in velar stops than in alveolar stops. There is no other speaker for whom none of the tokens of C2 contain closure voicing.

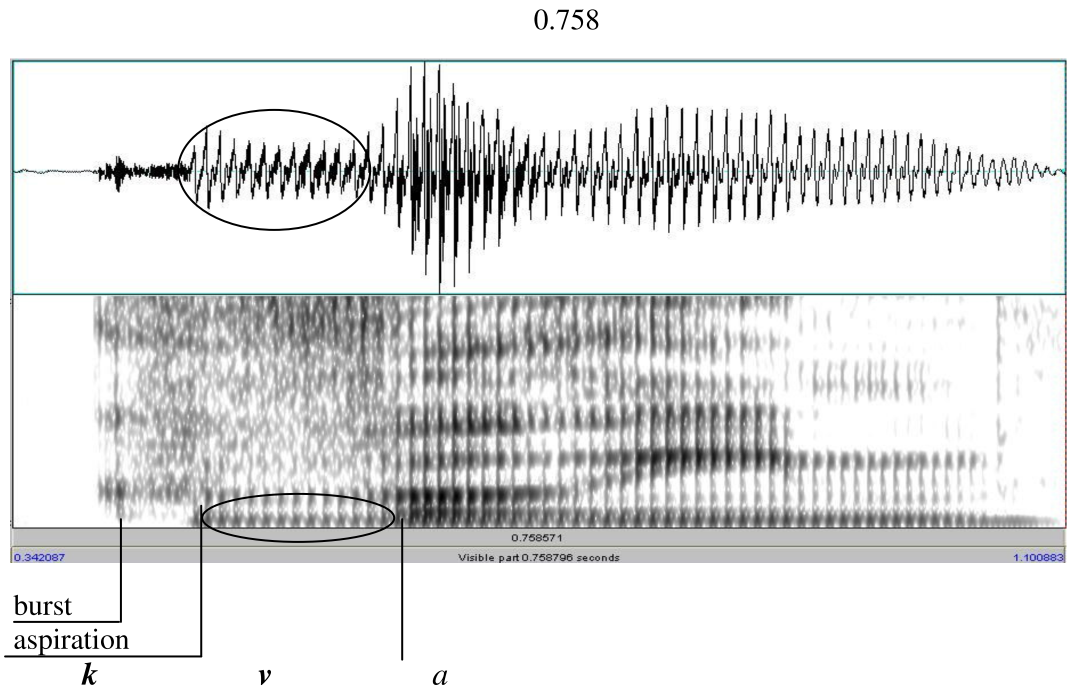
Table 5.19: Voiced stops as a second member of a cluster.

C2	**d**						
	closure voicing	**burst dur.**	**closure dur.**	**% cl. voice**	**cl. no voice.**	**release**	***f*0**
s. 1	0.107 (0.022)	0.006 (0.002)	104 (0.008)	103%	0/4	0.008 (0.004)	116 (10.7)
s. 2	0.051 (0.037)	0.012 (0.005)	0.068 (0.016)	75%	2/4	0.009	146 (22)
s. 3	0.090 (0.019)	0.013 (0.003)	0.117 (0.026)	77%	0/6	0.009 (0.004)	140 (1.5)
s. 4	0.048 (0.017)	0.008 (0.002)	0.086 (0.01)	56%	1/7	0.012 (0.004)	202 (20)
s. 5	0.039 (0.022)	0.009 (0.002)	0.088 (0.016)	44%	5/7	0.011 (0.006)	252 (14)*
s. 6	0.083 (0.033)	0.011 (0.006)	0.072 (0.010)	115%	2/4	0.015 (0.003)	244 (15)
	g						
	closure voicing	**Burst dur.**	**closure dur.**	**% cl. voice**	**cl. no voice**	**release**	***f*0**
s. 1	0.088 (0.036)	0.006 (0.002)	0.056 (0.009)	157%	0/4	0.025 (0.006)	106 (6.34)
s. 2	0.074 (0.022)	0.021 (0.006)	0.061 (0.01)	121%	0/4	0.015	129.5 (4.4)
s. 3	n/a	n/a	n/a	n/a	n/a	n/a	n/a
s. 4	0.057 (0.027)	0.015 (0.009)	0.077 (0.017)	74%	1/4	0.011 (0.006)	208 (24)
s. 5	none	0.010 (0.005)	0.082 (0.008)		4/4	0..022 (0.008)	235 (16)
s. 6	0.057 (0.006)	0.009 (0.005)	0.087 (0.006)	66%	1/4	0.02 (0.003)	249* (5)

*** all numbers for spk 5 for C2=d and spk. 6 C2=g were all obtained manually.**

5.4.2. (ii) [−v]$_{stop}$[+v]$_{fricative}$

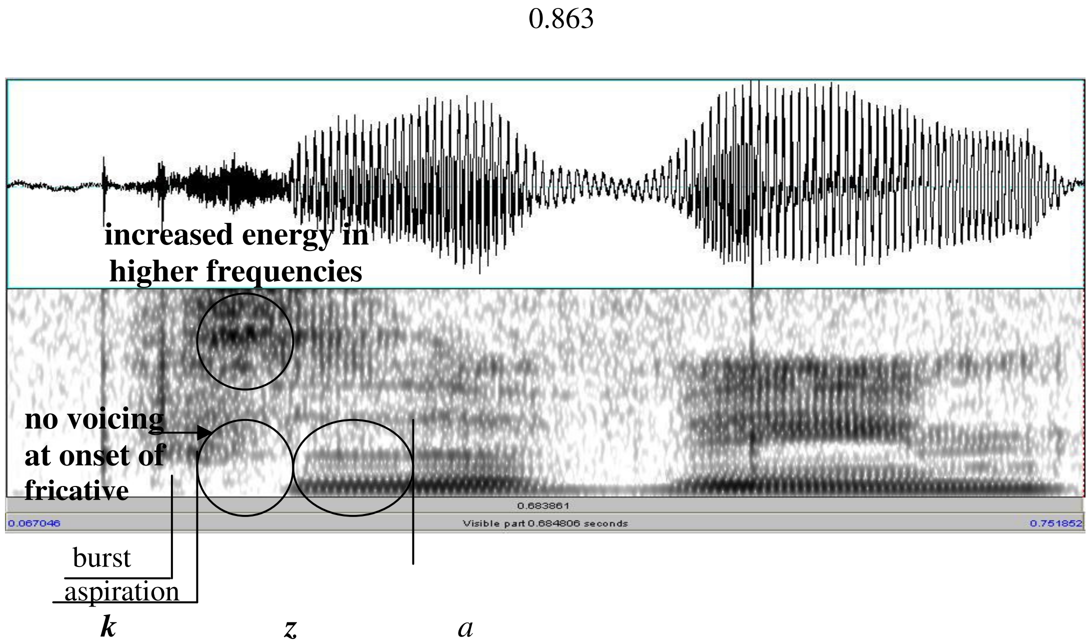
As previously discussed in this chapter, the only fricatives that may occupy a C2 position are either *z* or *v*. Therefore, the only examples of a [−v]$_{stop}$[+v]$_{fricative}$ cluster type are the clusters C1*v* or C1*z*. An example of the cluster *kv* in the word *kvarim* ‘graves’, is illustrated in spectrogram (5.31).



Spectrogram 5.31: Speaker – 1 – ***kv****arim* ‘grave (pl.)’.

Evident from spectrogram (5.31), the stop is fully released and heavily aspirated, more so than when the voiceless velar stop *k* preceded a voiceless fricative. The amount of aspiration varies between various tokens and across speakers, as will be evident by the considerable standard deviations in the qualitative section. In spectrogram (5.31) the initiation of voicing is aligned with the beginning of the fricative. However, that is not always the case.

While in C1*v* clusters, voicing aligning with the onset of the voiced fricative is the most common pattern, there are $[-v]_{stop}[+v]_{fricative}$ cases where voicing is delayed and a portion of the fricative is voiceless. In these cases voicing begins only mid-fricative rather than aligning perfectly with the onset of the fricative. This is more obvious when the second segment, is a *z*, as can be seen in spectrogram (5.32).[9]

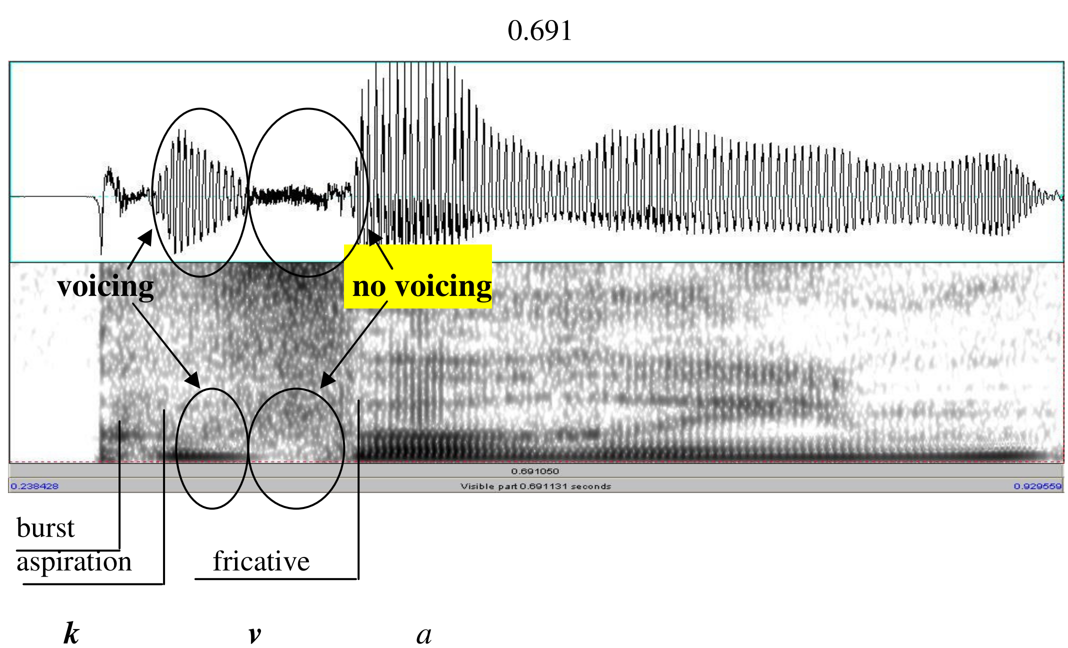


Spectrogram 5.32: Speaker – 4 – ***kz**avim* ‘lies’.

In the circled portions, the voicing of the second member of the cluster, does not begin until the middle of the fricative. Thus, voicing is delayed and the fricative is produced with only partial voicing. There is misalignment between the initiation of the fricative and the initiation of voicing. The distinction between the aspirated portion of the voiceless velar stop and the voiceless portion of the fricative is determined by audition and also by the increased energy in higher frequencies at the onset of the fricative. This pattern is quite common in these types of clusters as will be evident from the quantitative data.

[9] Possibly because for *v* there is an additional variation, producing *v* as an approximant.

The last type of realization of $[-v]_{stop}[+v]_{fricative}$ is more rare and occurs only for speakers 5 and 6. It represents voicing decay towards the end of the second fricative member of the cluster and occurs only when C2 is a fricative. In this case voicing begins at the onset of the fricative but is not maintained through the entire duration of the fricative. Rather, voicing slowly subsides towards the offset of the fricative and is initiated again during the vowel. This realization is phonetically curious since we would expect that once voicing is initiated during the fricative, is would be maintained all the way through to the vowel. After all, initiating vocal fold vibration is much more physiologically demanding than maintaining vocal fold vibration. Therefore, it is somewhat puzzling that speakers let the vibration decay only to initiate it again for he vowel. This case is illustrated in spectrogram (5.33):



Spectrogram 5.33: Speaker – 5 – ***kv**avim* 'graves'.

One explanation for this pattern may be that the differences between subglottal pressure and supraglotta pressure equalize fairly quickly after the onset of the fricative. A greater occlusion during the production of the fricative can cause a rapid pressure increase.

The logical possibility of realizing C1, the underlyingly voiceless segment, with closure voicing does not occur in any of the tokens collected by the author.

5.4.2.1. Quantitative results for $[-v]_{stop}[+v]_{fricative}$

For $[-v]_{stop}[+v]_{fric}$ clusters there is no voiced release after the stop and before the fricative as was the case for $[+v]_{stop}[+v]_{fric}$ clusters, although, there is still some aspiration in the stop segments even before the voiced fricative. These data show similarity to the release and aspiration values of voiceless stop consonants in C1 position when preceding another voiceless consonant.

Table 5.20: Voiceless stops as first members of a cluster.

C1	**k**			
	aspiration	**burst dur.**	**# token v.r**	**%**
s. 1	0.025 (0.02)	0.02 (0.008)	0/8	0
s. 2	0.023 (0.015)	0.02 (0.007)	0/8	
s. 3	0.044 (0.023)	0.017 (0.008)	0/4	0
s. 4	0.039 (0.027)	0.014 (0.01)	0/7	0
s. 5	0.046 (0.021)	0.01 (0.004)	0/8	
s. 6	0.042 (0.019)	0.133 (0.008)	0/8	0

Data for realization of voicing in C2 is presented in table (5.21). The duration of the fricative is given in the first column **fric. Dur.** The duration of voicing is given in the following column under **voice dur.** Voicing duration is always shorter than the duration of the fricative, meaning that no fricative is fully voiced. In the column **%fric**

voiced we find averages of the portion of the fricative which is voiced. The number of tokens with delayed voicing is given in the column **delay voice**. For speaker 1, 1 token of 4 for *v* and 2 tokens of 4 for *z* have delayed voicing. For the rest of the tokens, which do not have delayed voicing, voicing initiation is aligned with the onset of the fricative. For speakers 2 and 4 most of the tokens have delayed voicing.

Table 5.21: Voiceless fricative as a second member of a cluster.

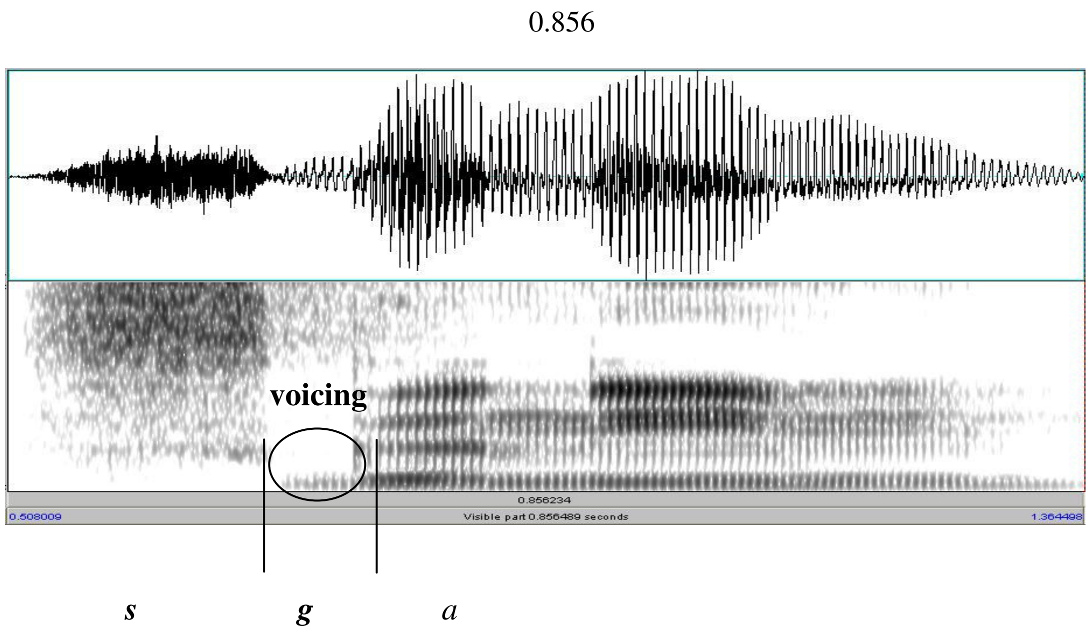
C2	**v**				
	fric. Dur.	**voice dur.**	**% fric voiced**	**delay voice**	***f*0**
s. 1	0.143 (0.02)	0.118 (0.04)	83%	1/4	105 (2.25)
s. 2	0.107 (0.022)	0.05 (0.018)	47%	4/4	127 (4)
s. 3	0.124 (0.041)	0.101 (0.008)	81%	1/4	135 (3.4)
s. 4	0.112 (0.029)	0.069 (0.028)	62%	4/4	195 (5)
s. 5	0.128 (0.043)	0.058 (0.023)	45%	1/4	221 (9)
s. 6	0.118 (0.009)	0.099 (0.036)	84%	1/4	224 (9)
	z				
	fric. dur.	**voice dur.**	**% fric voiced**	**delay voice**	***f*0**
s. 1	0.148 (0.015)	0.107[10] (0.073)	72%	2/4	106 (4)
s. 2	0.117 (0.017)	0.092 (0.036)	79%	2/4	122 (4)
s. 3	n/a	n/a	n/a	n/a	n/a
s. 4	0.123 (0.025)	0.087 (0.029)	71%	2/3	190 (7)
s. 5	0.141 (0.019)	0.043 (0.018)	30%	1/3	224 (45)
s. 6	0.148 (0.024)	0.090 (0.043)	61%	0/4	224 (7)

5.4.3. (iii) $[-\mathrm{v}]_{\mathrm{fricative}}[+\mathrm{v}]_{\mathrm{stop}}$

As in many other languages, in Modern Hebrew as well, the most common voiceless fricative in C1 position is *s*. The voiceless alveolar fricative *s* combines fairly liberally with all obstruents other than *z*. It may combine with all voiced and voiceless stops and some fricatives (*f, x, v*). The other voiceless fricatives: the labial *f* and the uvular *x*, are much more restricted in distribution in C1 position as discussed previously in chapters 1 and 3. Another voiceless fricative which may appear in C1 position is *š*, but

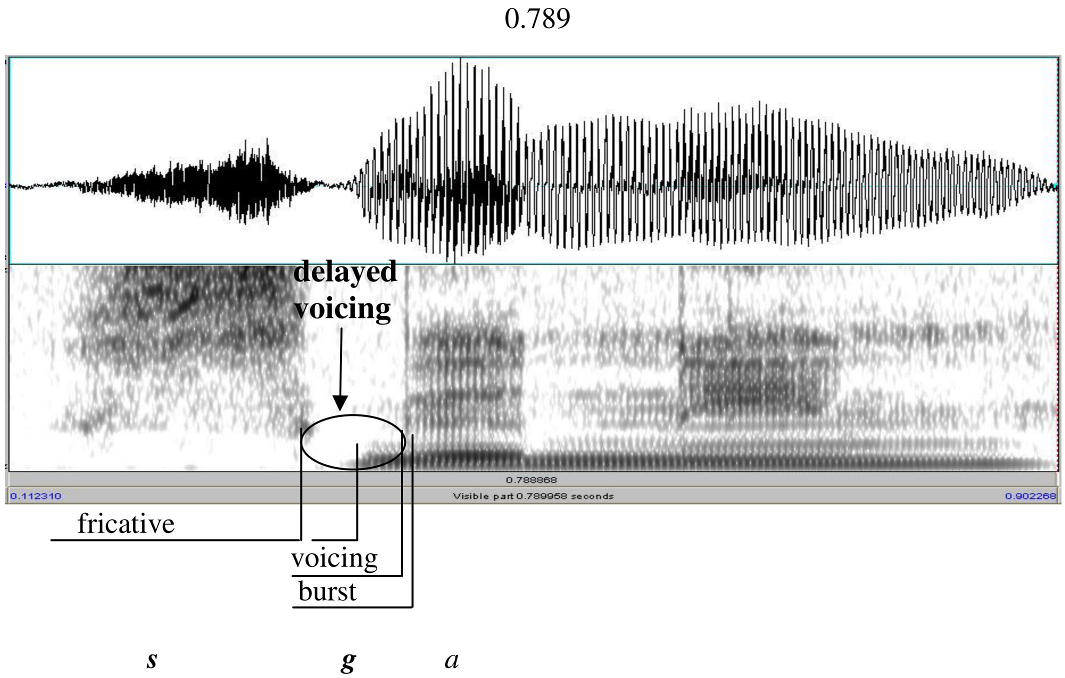
[10] Voicing in one of the tokens in this case is quite delayed and appears only 0.007 before the vowel, which results in a distortion of the standard deviations.

this fricative is not discussed here. The case presented in spectrogram (5.34) is an example of a voiceless fricative followed by a voiced stop in the cluster *sg*. As in all other cases, there are several different strategies for producing this cluster, which will be presented in subsequent spectrograms.



Spectrogram 5.34: Speaker – 3 – ***sg****anim* 'vice …'.

In the example in spectrogram (5.34), voicing is well aligned with the beginning of the closure for the second member of the cluster, the stop *g*. However, as in all other cases demonstrated prior to this, in [–v][+v] clusters voicing may be delayed and begin towards the middle of the closure as in spectrogram (5.35). As can be seen in spectrogram (5.35), voicing does not begin until the middle of the closure for the second consonant, *g*. This pattern is not uncommon, as will be evident from the quantitative data.



Spectrogram 5.35: Speaker – 4 – ***sg****anim* 'vice …'.

5.4.3.1. Quantitative results for $[-\mathrm{v}]_{\mathrm{fricative}}[+\mathrm{v}]_{\mathrm{stop}}$

The only measurement relevant for C1 in $[-\mathrm{v}]_{\mathrm{fricative}}[+\mathrm{v}]_{\mathrm{stop}}$ cluster type is the duration of the fricative *s*, which is presented in table (5.22). When compared to the duration of *s* in either C1 in a $[-\mathrm{v}]_{\mathrm{fricative}}[-\mathrm{v}]_{\mathrm{stop}}$ cluster or a $[-\mathrm{v}]_{\mathrm{fricative}}[-\mathrm{v}]_{\mathrm{fricative}}$ cluster or C2 in $[-\mathrm{v}]_{\mathrm{fricative}}[-\mathrm{v}]_{\mathrm{fricative}}$ cluster, the duration of *s* in a mixed voicing cluster is longer on average than in any of the other clusters.

Table 5.22: Voiceless fricatives as first member of a cluster.

C1	**s**
	fric. dur.
s. 1	n/a
s. 2	n/a
s. 3	0.193 (0.015)
s. 4	0.151 (0.023)
s. 5	0.189 (0.024)
s. 6	0.158 (0.009)

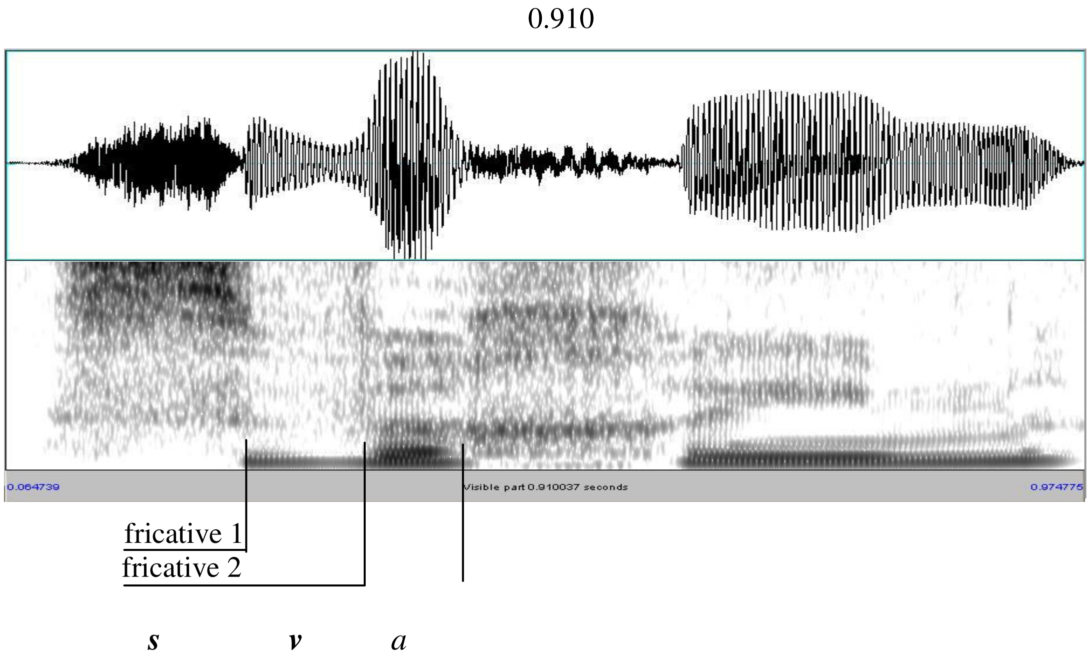
Table 5.23: Voiceless stops as a second member of a cluster.

C2	g					
	burst dur.	**closure dur.**	**cl. voicing**	**cl. no voice**	**asp.**	***f*0**
s. 1	n/a	n/a	n/a	n/a	n/a	n/a
s. 2	n/a	n/a	n/a	n/a	n/a	n/a
s. 3	0.014 (0.002)	0.066 (0.001)	0.064 (0.018)	0/3	n/a	137 (0.2)
s. 4	0.015 (0.006)	0.068 (0.01)	0.074 (0.013)	1/8	0.022 (0.004)	198 (20)
s. 5	0.017 (0.011)	0.071 (0.005)	0.022	3/4	0.017 (0.007)	242 (8)
s. 6	0.010 (0.003)	0.087 (0.015)	0.090 (0.019)	0/4	0.011 (0.007)	229 (11)

Table (5.23) lists the data for C2 of a $[-\mathrm{v}]_{\mathrm{fricative}}[-\mathrm{v}]_{\mathrm{stop}}$ cluster. The burst duration of the stop is listed under the column **burst dur.** The closure duration of the second segment is presented in the column entitled **closure dur.** followed by **cl.voicing**, which presents the data for the duration of the closure which is voiced. The column entitled **cl. no voice** indicates the number of tokens that contain no voiced portion during closure. From this column we can see that speaker 5 is exceptional in that 3 of the 4 tokens (75%) are not voiced at all which means that 3 of the 4 analyzed tokens are phonetically realized as [−v][−v]. Some tokens, although voiced, contain a small aspirated portion. When voicing decays or is considerably diminished prior to the burst, the release may contain a small portion of aspiration after the release. Note the aspiration is short compare to the aspiration in a voiceless stop and that once again standard deviations are quite large due to some tokens which contain no aspiration at all.

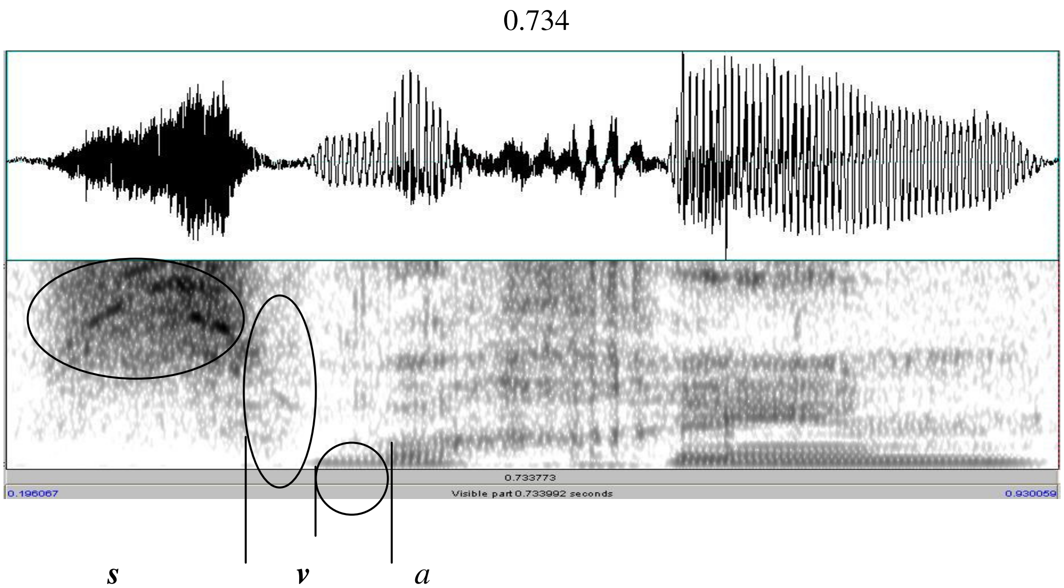
5.4.4. (iv) $[-\mathrm{v}]_{\mathrm{fricative}}[+\mathrm{v}]_{\mathrm{fricative}}$

Clustering possibilities for two fricatives are very limited. As discussed above, the only voiced fricatives that may occupy C2 are *z* and *v*. Combined with the fact that *s* and *š* are the most common voiceless fricatives in C1 position, and that Modern Hebrew does not allow two segments with the same place of articulation, clustering possibilities are reduced to a total of one. Therefore, the only $[-v]_{fricative}[+v]_{fricative}$ cluster that occurs in Modern Hebrew is *sv*, found in spectrogram (5.36):



Spectrogram 5.36: Speaker – 6 – ***sv**axim* 'entangled bushes'.

Contrast this cluster with the contrastive *zv* cluster in the minimal pair ***zv**axim* 'sacrefices' ~ ***sv**axim* 'entangled bushes.' Spectrograms (5.23-5.25) for *zvaxim* are in section 5.3.4. For most *sv* tokens voicing aligns fairly well with the beginning of the second member of the cluster. However, as in previous examples, there are a few cases of delayed voicing as exemplified below in spectrogram (5.37) and perhaps should be transcribed more appropriately as *sf^{v}axim*. The end of the fricative *s* is clear in the spectrogram. Higher concentration of energy in higher frequencies drops significantly when the labial *v* begins. However, at the onset of *v* there is still no voicing which results in part of the *v* appearing phonetically similar to *f*. The part preceding voicing initiation in *v*, indeed sounds like *f* to the author. However, once voicing begins the segment as a whole sounds like *v*. This is presumably due to additional cues for voicing such as the transition into the vowel, which are present towards the end of the segment but are missing from the onset of the segment.



Spectrogram 5.37: Speaker – 4 – ***sv****axim* ‘entangled bushes’.

5.4.4.1. Quantitative results for $[-\text{v}]_{\textbf{fricative}}[+\text{v}]_{\textbf{fricative}}$

Table (5.24) lists the duration of the first fricative member of a $[-\text{v}]_{\text{fricative}}[+\text{v}]_{\text{fricative}}$ cluster. The *s* is longer before voiced segments than before voiceless segments possibly to allow the larynx to adjust to the laryngeal configuration necessary for the following voiced segment. No such laryngeal adjustment is necessary in [−v][−v] clusters and therefore the first fricative member of a [−v][−v] cluster are shorter.

Table 5.24: Voiceless fricatives as a first member of a cluster.

C1	s
	fric. dur.
s. 1	0.174 (0.035)
s. 2	0.15 (0.027)
s. 3	n/a
s. 4	0.174 (0.018)
s. 5	0.208 (0.037)
s. 6	0.149 (0.018)

Data for C2 of a $[-v]_{fricative}[+v]_{fricative}$ cluster is presented in table (5.25). The voiced fricative *v* is shorter in duration in a $[-v]_{fricative}[+v]_{fricative}$ clusters than in $[+v]_{stop}[+v]_{fricative}$ cluster.

Table 5.25: Voiced fricative as a second member of a cluster.

C2	v				
	fric. dur.	**voice dur.**	**delayed voicing**	***f*0**	**no voicing**
s. 1	0.112 (0.008)	0.112 (0.008)	0/4	105 (2.25)	0/4
s. 2	0.093 (0.026)	0.071 (0.033)	1/4	122 (4)	0/4
s. 3	n/a	n/a	n/a	n/a	n/a
s. 4	0.071 (0.017)	0.067 (0.005)	1/4	194 (28)	2/4
s. 5	0.095 (0.022)	0.052	0/3	209 (23)	2/3
s. 6	0.102 (0.009)	0.102 (0.009)	0/4	211 (3.6)	0/4

5.4.5. Discussion

The most obvious and optimal realization of a [–v][+v] cluster is with the first member of the cluster realized with no voicing whereas voicing for the second member begins at the onset of C2 onset. For stop clusters this means that the first segment is realized as voiceless, with a possible aspirated release, and voicing is expected to being at the onset of closure for C2. This perfect alignment of voicing between C1 and C2 is illustrated in the diagram in figure (5.3). There is no misalignment between the initiation of vocal fold vibration and the closure for C2.

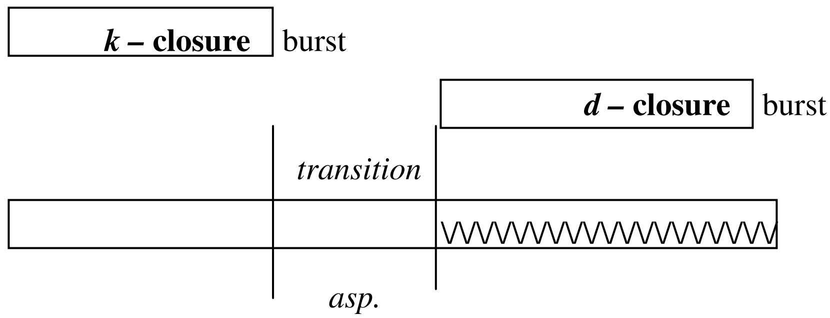


Figure 5.3: Illustration of vocal fold vibration in a $[-\mathrm{v}]_{stop}[+\mathrm{v}]_{stop}$ cluster.

However, misalignment between target laryngeal configurations and configurations of the articulators is quite common. This misalignment leads to several possible realizations of voicing in [−v][+v] clusters. One such possible misalignment in the case of $[-\mathrm{v}]_{stop}[+\mathrm{v}]_{stop}$ clusters is when vibration of vocal folds begins early, in preparation for the actual voicing of the closure of the second stop member of the cluster, resulting in what looks like a voiced release of the first stop member. Consequently, vibration of the vocal folds may begin during the release portion of the first stop and resemble a voiced release. Such a case is illustrated in figure (5.4), where part of the aspirated released portion of the *k* is voiced and voicing persists through the entire closure portion of the *d*, resulting in what looks like a voiced release with a continuous stretch of voicing following during the closure of the *d*. The diagram in figure (5.4) schematizes the case demonstrated by spectrogram (5.30).

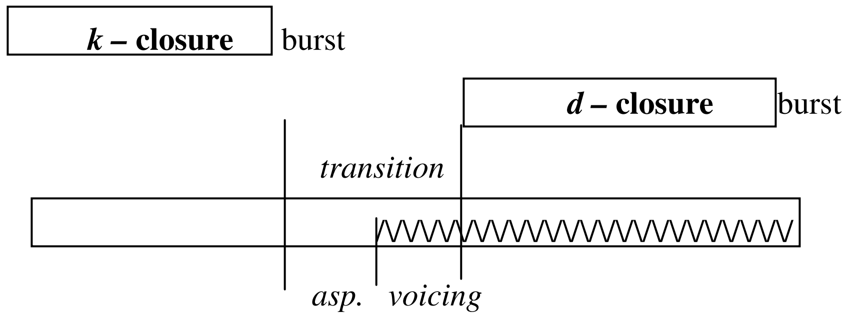


Figure 5.4: Illustration of vocal fold vibration in a $[-v]_{stop}[+v]_{stop}$ cluster.

As evident from figure (5.4), some $[-v]_{stop}[+v]_{stop}$ clusters also contain a voiced transition between C1 and C2. However, while in $[+v]_{stop}[+v]_{stop}$ clusters closure voicing must begin during the closure portion of C1, in $[-v]_{stop}[+v]_{stop}$ clusters C1 never contains any closure voicing. Although a certain voiced transition occurs between the voiceless and voiced segment in $[-v]_{stop}[+v]_{stop}$ clusters, it is caused by a laryngeal adjustment that is made in prepartion for the following voiced stop. The voiced release is caused by a misalignment between initiation of vocal fold vibration which is the target voicing specificion of C2 and the end of the release of C1. Premature muscular activity in preparation for initiation of vocal fold vibration causes an overlap between the release of C1, which usually contains an aspirated portion, and the beginning of voicing. Thus, this seemingly epenthetic vowel like portion is, in fact, a misalignment between beginning of voicing for C2 and the end of the release in C1. Since this misalignment is caused by imprecise realization of targets, it is not systematic and therefore occurs sporadically. We are only able to predict that it can occur in $[-v]_{stop}$ $[+v]_{stop}$ clusters but cannot predict which tokens will contain this vocalic portion and which will not. There is great variation in the realization of these clusters among speakers and some speakers, like speaker 1, for example, have a voiced release for most tokens (6/8=75%), while other speakers, like speaker 3 for example, do not have voiced releases for any of the tokens. Speaker 6 is an "in between" case for whom only 4/8 (50%) of tokens have a voiced release. These voiced releases do not occur in any other cluster type but $[-v]_{stop}[+v]_{stop}$ clusters.

In most cases the duration of this voiced release is 2 to 3 periods long and usually does not exceed 30 ms. Clearly it is too short to be an underlying vowel and therefore, is nothing more than a mere voiced transition. We conclude that the vocalic release is a transition, but not an epenthetic schwa, as the duration of the transitional portion is too short to be a phonological target vowel.

Another possible misalignment between initiation of voicing and consonant onset occur when voicing begins late in C2 and does not begin until the middle of the segment. The less extreme case of delayed vocal fold vibration is that in which vocal fold vibration begins in the middle of C2 as illustrated in figure (5.5). This pattern is more common in $[-v]_{stop}[+v]_{fricative}$ than in $[-v]_{stop}[+v]_{stop}$ or in $[-v]_{fricative}[+v]_{fricative}$ clusters as can be seen from the quantitative data for speaker 2, who has delayed voicing in all his $[-v]_{stop}[+v]_{fricative}$ tokens but only one token with delayed voicing in his realization of $[-v]_{fricative}[+v]_{fricative}$ clusters.

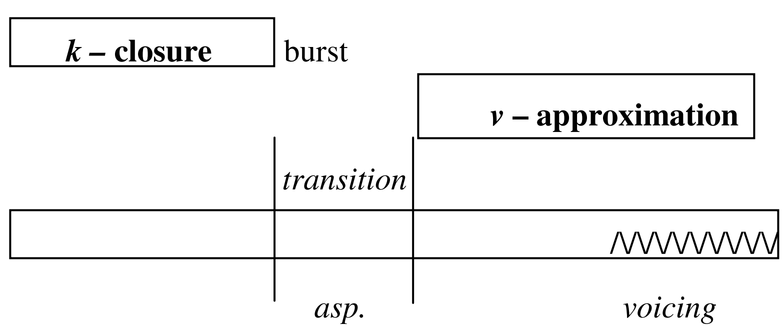


Figure 5.5: Illustration of vocal fold vibration in a $[-v]_{stop}[+v]_{fric.}$ cluster.

In the extreme case of voice delay, voicing of the second member of the cluster is delayed until the onset of the following vowel and thus no voicing is found in the second member of the cluster, as illustrated in the diagram in figure (5.6). This does occur in $[-v]_{stop}[+v]_{stop}$ clusters and is generally more common in clusters with a stop C2. This is attributed to the physiological difficulty in initiating vocal fold vibration in stops, which increases in a sequence of two stops. As a result of voicing delay, both consonants are realized as voiceless:

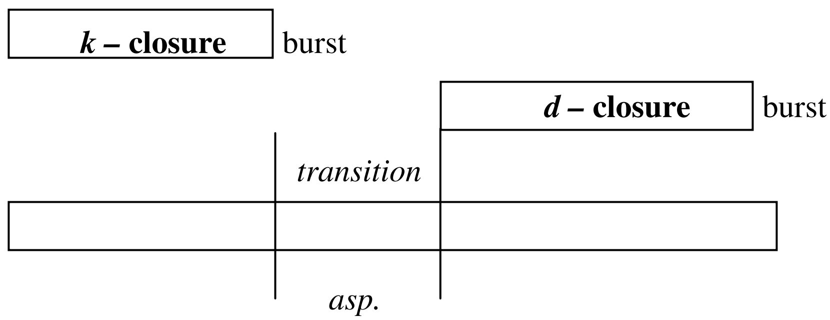


Figure 5.6: Illustration of vocal fold vibration in a $[-v]_{stop}[+v]_{stop}$ cluster.

The realization of voicing illustrated in figure (5.6) is not very common but there is speaker variation regarding how frequently it occurs. For example, for speaker 5 it is much more common with 12 out of 15 tokens (80%) with a stop in C2 position, in both $[-v]_{stop}[+v]_{stop}$ and $[-v]_{fricative}[+v]_{stop}$ clusters, being realized with no voicing in the second member of the cluster. This results in what looks like a completely voiceless cluster for 80% of the tokens which have a second stop member. This is not the case in clusters with a fricative second member, where the same speaker realizes only 4/10 of the tokens (40%) with no voicing in the fricative.

Chitoran (1999) explains her findings for Georgian [aspirated-voiced] clusters:

> Voicing control is more difficult in this [aspirated-voice] sequence, especially word-initially, where the glottis, which is wide open during C1, cannot close fast enough to start vocal fold vibration for C2. (Chitoran 1999:104)

Given the amount of aspiration we find in voiceless stops in Modern Hebrew, $[-v]_{stop}[+v]$, these clusters are reminiscent of Georgain [aspirated-voiced] clusters and therefore Chitoran's explanation may be extended to include $[-v]_{stop}[+v]$ clusters in Modern Hebrew. Thus, we are able to account for those patterns of realization in which both members of the cluster surface as voiceless. Given the compulsory release of the voiceless segment, and the difficulty in preparing the glottis for vibration, the result is lack of vocal fold vibration during both segments. The glottis simply does not close quickly enough to initiate vibration of vocal folds in the second member of the cluster.

Another case of misalignment is demonstrated by spectrogram 5.33 where voicing gradually subsides towards the end of C2. This case is very rare but occurs in $[-v]_{stop}[+v]_{fricative}$ clusters. In this case voicing initiation is aligned properly and begins at the onset of the fricative but does not last through the entire portion of the fricative. Voicing is initiated for the production of the fricative but due to physical limitation and a high pressure build up in the supra-glottal cavity, voicing slowly decays towards the end of the fricative and must be initiated again for the vowel. This case is illustrated by figure (5.7).

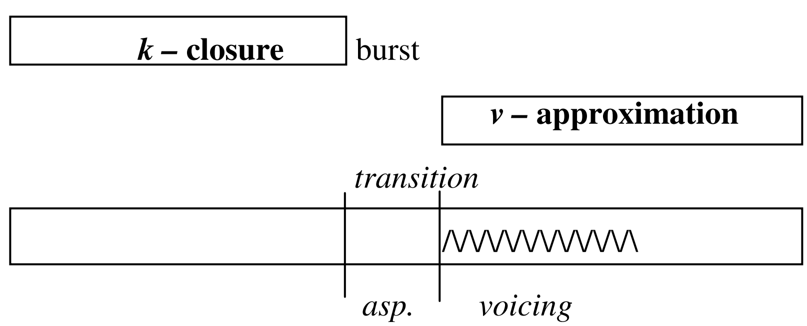


Figure 5.7: Illustration of vocal fold vibration in a $[-v]_{stop}[+v]_{fric.}$ cluster.

The diagrams in figures (5.3-5.7) show that there is quite a bit of variation in the realization of C2 in [−v][+v] clusters but very little variation in the realization of C1. Vocal fold vibration during the production of the first member of the cluster is not attested. The only voicing during C1 occurs in those cases where vibration of vocal folds begins during the release portion of C1, but never before. That is, voicing is never realized during the closure portion of C1 when C1 is a stop, or during the frication portion when C1 is a fricative. This non-existing case is diagramed in figure (5.8):[11]

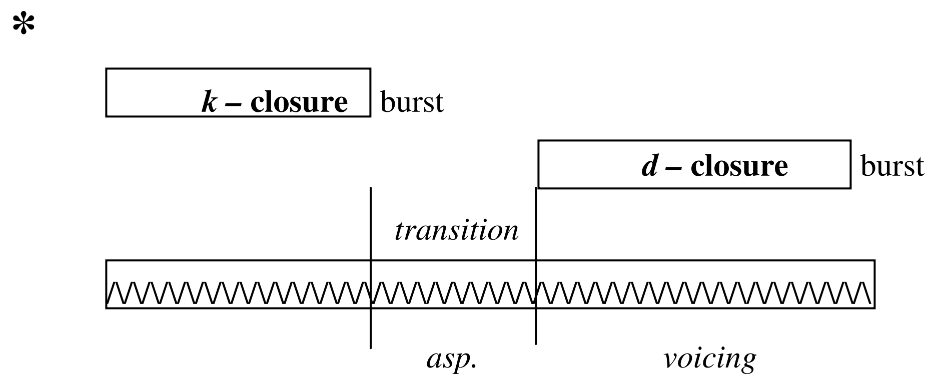


Figure 5.8: Illustration of vocal fold vibration in a $[-v]_{stop}[+v]_{stop}$ cluster.

We conclude the discussion and summarize the possible phonetic realizations for phonological /–v +v/ clusters in (8), where /–v +v/ clusters may be phonetically realized as [−v][+v] or [−v][−v] but never as *[+v][+v] or *[+v][−v].

[11] Although there is one attested token of this kind of realization, I refer to it as speaker error and neglect this token since it is the only case and does not represent a systematically occurring pattern, unlike other cases which are much more systematic across speakers.

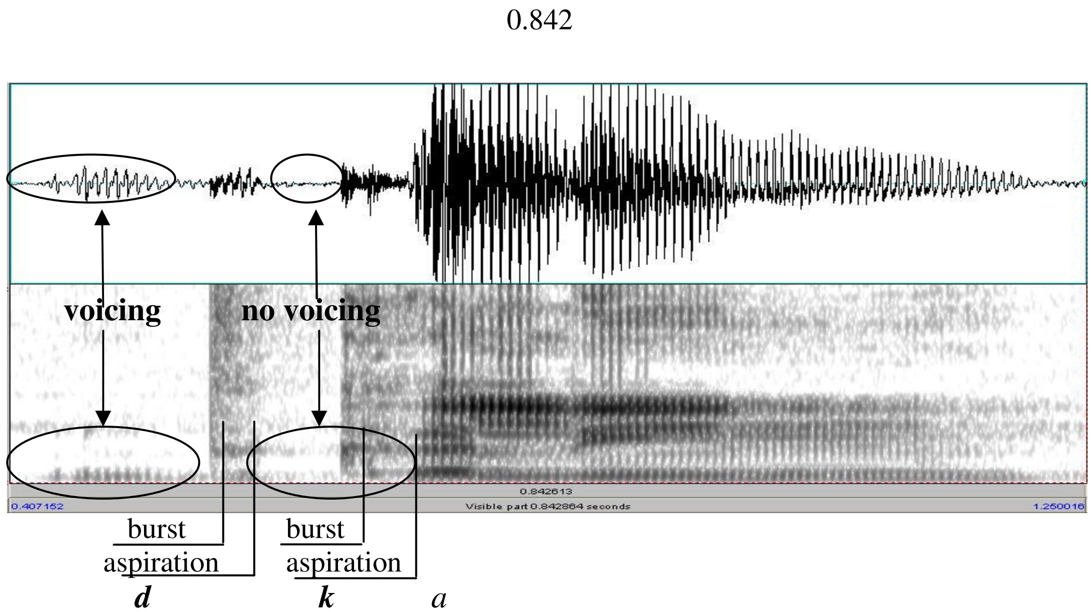
8)	Underlying target cluster		Phonetic realization
	/−v +v/	→	[−v][+v]
		→	[−v][−v]
			*[+v][+v]
			*[+v][−v]

5.5. [+v][−v] Clusters

5.5.1. (i) $[+v]_{stop}[-v]_{stop}$

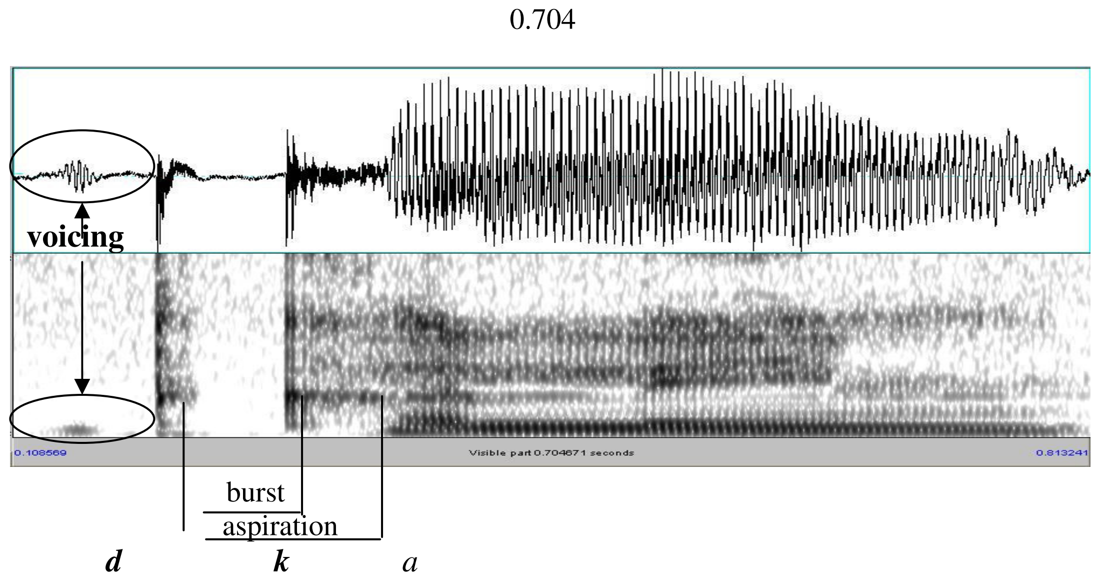
This rare cluster type, which has been extensively discussed in chapters 2 and 3 is quite varied in its phonetic realization. In this section my goal is to provide phonetic evidence for the existence of this cluster type and to counter claims in the literature that such clusters do not exist (Lindblöm 1983, Lombardi 1991).

In spectrogram (5.38), the word *dkalim* 'palm trees' illustrates that it is physiologically possible to realize voicing in the first voiced stop segment, while the second member of the cluster, the voiceless stop *k*, is phonetically realized as voiceless. Thus, voicing is realized in the first segment but not in the second. Voicing cessation is well aligned with the beginning of the closure of the second stop.

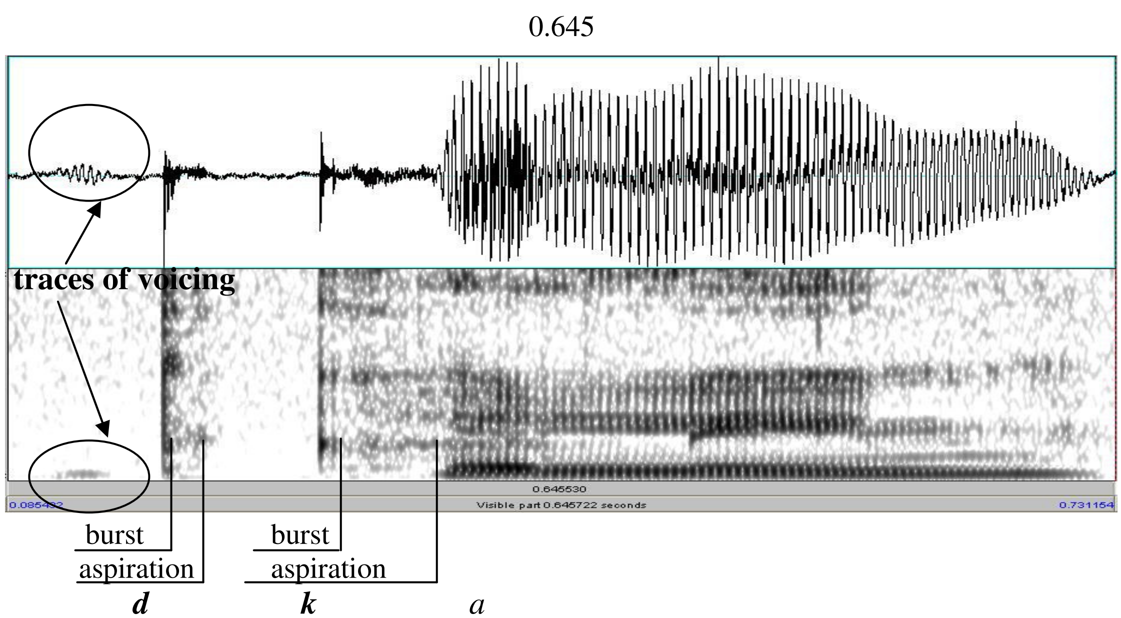


Spectrogram 5.38: Speaker – 3 – ***dk****alim* ‘palm trees’

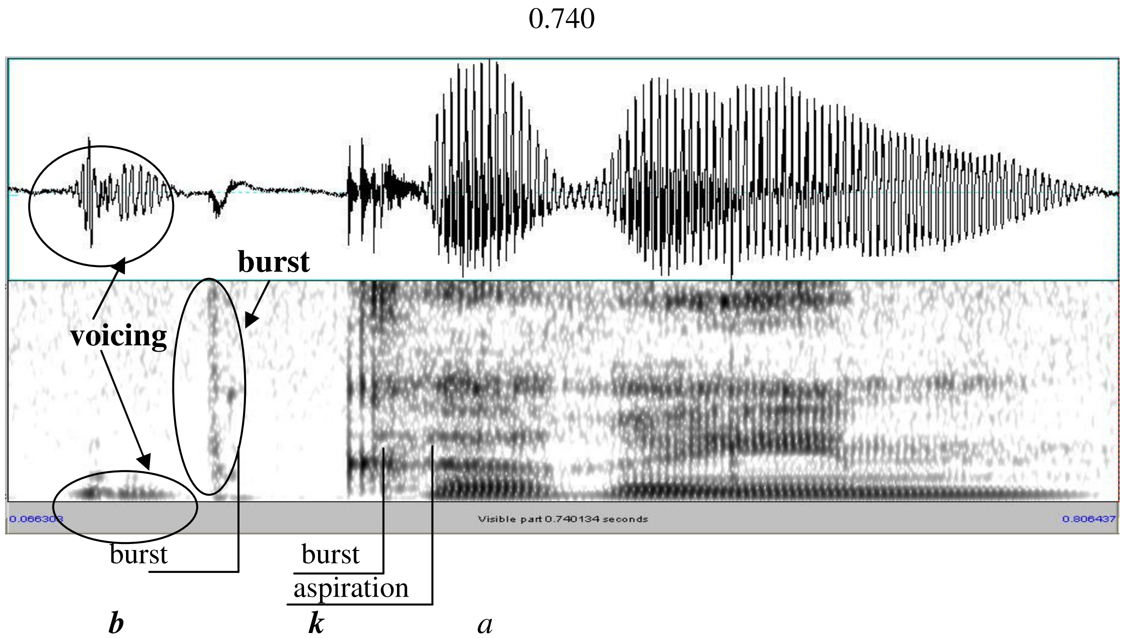
A similar pattern is found in spectrograms (5.39) and (5.40). However, in both spectrograms voicing ceases completely before the burst rather than gradually declining as it does in spectrogram (5.38). In fact, voicing is so weak and ceases quite considerably before the burst that it appears as if it “floats” freely, independently of the segment. Although voicing is present in both spectrograms (5.39) and (5.40), it is weaker in spectrogram (5.40) than in spectrogram (5.39). In both spectrograms the following segment, C2, contains no voicing. This pattern occurs in other $[+v]_{stop}[-v]_{stop}$ clusters such as *bk* as illustrated in spectrogram (5.41).



Spectrogram 5.39: Speaker – 4 – ***dk****alim* ‘palm trees’.

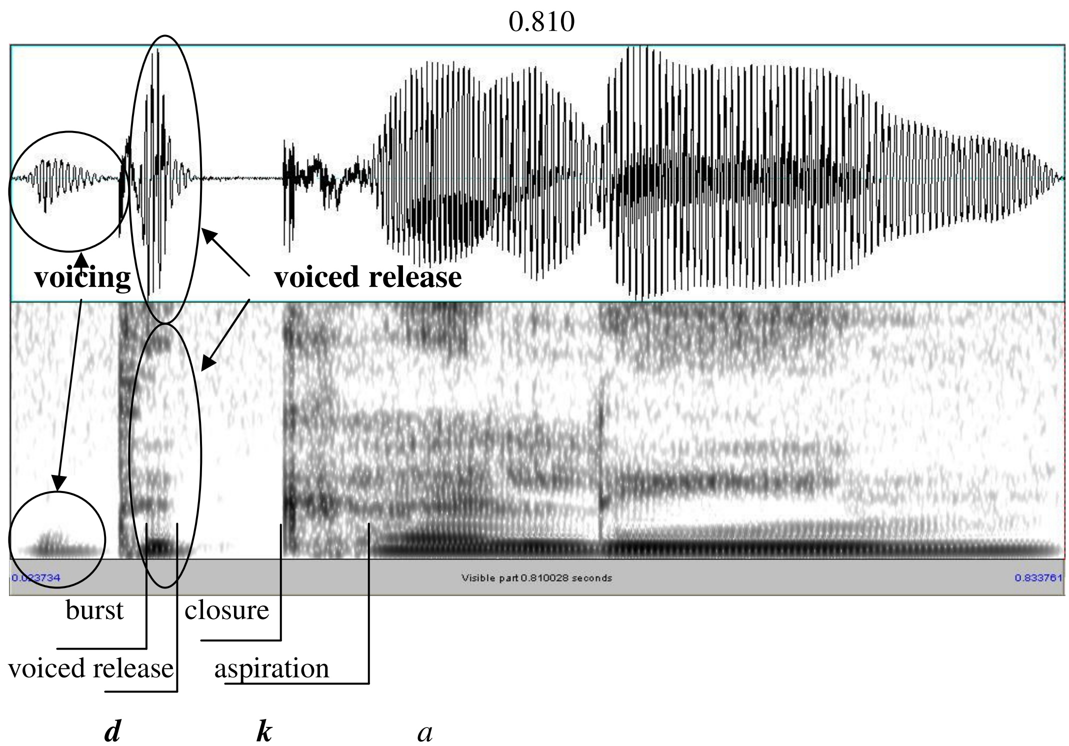


Spectrogram 5.40: Speaker – 4 – ***dk****alim* ‘palm trees’.



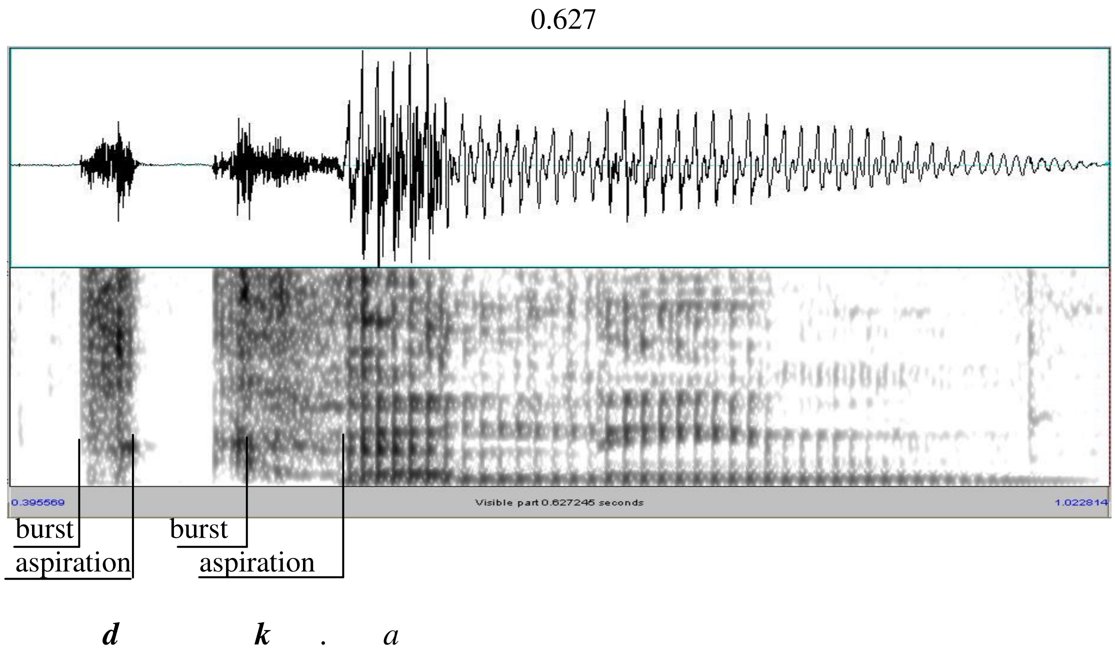
Spectrogram 5.41: Speaker – 4 – ***bk****arim* ‘morning’

We have seen in $[+v]_{stop}[+v]_{stop}$ clusters that when voicing is realized in C1, sometimes the vocal folds continue to vibrate through the burst resulting in a voiced release of the burst. For $[+v]_{stop}[+v]_{stop}$ clusters we called this voiced release a voiced transition between C1 and C2. Since the first member of the cluster $[+v]_{stop}[-v]_{stop}$ is also voiced, it is only reasonable to expect voiced releases in this cluster type as well. Vocal folds may vibrate until the speaker is able to cease their vibration and open the glottis for the voiceless segment. The voiced release, or voiced transition occurs in $[+v]_{stop}[-v]_{stop}$ clusters, albeit they are more rare in this context than in $[+v]_{stop}[+v]_{stop}$ clusters and vary from speaker to speaker, as will be evident in the quantitative data. An instance of voiced release in $[+v]_{stop}[-v]_{stop}$ clusters is demonstrated in spectrogram (5.42). In $[+v]_{stop}[-v]_{stop}$ clusters C2 is always voiceless even when following a voiced release but in $[+v]_{stop}[+v]_{stop}$ clusters C2 is always voiced when following a voiced release. This fact further supports the claim that an additional laryngeal gesture is necessary to initiate vibration in C2 in a $[+v]_{stop}[+v]_{stop}$ cluster, which does not occur in $[+v]_{stop}[-v]_{stop}$ clusters.



Spectrogram 5.42: Speaker – 6 – ***dk**alim* ‘palm trees’

Since [+v][−v] clusters are physiologically quite difficult to realize, we predict that in at least some cases voicing will not be realized in the first segments, as we predicted in chapter 4. Thus, one of the logically predicted realization of [+v][−v] clusters is without any voicing in C1, which results in neither segment containing any voiced portion. Consequently, both segments are completely voiceless as demonstrated by spectrogram (5.43):



Spectrogram 5.43: Speaker – 1 – ***dk****alim* ‘palm trees’

For $[+v]_{stop}[-v]_{stop}$ the phonetic realizations demonstrated in spectrograms (5.38)-(5.43) are the only attested realizations. There are no other phonetic variations in the realization of this cluster type.

5.5.1.1. Quantitative results for $[+v]_{stop}[-v]_{stop}$

In table (5.26) data is presented for the phonetic realization of C1 of a $[+v]_{stop}[-v]_{stop}$ cluster. In the first column data are presented for closure voicing (**c.v**). From the data it is clearly evident that speakers 2 and 5 do not realize voicing in C1 in any of their tokens. Other speakers are not consistent and may realize voicing in *b* but not in *d* or vice versa. For example, speaker 1 voices more *b* segments than *d* segments. The large standard deviations suggest that only the presence or absence of voicing, and not its duration, are significant.

Table 5.26: Voiced stops as first member of a cluster.

C1	b						
	c.v	**# c.v**	**burst dur.**	**v. r**	**# v.r**	no cv + no vr	**asp.**
s. 1	0.019	1/4	0.005 (0.001)	0.032	1/4	3/4	0.031 (0.012)
s. 2	0	0/2	0.016 (0.008)	0	0/2	4/4	n/a
s. 3	n/a	n/a	n/a	n/a	n/a	n/a	n/a
s. 4	0.064 (0.005)	6/8	0.007 (0.006)	0.026	1/8	1/8	0.02 (0.011)
s. 5	0	0/4	0.011 (0.004)	0.021	1/4	0	0.019 (0.010)
s. 6	0.059 (0.015)	3/9	0.011 (0.004)	0.025 (0)	2/9	6/9	0.016 (0.010)
	d						
	c.v	**# c.v**	**burst dur.**	**v. r**	**# v.r**	no cv + no vr	**asp.**
s. 1	0	0/4	0.006 (0.002)	0	0/4	4/4	0.027 (0.001)
s. 2	0	0/4	0.008 (0.006)	0	0/4	4/4	0.028 (0.001)
s. 3	0.088 (0.034)	3/4	0.009 (0.005)	0.029 (0.005)	2/4	1/4	0.017 (0.005)
s. 4	0.030 (0.007)	2/4	0.006 (0.002)	0	0/4	2/4	0.015 (0.006)
s. 5	0	0/4	0.009 (0.005)	0	0/4	4/4	0.033 (0.009)
s. 6	0.049	1/4	0.008 (0.003)	0.025	1/4	3/4	0.023 (0.007)

In column **#c.v** we see the number of tokens that contain a voiced portion. For speaker 1, only 1 out of 4 tokens beginning with *b* contains closure voicing, while speaker 4 voiced 6 of the 8 tokens beginning with *b*. The column **burst dur.** contains information about the duration of the burst, which, here too, contains very large standard deviations, suggesting that the burst is not a reliable cue for detecting voicing. The column **v.r**, lists the duration of the voiced release portion, which is not as common in this cluster type as in the [+v][+v] cluster type. The number of tokens with voiced releases is found in the next column **#v.r**. The column **no cv + no vr** includes information about tokens that contain neither closure voicing nor a voiced release portion. Theoretically, a voiced release can function as a secondary cue in the event that voicing is not realized. That is, speakers may choose to use a voiced release as a secondary phonetic cue to indicate a phonologically voiced target in the absence of closure voicing. This, however, does not seem to be the case. It is not the case that if closure voicing is missing speakers compensate for the lack of closure voicing with a voiced release. In the table, we see tokens that have closure voicing or voiced release or both closure voicing and voiced release or neither one. All combinations exist suggesting that voiced release is not used as a phonetic cue to distinguish phonological voicing but rather, functions as a transitional byproduct as in the $[+v]_{stop}[+v]_{stop}$ cluster type. Lastly, some segments, despite being phonologically voiced, contain a certain aspirated portion. This occurs more frequently in segments that are realized as phonetically voiceless or in voiced segments where voicing peters out prior to the burst.

Speakers vary greatly regarding the phonetic realization of these clusters. Speakers 3 and 4 voice C1 in more than half their tokens (speaker 3 voiced 3/4=75% and speaker 4 voiced 8/12=75%) while speakers 1 and 6 voice C1 only some of their tokens (speaker 1 voiced 1/8=12.5% and speaker 6 voiced 4/13=31%). Speakers 2 and 5 voice none of the stop C1 in any of the tokens collected by the author.

Table (5.27) presents data for C2 of a $[+v]_{stop}[-v]_{stop}$ cluster type. The first column in table (5.27) demonstrates that there are no cases of C2 which contain any voiced portion. The following column, **closure dur.**, provides the duration of the closure portion of C2, followed by **asp.** (aspiration)**,** which provides the length of aspiration in this voiceless stop segment. The following column **burst dur.** provides information about the duration of the burst. The large standard deviations in this column suggest, once again, that the duration of the burst is not a reliable cue for voicing. The last column ***f*0** stands for fundamental frequency, in which the standard deviation is quite small. As will be seen later, *f*0 proves to be a fairly reliable cue for voicing.

Table 5.27: Voiceless stops as a second member of a cluster.

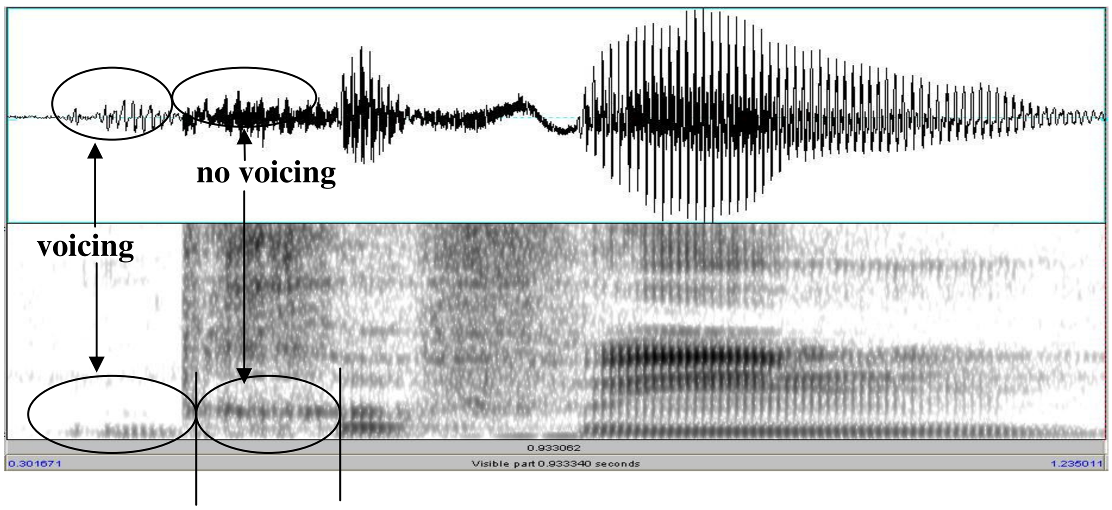
C2	k				
	c.v	**closure dur.**	**asp.**	**burst dur.**	***f*0**
s. 1		0.053 (0.011)	0.063 (0.009)	0.009 (0.006)	133 (9)
s. 2		0.053 (0.006)	0.031 (0.009)	0.019 (0.005)	141 (2)
s. 3		0.071 (0.008)	0.038 (0.012)	0.023 (0.012)	154 (11)
s. 4		0.070 (0.013)	0.037 (0.012)	0.022 (0.009)	217 (7)
s. 5		0.085 (0.011)	0.06 (0.014)	0.016 (0.009)	246 (18)
s. 6		0.080 (0.012)	0.047 (0.012)	0.011 (0.005)	246 (14)

5.5.2. (ii) $[+\mathrm{v}]_{\mathrm{stop}}[-\mathrm{v}]_{\mathrm{fricative}}$

A voiced stop followed by a voiceless fricative can be found in words like ***dx****afim* 'urges' and ***gf****anim* 'grape vines'. There are several strategies for realizing these clusters. The examples in spectrograms (5.44) for the cluster *dx* and (5.45) for the cluster *gf*, illustrate clusters produced with voicing in the first segment and no voicing in the second segment. The first voiced stop contains closure voicing which is well aligned with the onset of the second voiceless fricative. The fricative *x* is entirely voiceless. In spectrogram (5.44) the boundary between the first member of the cluster, the voiced stop *d*, and the second member of the cluster, the voiceless fricative *x*, was determined based on audition and by waveform.

Another pattern of realization often seen in voiced segments is that in which voicing slowly decreases and completely ceases before the burst. This pattern is exemplified by spectrogram (5.46) for the cluster *dx* and in spectrogram (5.47) for the cluster *gf*.

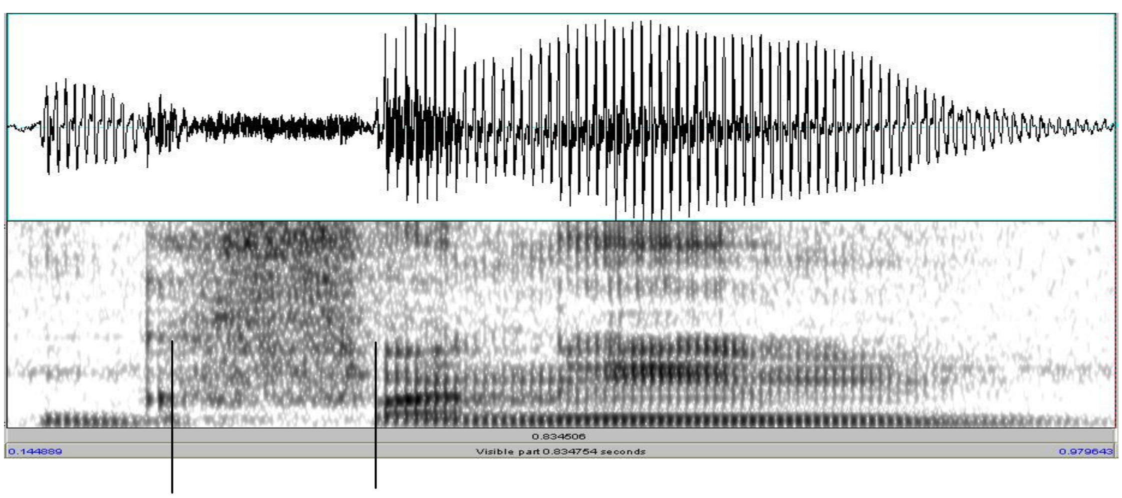
0.933



d ***x*** *a*

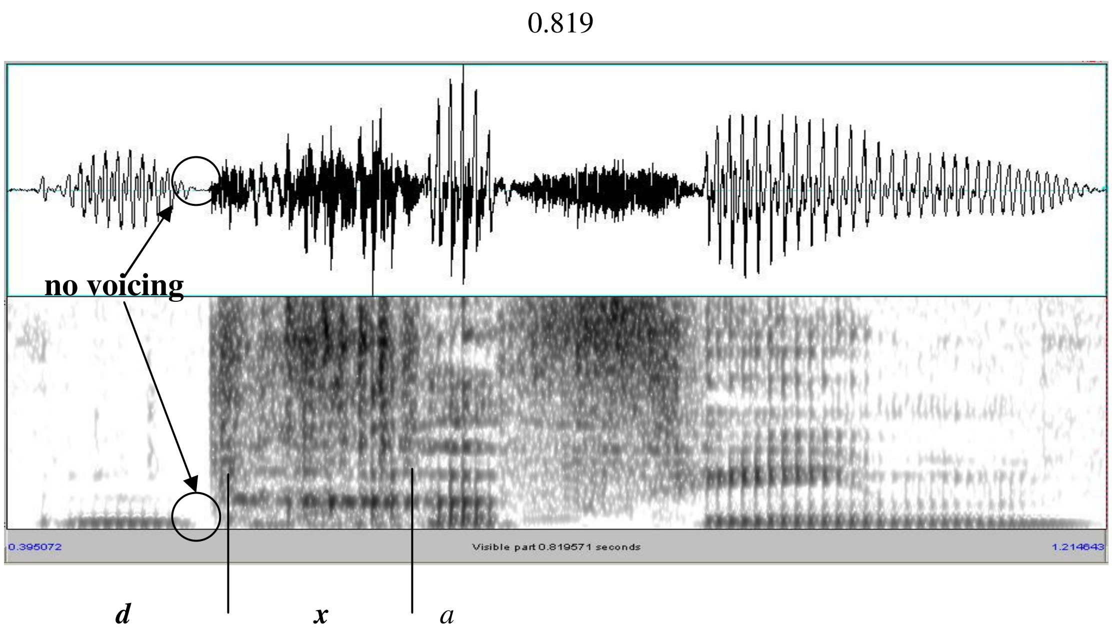
Spectrogram 5.44: Speaker – 3 – ***dx****afim* ‘urges’

0.834

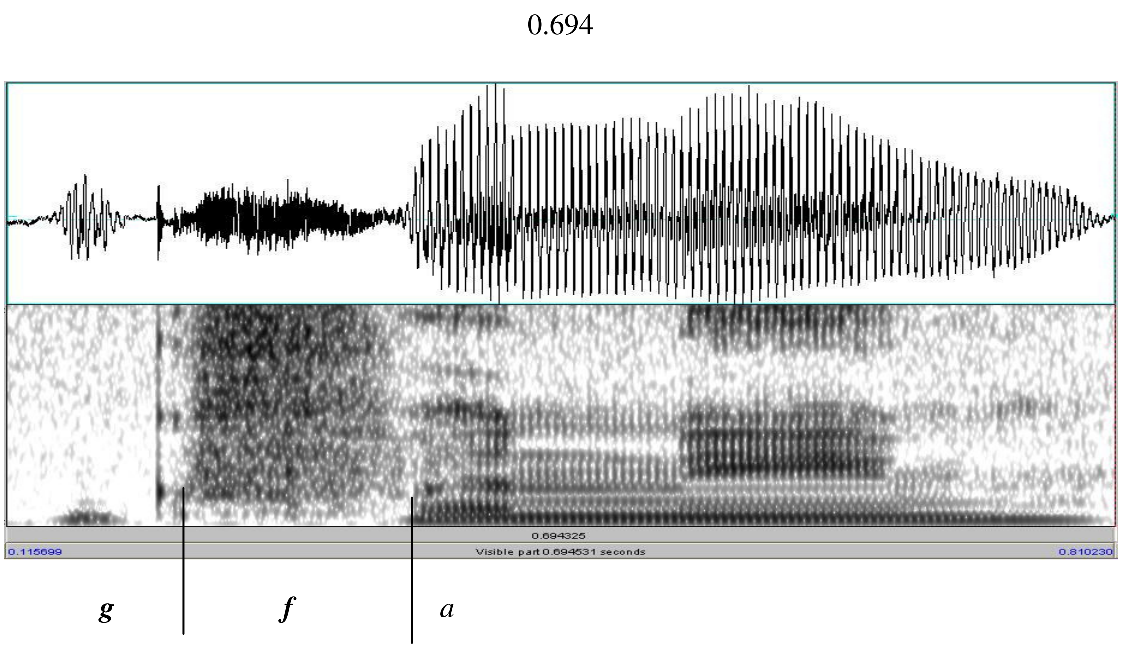


g ***f*** *a*

Spectrogram 5.45: Speaker – 3 – ***gf****anim* ‘grape vines’



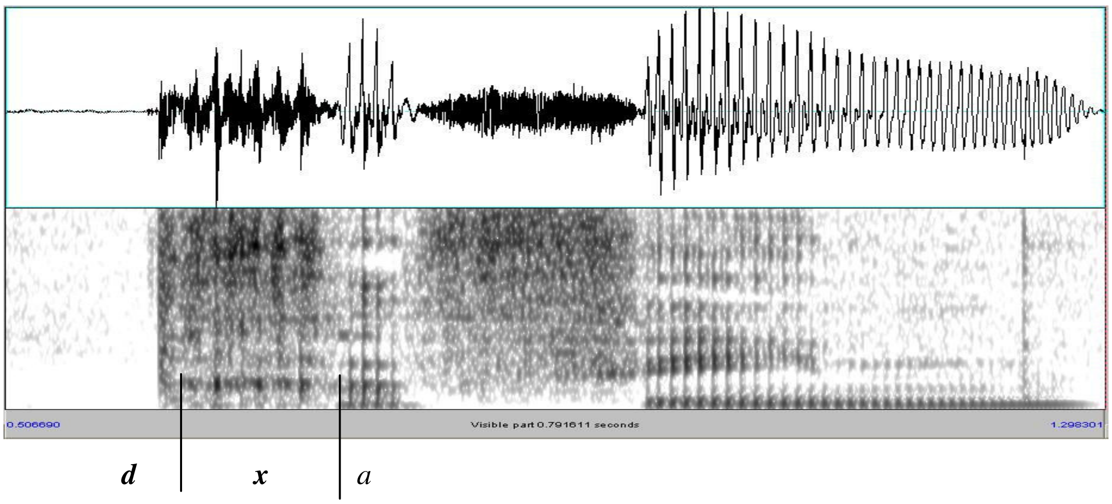
Spectrogram 5.46: Speaker – 1– ***dx****afim* ‘urges’



Spectrogram 5.47: Speaker – 4 – ***gf****anim* ‘grape vines’

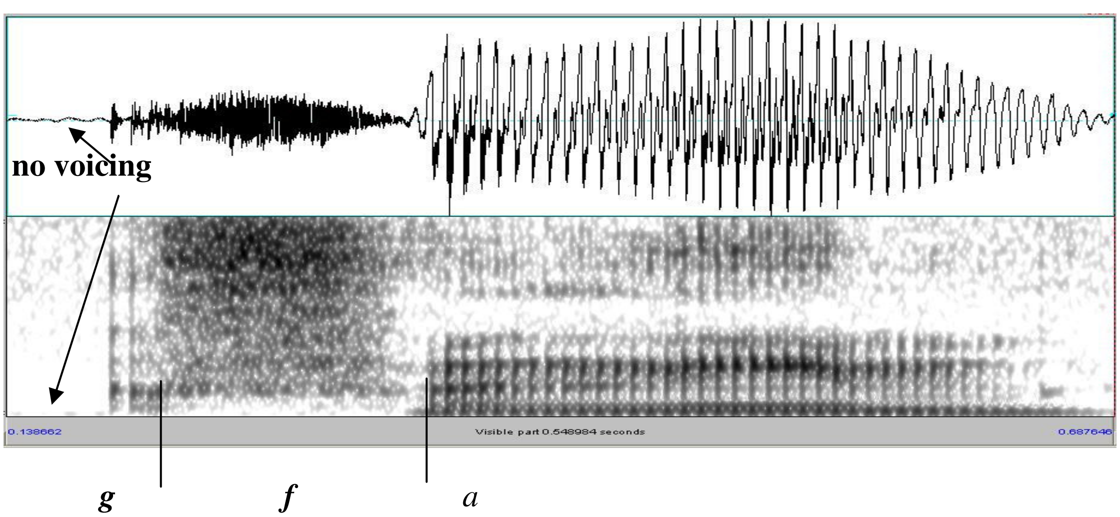
In the last example of realization of the $[+v]_{stop}[-v]_{fricative}$ cluster, is one in which both the first member of the cluster and the second member of the cluster are realized with no voicing as in spectrograms (5.48) and (5.49).

0.791



Spectrogram 5.48: Speaker – 1 – ***dx****afim* ‘urges’

0.548



Spectrogram 5.49: Speaker – 2 – ***gf****anim* ‘grape vines’

5.5.2.1. Quantitative results for $[+v]_{stop}[-v]_{fricative}$

Table (5.28) presents data for C1 in a $[+v]_{stop}[-v]_{fricative}$ cluster. In the first column, **c.v,** the duration of closure is presented. Evident from the table is that speakers 2, 5 and 6 do not voice any of their tokens, while speakers 1, 3 and 4 voice the majority of their tokens. The number of tokens which contain closure voicing is given in the column titled **# c.v** and the percentage of voiced tokens is given in the next column **%**. Burst duration and aspiration are listed in the next two columns respectively **burst dur.** and **asp.**

In $[+v]_{stop}[-v]_{fricative}$ context speaker 1 has a higher rate of voice production, 62.5% of all tokens were produced with voicing, as opposed to only 12.5% in the context of $[+v]_{stop}[-v]_{stop}$. This is not true for speaker 6 for whom 31% of C1 in $[+v]_{stop}[-v]_{stop}$ clusters but no tokens in $[+v]_{stop}[-v]_{fricative}$ clusters contained any voiced portion. Thus, it is impossible to predict which cluster contains more voiced C1, $[+v]_{stop}[-v]_{stop}$ or $[+v]_{stop}[-v]_{fricative}$, since it is entirely speaker dependent. Some speakers realize voicing more frequently in C1 in $[+v]_{stop}[-v]_{stop}$ clusters while others realize voicing more frequently in C1 in $[+v]_{stop}[-v]_{fricative}$ context. However, the duration of closure voicing is noticeably longer in $[+v]_{stop}[-v]_{fricative}$ clusters than in $[+v]_{stop}[-v]_{stop}$ clusters.

Table 5.28: Voiced stops as a first member of a cluster.

C1	d				
	c.v	**# c.v**	**%**	**burst dur.**	**asp.**
s. 1	0.091 (0.024)	3/4	75	0.015 (0.006)	0.02
s. 2	0	0/4	0	0.011 (0.008)	0.021 (0.001)
s. 3	0.103 (0.026)	4/4	100	0.015 (0.005)	0.024 (0.006)
s. 4	0.055	1/4	25	0.008 (0.005)	0.009 (0.001)
s. 5	0	0/4	0	0.008 (0.002)	0.014 (0.005)
s. 6	0	0/4	0	0.012 (0.002)	0.019 (0.007)
	g				
	c.v	**# c.v**	**%**	**burst dur.**	**asp.**
s. 1	0.037 (0.013)	2/4	50	0.018 (0.007)	0.012
s. 2	0	0/4	0	0.024 (0.005)	0.008
s. 3	0.072 (0.019)	3/4	75	0.012 (0.004)	0.012 (0.003)
s. 4	0.047 (0.009)	3/4	75	0.014 (0.007)	No asp.
s. 5	0	0/4	0	0.018 (0.011)	0.047 (0.034)
s. 6	0	0/4	0	0.014 (0.012)	0.014 (0.016)

In $[+v]_{stop}[-v]_{fricative}$ clusters, as in other [+v][−v] clusters, none of the C2 segments contain any voiced portion. None of the tokens are realized with closure voicing as evident from the column **voicing dur**.

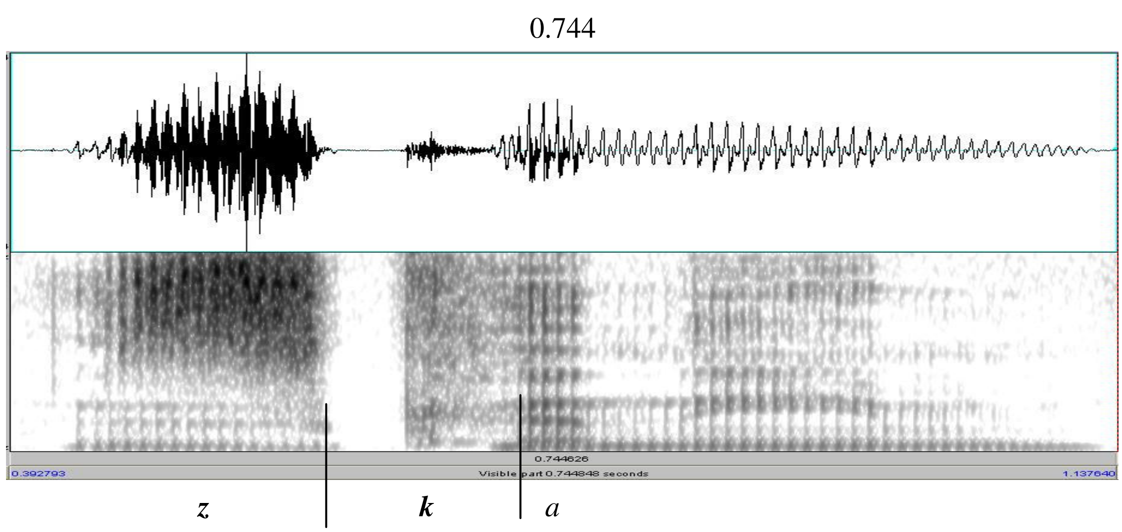
Once again, the pitch tracker proves to be unreliable in this context and fails to retrieve the fundamental frequency (***f*0**) values for speaker 3. These were obtained manually wherever possible. Nonetheless, and not surprisingly, *f*0 is higher in this context, after a voiceless segment than after voiced segments, as was the case in both [+v][+v] and [–v][+v] clusters.

Table 5.29: Voiceless fricatives as a second member of a cluster.

C2	f		
	voicing dur.	**fric. dur.**	***f*0**
s. 1	0	0.189 (0.014)	134 (0.45)
s. 2	0	0.119 (0.008)	143 (7)
s. 3	0	0.157 (0.011)	136 (3.75)
s. 4	0	0.147 (0.005)	226 (10.8)
s. 5	0	0.172 (0.019)	271 (24)
s. 6	0	0.169 (0.013)	236 (6)
	x		
	voicing dur.	**fric. dur.**	***f*0**
s. 1	0	0.125 (0.010)	116 (12)
s. 2	0	0.090 (0.001)	139 (10)
s. 3	0	0.115 (0.007)	152 (0.25)
s. 4	0	0.136 (0.018)	214 (16)
s. 5	0	0.118 (0.012)	260 (30)
s. 6	0	0.106 (0.004)	227 (9)

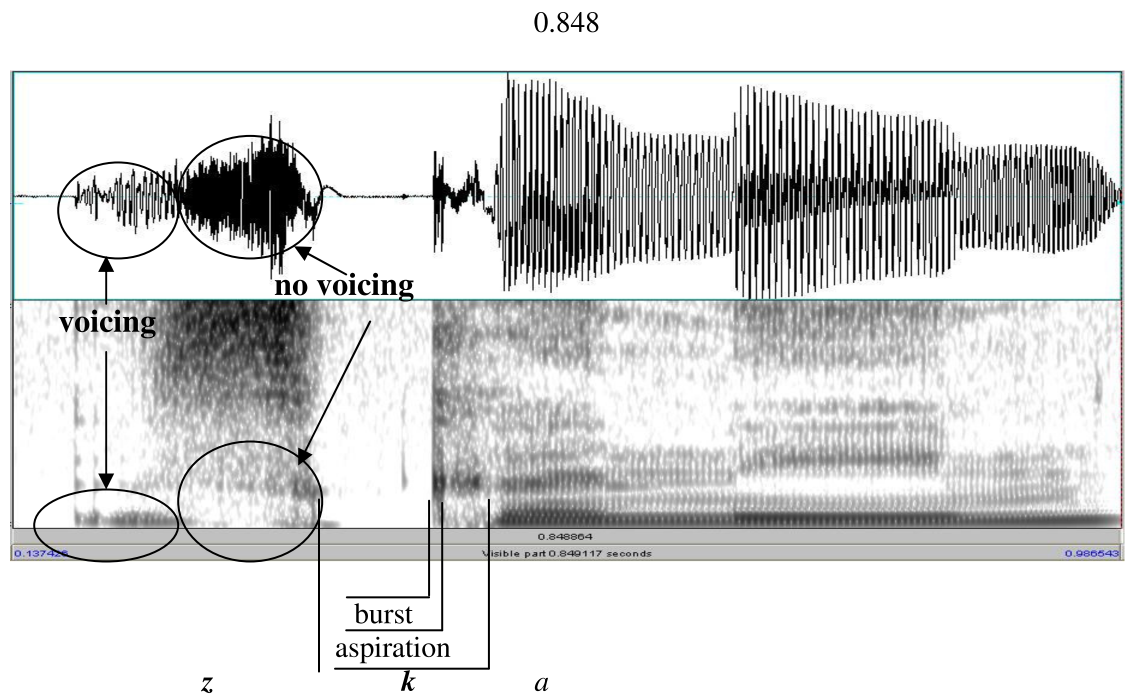
5.5.3. (iii) $[+v]_{fricative}[-v]_{stop}$

A voiced fricative preceding a voiceless stop occurs in words like *zkanim* 'beards'. In spectrogram (5.50) the fricative *z* is voiced through the entire duration of the fricative. Voicing ceases at the onset of closure for the following voiceless stop, *k*.

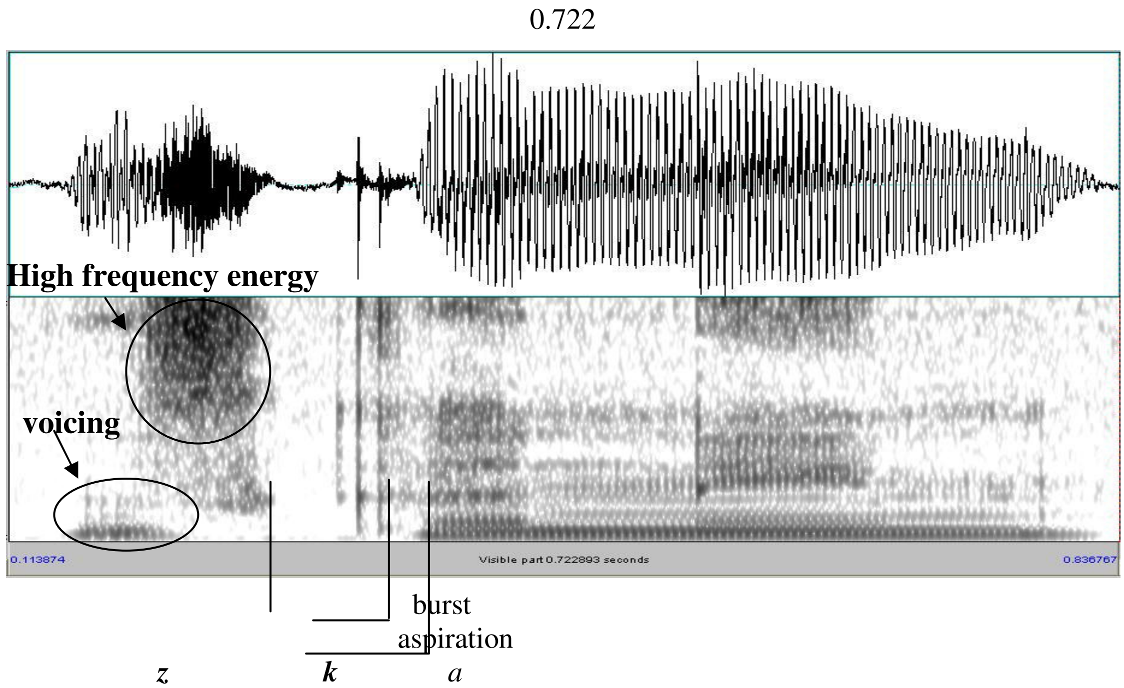


Spectrogram 5.50: Speaker – 1 – ***zk****anim* 'beards'

In spectrogram (5.51), the fricative is partially voiced. Voicing is initiated in the beginning of the fricative and slowly decreases towards the offset of the fricative. In the circled portion in second half of the fricative there is an energy increase visible in higher frequencies but no voicing. The increased energy in higher frequencies implies that the second portion of the fricative is indeed voiceless. The second part of the fricative resembles the voiceless alveolar fricative *s* while the first part of the fricative appears closer to a voiced alveolar fricative *z*. This pattern is more prominent in the spectrogram in (5.52), where the beginning of the fricative contains very little energy at high frequencies but quickly changes to resemble a robust *s* with considerable amounts of energy in higher frequencies.

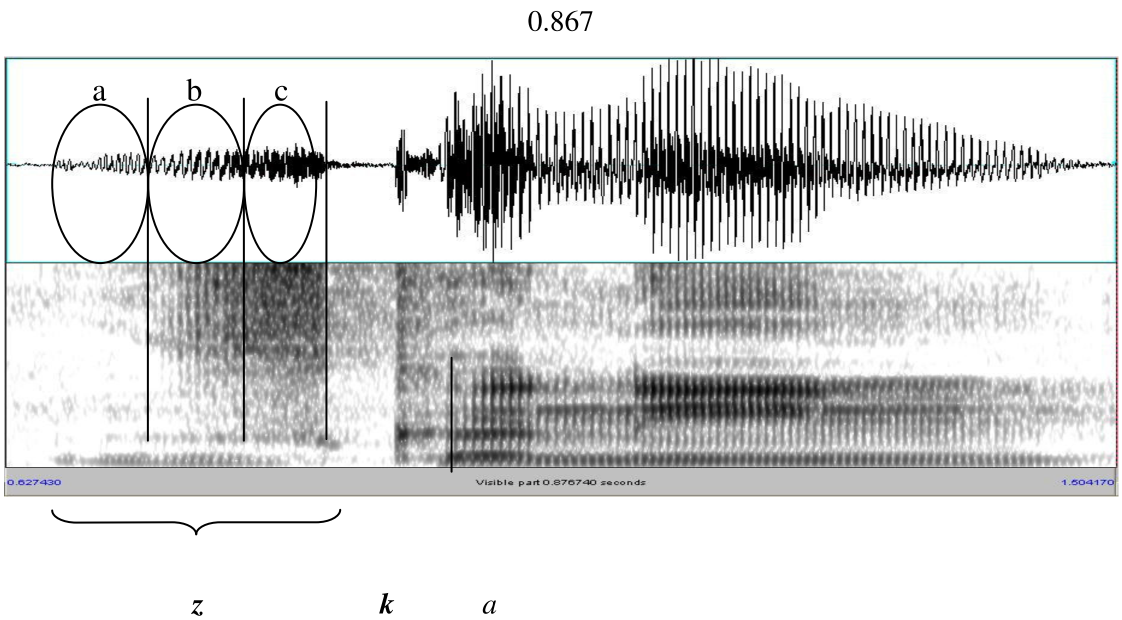


Spectrogram 5.51: Speaker – 6 – ***zk****anim* ‘beards’



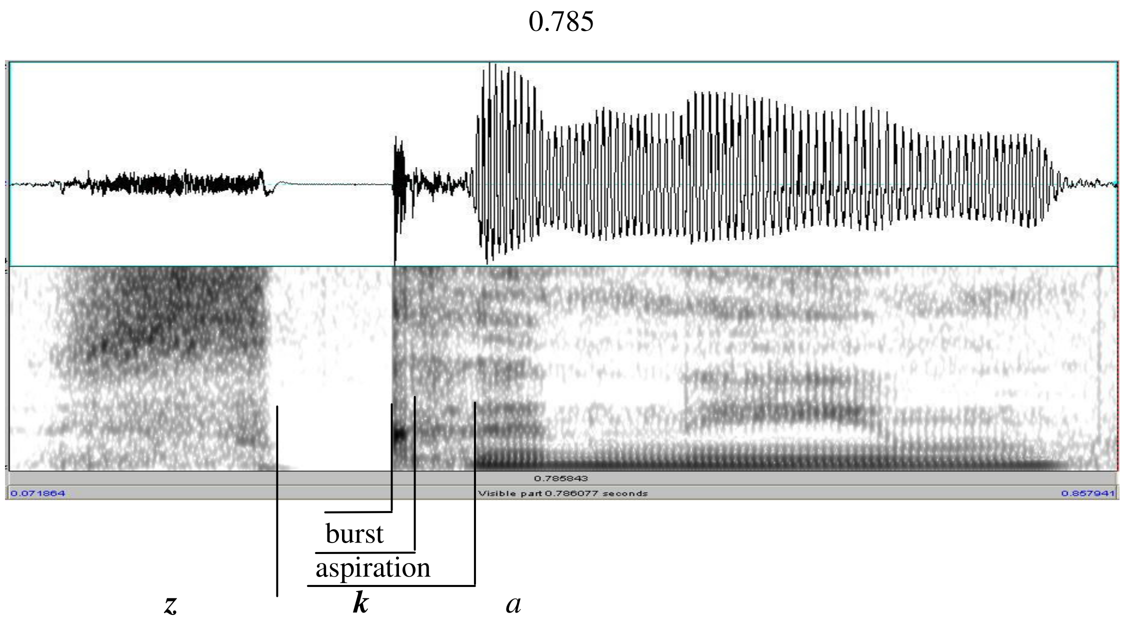
Spectrogram 5.52: Speaker – 4 – ***zk****anim* ‘beards’

In spectrogram (5.52) only the beginning of the fricative is voiced. The second portion of the fricative contains a high concentration of energy in higher frequencies but contains no voiced portion in the form of a voice bar in the lower frequencies. This pattern is somewhat similar to, though slightly different from, the pattern illustrated in spectrogram (5.53), where the fricative appears to have three parts: (a) leading voicing (b) a voiced fricative (c) a voiceless fricative. At the onset of the fricative there is no energy at higher frequencies but only a voice bar. In the second portion of the fricative there are both components: high frequency noise and a voice bar. In the last portion of the fricative the voice bar almost disappears but the high frequency energy is increased and the waveform exhibits irregular pulses. The three stages of the fricative illustrate the transitions that occur during the production of the fricative. The segment sounds to the author as the voiced alveolar fricative *z*, despite the three different components evident during the production of this single segment, suggesting that the three parts of the fricative are not crucial for the perception of voicing. A voiced portion of the fricative, whether larger or smaller, is sufficient for perceiving the entire segment as voiced.



Spectrogram 5.53: Speaker – 3 – ***zk****anim* ‘beards’

One of the more common realizations of the $[+v]_{fricative}[-v]_{stop}$ cluster, much like other [+v][−v] clusters, is one in which both segments are voiceless as in spectrogram (5.54). It should be noted that although the first member of the cluster, the fricative *z*, is realized with no voicing, there is not a complete neutralization with the voiceless alveolar fricative *s*. The amplitude of the surface voiceless *z* is significantly lower than that of the surface *s*, resulting in a slightly different energy pattern. In other words, the *z* in spectrogram (5.54) is produced with significantly lower energy in high frequencies than the voiceless alveolar fricative *s*, as evidenced in spectrograms (5.34), (5.35) and (5.36), where *s* appears to have considerably greater amplitude. This suggests that despite the lack of voicing in *z*, it is not completely neutralized with *s*. Rather, this pattern suggests failure in realization of voicing and not a categorical change from the voiced *z* to the voiceless *s*.



Spectrogram 5.54: Speaker – 5 – ***zk****anim* 'beards'

5.5.3.1. Quantitative results for $[+v]_{fricative}[-v]_{stop}$

Table (5.30) lists data for C1 in a $[+v]_{fricative}[-v]_{stop}$. The first column **voicing dur.** presents the duration of voicing in the fricative, while the second column **fric. dur.,** lists the duration of the fricative. The column entitled **tokens vcd.** lists the number of tokens that are voiced or contain a voiced portion. For speaker 1, 4 tokens out of 4 tokens are voiced, which means that all his tokens in the *zk* cluster are realized with some voicing. The amount of voicing is listed in the next column, **%v. of fric.,** which stands for the percentage of the fricative which is voiced. The percentage is calculated by dividing the duration of the voiced portion of the fricative by the duration of the entire fricative (=**voicing dur./fric.dur.**). This column can be misleading since the column listing voicing duration includes tokens which are voiceless, causing a distortion of the averages. For example, for speaker 6 results suggest that 47% of the fricatives are voiced; that is, on average 47% of each fricative is voiced. However, a closer inspection reveals that two tokens contain no voicing while the remaining two tokens contain 44% and 40% voicing respectively. However, since the column lists averages large standard deviations are evident. The last column lists the number of tokens which are fully voiced.

Table 5.30: Voiceless fricatives as first member of a cluster.

C1	**z**				
	voicing dur.	**fric. dur.**	**tokens vcd**	**% v. of fric**	**# fric. fully v.**
s. 1	0.141 (0.033)	0.166 (0.012)	4/4	85	2/4
s. 2	0	0.134 (0.009)	0/2	0	0/2
s. 3	0.116 (0.056)	0.175 (0.036)	3/3	79	0/3
s. 4	0.065 (0.010)	0.141 (0.009)	4/4	52	0/4
s. 5	0	0.169 (0.016)	0/4	0	0/4
s. 6	0.081 (0.005)	0.172 (0.028)	2/4	47	0/4

Speakers 1, 3 and 4 voice all their tokens but this is not the case for speakers 2, 5 and 6. Speakers 2 and 5 voice none of their tokens while speaker 6 voices only 50% of her tokens. For speakers 1, 3, 4 and 6, for those fricatives where voicing is realized (2/4 tokens for speaker 1, 3/3 tokens for speaker 3, 4/4 tokens for speaker 4 and 2/4 tokens for speaker 6), it is realized with only partial voicing, which increasingly decays towards the offset of the fricative. That is, in most cases voicing onset is aligned with fricative onset and decreases towards the end of the fricative. Only 2 out of 4 tokens for speaker 1 are realized with 100% voicing. There are no cases of what was referred to as "delayed voicing". That is, there are no tokens where voicing begins well after the onset of the fricative. This result is to be expected considering the fact that the following C2 segment is voiceless and laryngeal configuration for the ensuing voicless C2, needs to be achieved. Next we examine the production of C2, which is never voiced.

Table 5.31: Voiceless stop as second member of a cluster.

C2	**k**				
	c.v	**closure dur.**	**asp.**	**burst dur.**	***f*0**
s. 1	0	0.064 (0.012)	0.046 (0.011)	0.018 (0.005)	147 (14)
s. 2	0	0.055 (0.011)	0.020 (0.004)	0.013 (0.003)	125 (2)
s. 3	0	0.057 (0.005)	0.030 (0.005)	0.014 (0.003)	154
s. 4	0	0.060 (0.017)	0.030 (0.014)	0.025 (0.010)	214 (9)
s. 5	0	0.081 (0.008)	0.041 (0.006)	0.014 (0.007)	243
s. 6	0	0.092 (0.005)	0.040 (0.005)	0.009 (0.002)	237 (7)

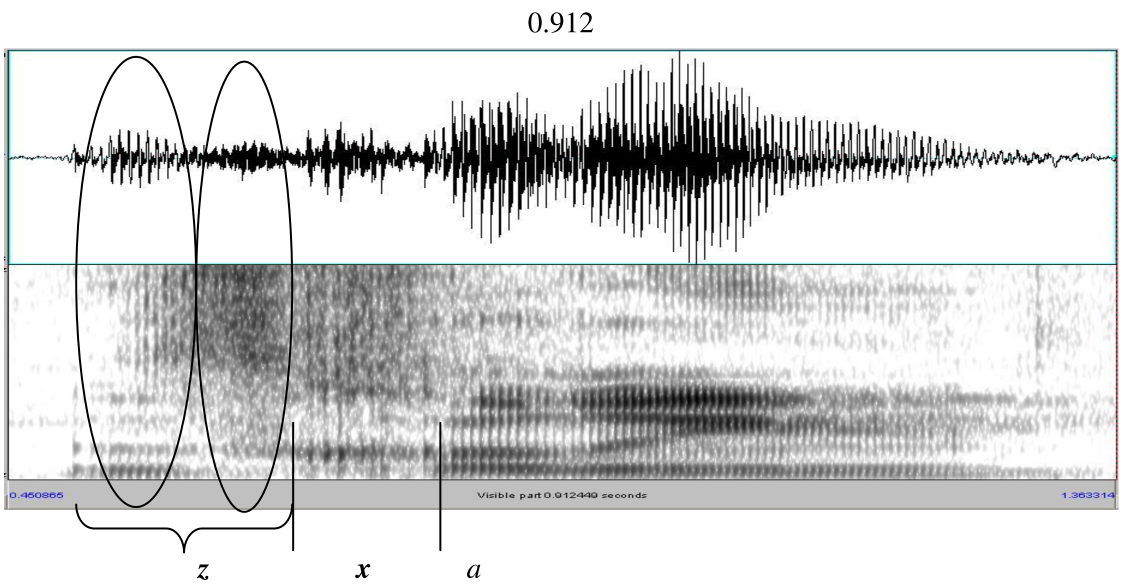
Similarly to other clusters where *k* is in C2 position, in this cluster (***zk***), *k* is also aspirated, although much more so for speakers 1, 5 and 6 than for speakers 2, 3 and 4. As in previous instances, here too, *f*0 is higher after this voiceless stop than after voiced stops. The closure duration, however, is not too different from that of *k* following a voiceless fricative.

5.5.4. (iv) $[+v]_{fricative}[-v]_{fricative}$

Fricative voiced segments in C1 position are limited to *z*, the only fricative that may appear in C1 position. This fact limits the fricative segments that may occupy C2 position to *x*[12] resulting in *zx* being the only $[+v]_{fric.}[-v]_{fric.}$ that occurs in Modern Hebrew. Below is a spectrogram of the word *zxarim* 'males'. There are no instances where voicing in the first segment *z* is perfectly aligned with both the onset and the

[12] Recall that *f* is very rare in C2 position and **zs* clusters are prohibited.

offset of the fricative. That is, cases in which voicing begins at the beginning of the fricative and ends at the end of it do not exist. The alignment between voicing and frication is never perfect and in all instances where there is voicing, it decays towards the end of the fricative.

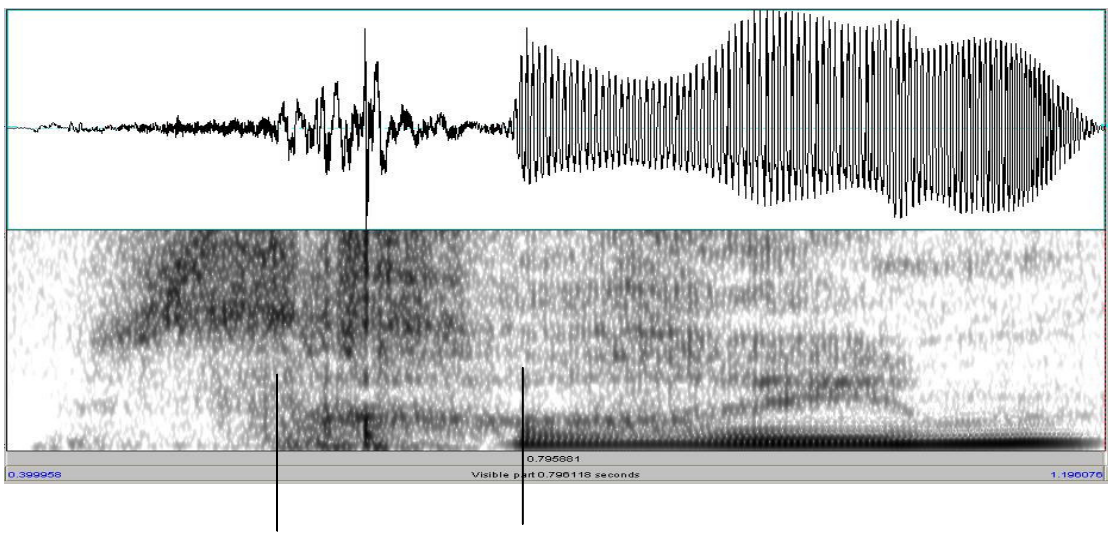


Spectrogram 5.55: Speaker – 3 – ***zx****arim* 'males'

The two circled portions emphasize the changes within the fricative *z* from the beginning of the segment to the end of it. The most notable change during the segment is in voicing, which gradually decays towards the end of the fricative as the concentration of energy in higher frequencies in the second portion of the segment increases. Considering the ensuing voiceless fricative this pattern is not surprising. Voicing decays in anticipation of the voiceless fricative that follows the first voiced fricative.

The second pattern of realization of voicing in $[+v]_{\text{fricative}}[-v]_{\text{fricative}}$ is without voicing in either segment, as demonstrated by spectrogram (5.56), where the voiced alveolar fricative *z* is realized with no voicing. However, the underlyingly voiced fricative *z*, which surfaces as phonetically voiceless, is still acoustically distinct from the phonologically voiceless alveolar fricative *s* in that the phonetically voiceless *z* in this example is produced with considerably lower amplitude than its phonologically voiceless counterpart *s*.[13]

0.795



z ***x*** *a*

Spectrogram 5.56: Speaker – 5 – ***zx****arim* 'males'

5.5.4.1. Quantitative results for $[+v]_{fricative}[-v]_{fricative}$

In this section we present data for $[+v]_{fricative}[-v]_{fricative}$ for C1 in table (5.32) and for C2 in table (5.33).

[13] Spectrograms for *s* preceding a voiceless fricative can be found in section 5.2.4 in spectrogram (5.11).

Table 5.32: Voiced fricatives as a first member of a cluster.

C1	z			
	voicing dur.	**fric. dur.**	**token vcd**	**% v. of fric**
s. 1	n/a	n/a	n/a	n/a
s. 2	n/a	n/a	n/a	n/a
s. 3	0.080 (0.019)	0.161 (0.016)	4/4	50
s. 4	0.048	0.096 (0.013)	1/5	50
s. 5	0	0.154 (0.038)	0/4	0
s. 6	0	0.140 (0.020)	0/4	0

Results show that while speaker 3 voices all his tokens, none of the tokens are fully voiced. That is, although all tokens contain a portion which is voiced, all tokens are only partially voiced with voicing decreasing towards the offset of the fricative. Speaker 4 realized voicing in only one token and speakers 5 and 6 have only voiceless tokens.

Table 5.33: Voiceless fricatives as a second member of a cluster.

C2	x		
	voicing dur.	**fric. dur.**	***f*0**
s. 1	n/a	n/a	n/a
s. 2	n/a	n/a	n/a
s. 3	0	0.106 (0.014)	144 (13)
s. 4	0	0.141 (0.022)	221 (13)
s. 5	0	0.146 (0.023)	252 (14)
s. 6	0	0.136 (0.011)	225 (7)

When comparing table (5.33) with table (5.7), which lists the duration of voiceless fricatives when following a stop, we see that *x* is slightly longer after another fricative than when following a stop. The fundamental frequency, *f*0 is higher in this context of voiceless segment, than it is after voiced segments.

5.5.5. Discussion

In the case of [+v][–v] clusters, voicing is not always realized phonetically in the voiced C1 even though it is the phonological target. Sometimes vocal fold vibration fails due to unfavorable physical and physiological conditions, as discussed in chapter 4.

Results suggest that manner of articulation does influence the realization of voicing in clusters, however, the way in which it influences the realization of voicing is speaker dependent and varies greatly across speakers. In the following table the number in each column represents the percentage of tokens where C1 is voiced in [+v][–v] clusters. Notice that generally, for all 6 speakers clusters that begin with fricatives have a higher percentage of voiced C1s than those clusters that begin with a stop.

Table 5.34: Percentage of voiced C1 in various voicing combinations.

C1	$[+v]_{stop}[-v]_{stop}$	$[+v]_{stop}[-v]_{fric}$	$[+v]_{fric}[-v]_{stop}$	$[+v]_{fric}[-v]_{fric}$
Sp. 1	12.5%	62.5%	100%	n/a
Sp. 2	0%	0%	0%	0%
Sp. 3	75%	100%	100%	100%
Sp. 4	67%	50%	100%	20%
Sp. 5	0%	0%	0%	0%
Sp. 6	31%	0%	50%	0%

Voicing is realized differently in the different clusters types. In $[+v]_{stop}[-v]$ clusters voicing is realized as closure voicing, while in $[+v]_{fric}[-v]$ clusters it is vibration of vocal folds during the production of the fricative that makes C1 voiced.

Let us begin our discussion with $[+v]_{stop}[-v]$ clusters where vibration of vocal folds during closure is the only cue for voicing in the context of absolute initial position. Formant transitions, which usually provide information about the place of articulation of the segment and are present when the stop is followed by a vowel, are not present in this context, where the stop is followed by another consonant (Stevens 1998). Moreover, since the cluster is realized in utterance initial position, cues which are usually available for detection of voicing at the onset of the segment such as duration of preceding vowel are missing. Of all the cues listed in table (4.1) in chapter 4, repeated below in table (5.35) for the readers' convenience, only closure voicing is available in the context of utterance initial position, preceding a voiceless stop (context #8 in table (5.35)):

Table 5.35: Voicing cues in various word positions.

	Internal cues		**Offset cues**		**Onset cues**		
Context	**closure voicing**	**closure duration**	**S_1 duration**	**F1 values in S_1**	**Burst duration/ amplitdue**	**VOT**	**F0 and F1 transitions in S_2**
1. S_1 _ S_2	✓	✓	✓	✓	✓	✓	✓
2. O _ S_2	✓	✓			✓	✓	✓
3. # _ S_2	✓				✓	✓	✓
4. S_1 _ #	✓	✓	✓	✓	(✓)		
5. S_1 _ O	✓	✓	✓	✓			
6. O _ O	✓	✓					
7. O _ #	✓	✓					
8. # _ O	✓				[✓]		

Evident from the table is the fact that the only cue available for voicing is the duration of voicing during closure. Failing to produce voicing in this context, will result in the impossibility of the listener to recover the voicing specification of the segment.

Let us now focus on $[+v]_{stop}[-v]_{stop}$ clusters where closure voicing becomes crucial for the recovery of voicing specifications of C1. An optimal case where initiation of vocal fold vibration is perfectly aligned, and only the first segment has any vocal fold vibration realized, is illustrated in figure (5.9):

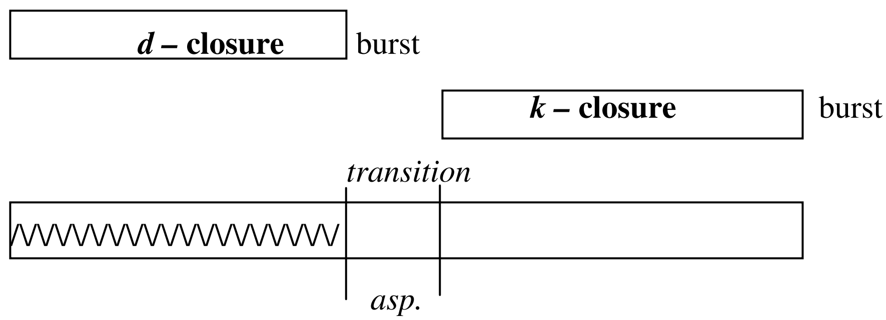


Figure 5.9: Illustration of vocal fold vibration in a $[+v]_{stop}[-v]_{stop}$ cluster.

In this case only the first member of the cluster has any vocal fold vibration and the second member of the cluster has no vocal fold vibration. However, if vocal folds do not cease vibrating in time, this may result in part of the transition being voiced, as in figure (5.10). In this case a voice released is detected after the burst of the stop C1 but voicing does not proceed into the closure of the following voiceless segment. The fact that a voiced release never occurs without the presence of closure voicing in C1 is evidence that the voiced release should be segmented with the first segment C1 and not with C2. The voiced release is realized when voicing continues through the burst of the first segment. C2 is always voiceless and never produced with any vocal fold vibration.

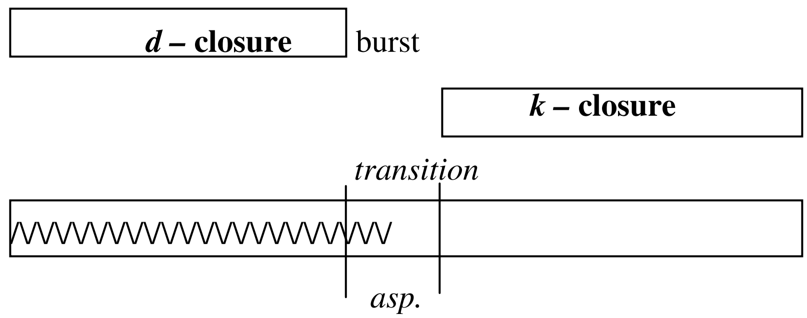


Figure 5.10: Illustration of vocal fold vibration in a $[+v]_{stop}[-v]_{stop}$ cluster.

Interestingly, the conditions for the voiced release in $[+v]_{stop}[-v]_{stop}$ clusters are different from those necessary for $[-v]_{stop}[+v]_{stop}$ clusters. In $[+v]_{stop}[-v]_{stop}$ clusters the voiced release is always correlated with closure voicing in the first segment C1. Conversely, the exact opposite is true for $[-v]_{stop}[+v]_{stop}$ clusters, where the voiced release is independent of closure voicing in C1, but is dependent on voicing in C2. That is, in $[+v]_{stop}[-v]_{stop}$ clusters closure voicing occurs only with voicing in C1 as opposed to $[-v]_{stop}[+v]_{stop}$ clusters where voicing is correlated with voicing in C2. This suggests that although vocalic releases occur in both cluster types, $[-v]_{stop}[+v]_{stop}$ and $[+v]_{stop}[-v]_{stop}$ and they function as a transition in both cases, they are, in fact, byproducts of different phonetic conditions.

In $[+v]_{stop}[-v]_{stop}$ clusters vocalic releases are a consequence of realization of voicing in C1, and are segmented with the first segment. They are conditioned by the presence of closure voicing in C1 and function as the offset of voicing of C1. It is the transitional period of vocal folds that are in a vibrating state prior to the burst and vibration does not cease before the movement of the articulators to the next target segment.

In $[-v]_{stop}[+v]_{stop}$ clusters vocalic transitions are initiation of vocal fold vibration for the voiced target which follows the voiceless segment and therefore, are a consequence of anticipatory voicing. They are conditioned by the presence of voicing in C2 and function as the onset to C2.

However, as seen throughout section 5.5 voicing fails to be realized in C1 in [+v][–v] clusters frequently due to unfavorable conditions in word initial position, as illustrated in figure (5.11):

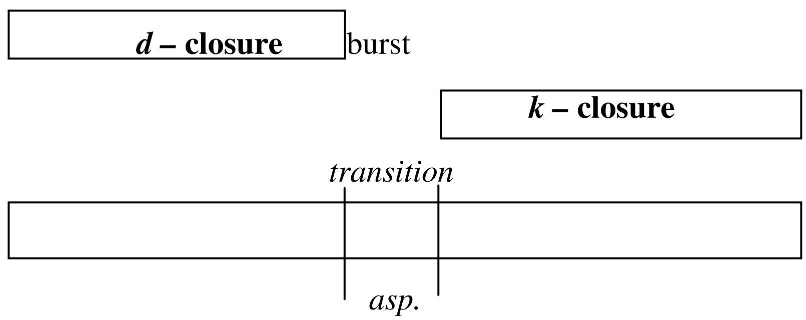


Figure 5.11: Illustration of vocal fold vibration in a $[+v]_{stop}[-v]_{stop}$ cluster.

As argued in chapter 4 and in this chapter this case is not a case of voicing assimilation where the first segment which has an underlying voiced target is devoiced, or that its categorical specification is altered from voiced to voiceless. Rather this is a case where voicing fails to occur due to insufficient pressure build up in the sub-glottal area, despite an underlying voiced target. This results in both C1 and C2 surfacing as phonetically voiceless. Although both segments are realized as voiceless phonetically, they have different phonological targets. The underlying phonological target of C1 is /+v/ whereas the underlying phonological target for C2 is /–v/. As we observed in section 5.3 of this chapter, even in voiced clusters, where both members of the cluster are phonologically specified for /+v/, C1 can be realized with no phonetic voicing, because voicing fails to be realized in the context of utterance initial position.

Cases that never occur are illustrated in figures (5.12) and (5.13). Voicing is never realized only during the transition period and then ceases as in (5.12). That is, voicing is never present in any part of the cluster if it is not realized in C1.

*

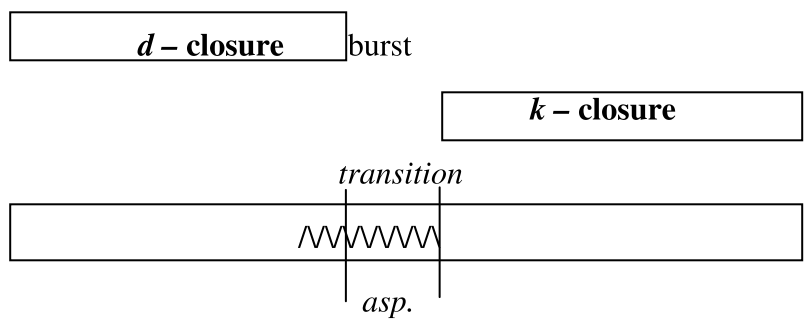


Figure 5.12: Illustration of vocal fold vibration in a $[+v]_{stop}[-v]_{stop}$ cluster.

There are no cases attested with voicing in C2, regardless of whether voicing is present in C1 or not, as in figure (5.13):

*

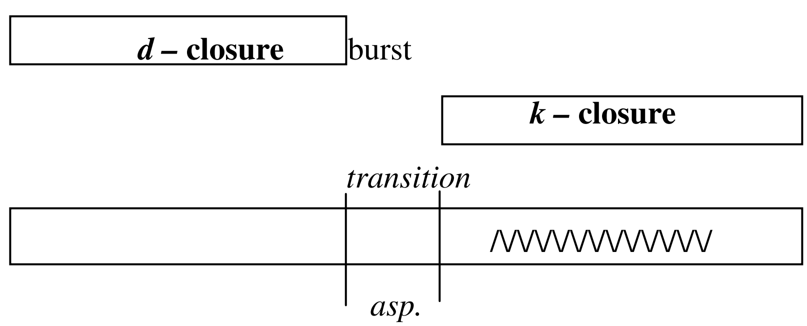


Figure 5.13: Illustration of vocal fold vibration in a $[+v]_{stop}[-v]_{stop}$ cluster.

Next we address $[+v]_{stop}[-v]_{fric}$ clusters in which voiced releases do not occur. All segments are realized with no voiced release despite the presentece of closure voicing in C1. The optimal realization of $[+v]_{stop}[-v]_{fric}$ is illustrated in (5.14):

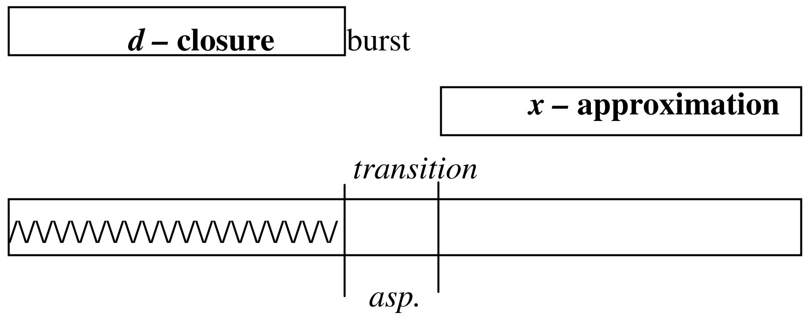


Figure 5.14: Illustration of vocal fold vibration in a $[+v]_{stop}[–v]_{fric}$ cluster.

Closure voicing varies among speakers though generally it tends to be shorter in $[+v]_{stop}[–v]_{fric}$ clusters than closure voicing in $[+v]_{stop}[–v]_{fric}$ clusters. This is not surprising considering the necessity to prepare the glottis for a non-vibrating configuration for the ensuing voiceless fricative in [+v][–v] clusters but not in [+v][+v] clusters, where vocal fold vibration must be maintained through the entire cluster. As in other cases of [+v][–v] clusters, in this case too, voicing may fail to be realized in C1 as illustrated in figure (5.15), which results in a phonologically /+v –v/ being phonetically realized as [–v][–v]:

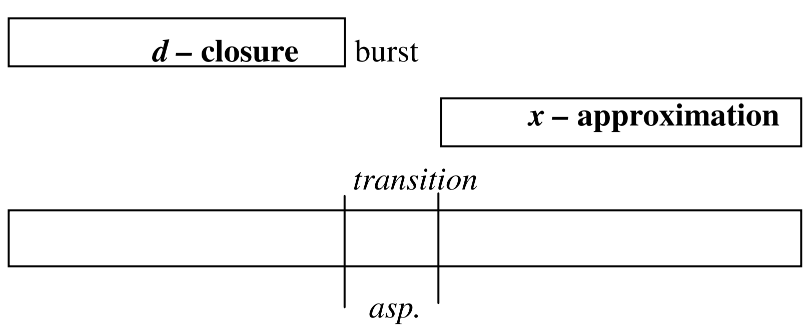


Figure 5.15: Illustration of vocal fold vibration in a $[+v]_{stop}[–v]_{fric}$ cluster.

Next I discuss $[+v]_{fric}[-v]_{stop}$ clusters in which C1 is voiced more often than C1 in $[+v]_{stop}[-v]$ clusters. However, the fricative C1, although usually voiced, is rarely fully voiced. The fricative is frequently voiced from its onset with voicing gradually decreasing towards the offset of the fricative, implying that initiation of vocal fold vibration is aligned with the onset of frication but most speakers fail to maintain voicing through the entire duration of the fricative, as illustrated by figure (5.16):

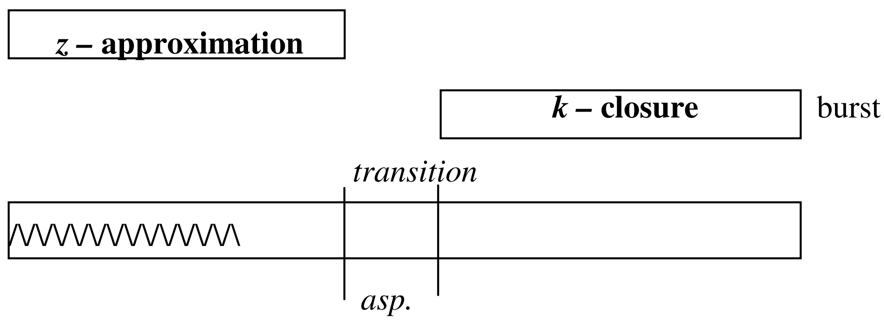


Figure 5.16: Illustration of vocal fold vibration in a $[+v]_{fric}[-v]_{stop}$ cluster.

$[+v]_{fric}[-v]_{fric}$ clusters show a similar pattern to $[+v]_{fric}[-v]_{stop}$ clusters. Only the first segment, the voiced C1, shows vibrations of vocal folds while the ensuing voiceless fircative never contains any voiced portion. The first segment, the fricative, is always realized with partial voicing and only the first portion of the fricative is voiced with voicing decaying as it approaches the voiceless C2 as in figure (5.17):

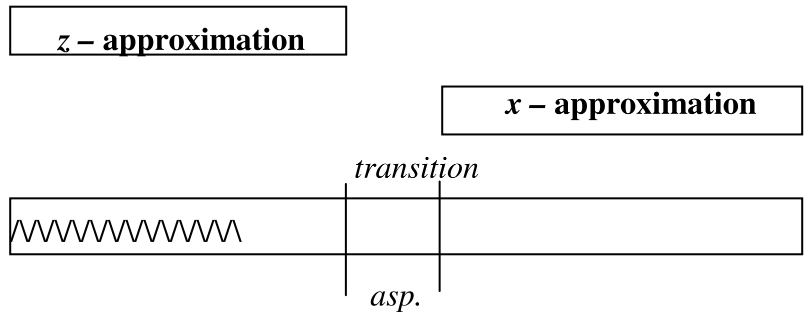


Figure 5.17: Illustration of vocal fold vibration in a $[+v]_{fric}[-v]_{fric}$ cluster.

We may summarize and say that the first member of a [+v][−v] cluster is not always realized as such phonetically. There is great speaker variation in the realization of this cluster type but only the patterns in (9) are attested across all speakers:

9)	Underlying target cluster		Phonetic realization
	/+v −v/	→	[+v][−v]
		→	[−v][−v]
			*[+v][+v]
			*[−v][+v]

Based on the data we observe in this chapter, it appears that only "voicelessness" spreads. That is, voiceless segments surface where they are not present underlyingly but voiced segments never do. This provides further evidence that surface voiceless onset cluster segments may have different underlying representations. They may be underlyingly voiceless, but may be underlyingly voiced. Underlyingly voiced segments may surface as voiceless when voicing fails to be realized physiologically due to unfavorable physiological and physical conditions.

5.6. Summary

To summarize, we have observed through this chapter that there are several different variations to realize the same underlying phonological representation. In figure (5.18), we see the predicted voicing configuration for each one of the clusters in optimal cases.

a. [−voice][−voice] –

b. [+voice][+voice] –

c. [−voice][+voice] –

d. [+voice][−voice] –

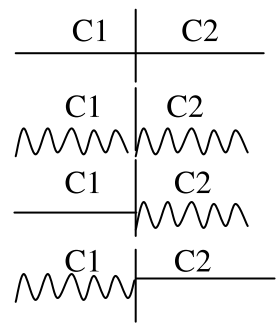
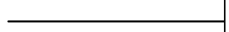
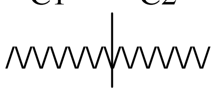
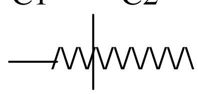
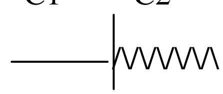
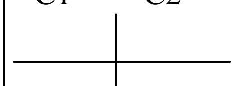
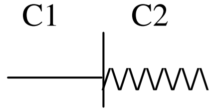
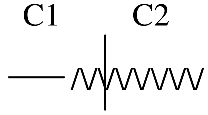
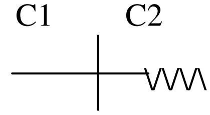
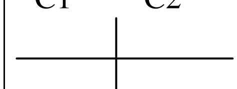
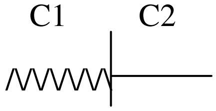
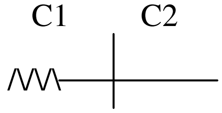
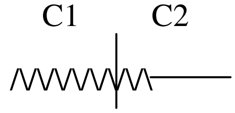
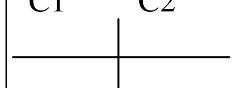


Figure 5.18: Optimal voicing configuration in initial clusters.

Table (5.36) gives a general summary of most of the configurations that were empirically attested for the various realizations of voicing with one realization missing from the [+v][+v] cluster types, which is specific to $[+v]_{stop}[+v]_{stop}$ clusters and that is the case where voicing ceases prior to the burst and is initiated again for the second member of the cluster. I draw the reader's attention to the fact that there are more variations that depend on the nature of the inherent nature of the segments (stops vs. fricatives), which are not represented in the table. The table gives an overview of possible phonetic realizations with the precise description of the misalignment in words. For example, in the second column, there is an identical configuration for both [+v][+v] and [–v][+v] clusters but while one is a result of delayed voicing initiation, the latter is a result of premature voicing. The table does not reflect more specific details like for example vocalic bursts (transitions), which occur in both [+v][+v] and [+v][–v] clusters and were discussed in sections 5.3 and 5.5 respectively.

Table 5.36: Actual voicing occurrences in initial clusters.

[–v][–v]	C1 C2  optimal alignment			
[+v][+v]	C1 C2  optimal alignment	C1 C2  delayed voicing	C1 C2  delayed voicing	C1 C2  delayed voicing
[–v][+v]	C1 C2  optimal alignment	C1 C2  premature voice	C1 C2  delayed voicing	C1 C2  voicing failure
[+v][–v]	C1 C2  optimal alignment	C1 C2  voicing decay	C1 C2  voicing persistence	C1 C2  voicing failure

It has been demonstrated throughout the chapter and in the summarizing table (5.36), which outlines optimal cases of voicing and the empirical realizations of voicing, that voicing is not always realized as expected and that there are several different realizations of the various laryngeal configurations for each cluster type. While [–v][–v] clusters represent a simple case, with no complex laryngeal specifications and are always realized as voiceless in their entirety, all other cluster types exhibit a range of realizations and laryngeal misalignments.

From this chapter we have observed that every segment which is phonetically realized with vibration of vocal folds is also underlyingly voiced, but not every underlyingly voiced segment is realized with phonetic voicing. Immediate phonetic environment plays a crucial role in the realization of vocal fold vibration and word initial environment is highly unfavorable for voicing, resulting in frequent failure to realize voicing in this context. Moreover, a segment preceding a vowel is in a phonetically more advantageous environment. In fact, the transition of the segment into the vowel provides many more cues that are not present in the transition of a consonant into another consonant. Some of the transitional cues that are found at the onset of a vowel can be more prominent than vocal fold vibration. A sharp rise in F1 seems to be a reliable cue for voicing, which in several cases seems to compensate for the lack of vocal fold vibration. Other cues such as higher *f*0 after voiced consonant or the duration of a fricative can also contribute to the perception of voicing in consonants. Other cues for voicing, which are found in the transition to the vowel were more prominent and contributed to the perception of voicing. These cues, however, cannot aid in the perception of voicing in the first member of a cluster. In C1 vocal fold vibration seems to be the only reliable cue for voicing and therefore its presence is crucial for the perception of voicing in C1 position. In its absence there are no cues that can function as a replacement for vocal fold vibration.

However, environment is not the only crucial parameter that influences realization of voicing. Strategies for realization voicing depend on the inherent nature of the segment. Although both stops and fricatives share the tendency of voicing to peter out towards the offset of the segment, which is typical for obstruents, in stops voicing tends to decrease and subside before a stop release, but in fricatives two patterns emerge. While voicing may decrease towards the offset of the fricative it can also be delayed at the onset of fricative and begin only towards the middle of the segment.

Whether it is the former or the latter voicing realization depends on the voicing specification of the following segment.

Cases of clusters with mixed voicing are much more rare cross linguistically, nonetheless, in this chapter I presented empirical evidence in the form of phonetic experiments to support the claims that both [–v][+v] and [+v][–v] cluster types exist. I also demonstrated that surface voicelessness does not always indicate underlying voicelessness but can be attributed to failure in realizing an underlying specification of voicing due to unfavorable phonetic conditions.

APPENDIX I: Word list[14]

The target cluster is bolded.

Meaning	Transliteration	Word in Hebrew
poor	**d**alim	1. דלים
waves	**g**alim	2. גלים
materials	**b**adim	3. בדים
easy/light	**k**alim	4. קלים
flying (masc.)	**t**asim	5. טסים
vase	**k**adim	6. כדים
strangers	**z**arim	7. זרים
ministers	**s**arim	8. שרים
bulls	**p**arim	9. פרים
villages	**kf**arim	10. כפרים
stains	**kt**amim	11. כתמים
vines	**gf**anim	12. גפנים
contents	**tx**anim	13. תכנים
paths	**dr**axim	14. דרכים
urges	**dx**afim	15. דחפים
sizes	**gd**alim	16. גדלים
precedence	**kd**amim	17. קדמים
flags	**dg**alim	18. דגלים
fissure	**bk**aʔim	19. בקעים
clothes	**bg**adim	20. בגדים
palms	**dk**alim	21. דקלים
surveys	**sk**arim	22. סקרים
vices	**sg**anim	23. סגנים
men/males	**gv**arim	24. גברים
damns	**sx**arim	25. סכרים
regulations	**tk**anim	26. תקנים
drill hole	**kd**axim	27. קדחים
walls	**kt**alim	28. כתלים
beards	**zk**anim	29. זקנים
graves	**kv**arim	30. קברים
males	**zx**arim	31. זכרים
faults	**pg**amim	32. פגמים
mornings	**bk**arim	33. בקרים
lies	**kz**avim	34. כזבים
money	**ks**avim	35. כספים

14 More words were recorded but only the words which were used in the experiment are listed..

books	**sf**arim	36. ספרים
bushes	**sv**axim	37. סבכים
sacrifices	**zv**axim	38. זבחים

CHAPTER 6

CONCLUSION

6.0. Summary of the dissertation

In this dissertation I present a study of the feature [sonorant] and the feature [voice] in biconsonantal onset clusters. I begin by showing, in chapter 2, that there are 4 logical combinations of the feature [sonorant] in clusters, OS, OO, SS and SO and 16 possible logical groupings of these clusters, each representing a theoretically, and logically, possible language type. A survey of 62 languages from 22 language families was conducted and results show that of the 16 logically possible combinations, only 4 emerge as occurring language types: {OS}, {OS, OO}, {OS, OO, SS} and {OS, OO, SS, SO}. The least marked cluster, OS, the cluster which is implied by all other clusters, is a cluster with a rise in sonority. That is, the value of the feature [sonorant] "increases" towards the nucleus from the obstruent, which is marked [–sonorant], to the sonorant whose value is [+sonorant]. The most marked cluster, SO, which implies the presence of all other clusters, is a cluster with falling sonority. That is, the value of the feature [sonorant] of the first member of the cluster, [+sonorant], is "greater" than that of the second member of the cluster, whose value is [–sonorant].

Next I address the typological distribution of the feature [voice], which also has 4 logical combinations and yields the following 4 possible clustering combinations: [–v][–v], [+v][+v], [–v][+v] and [+v][–v]. These 4 clusters also yield 16 logically possible groupings, and thus 16 logically possible language types. However, before addressing possible language types, an additional issue regarding these clusters had to be addressed. The cluster type [+v][–v] is a controversial cluster, and although it is documented in the grammars of several languages, it had been claimed to be an impossible cluster type, both phonetically and phonologically (Lindblöm 1983, Lombardi 1991). In chapter 2 I provide evidence for the existence of this cluster type in at least 3 languages. Once the existence of the [+v][–v] cluster type was established, I address the problem of possible language types. Of the 16 logically possible language types, 6 emerge as attested language types {[–v][–v]}, {[–v][–v], [+v][+v]}, {[–v][–v], [–v][+v]} {[–v][–v], [+v][+v], [–v][+v]}, {[–v][–v], [–v][+v], [+v][–v]}, {[–v][–v], [+v][+v], [–v][+v], [+v][–v]}. The distributional patterns which the feature [sonorant] follows are different from the distributional patterns of the feature [voice]. This suggests that, although these two features may interact in complex ways, they are not mutually dependent and cannot be reduced to a single pattern. That is, the phonological patterning of one of these features in clusters cannot be conjectured based on the other feature. The typological patterning of clusters based on the feature [sonorant] does not provide any clues about the phonological patterning of the feature [voice] in clusters. A language can be of one type in regards to one of these features, and another type in regards to the other. For example, Russian exhibits all possible clusters of the feature [sonorant], OS, OO, SS and SO, making it a type 4 language in terms of the feature [sonorant], yet only two combinations of the feature [voice] are permitted, [–v][–v] and [+v][+v], making it a type 2 language in terms of the feature [voice]. Modern Hebrew is the opposite example. It only allows two clusters in terms of the feature [sonorant], OS and OO, making it a type 2 language in terms of the feature [sonorant], but allows all possible voicing combinations, [–v][–v], [+v][+v], [–v][+v] and [+v][–v], making it a type 4 language in terms of the feature [voice].

Of interest to this work is the behavior of the rarely attested, yet existent cluster type [+v][–v]. After establishing the existence of this cluster, I address the question of its phonological representation. Are word initial [+v][–v] sequences tautosyllabic clusters? In chapter 3, which presents a case study of Modern Hebrew, I provide evidence that the sequence [+v][–v] is a “real” cluster and should be treated as such phonologically. I show that treating the voiced first member of the cluster as "extrametrical" is not the proper analysis of [+v][–v] word initial obstruent sequences. I present evidence from two separate morphological processes in Modern Hebrew, plural affixation in the Segolate noun class and diminutive reduplication. I show that the output of these processes are permissible clusters. When impermissible clusters such as SO clusters arise, they are broken by vowel epenthesis. However, a [+v][–v] sequence is never broken by vowel epenthesis, providing strong evidence that it is a licit cluster and further confirming its status as an onset clusters.

Lastly, of interest in this work is the actual phonetic realization of clusters with different voicing combinations. In chapter 4 I move away from the phonology of clusters and present the phonetic aspect of realizing voicing in clusters. I begin by giving an overview of the physiological mechanisms involved in producing voicing in general, and then discuss the challenges of producing voicing in clusters. Each voicing combination is discussed separately and predictions regarding the theoretically possible realizations of voicing in each voicing combination are presented. The main claim in chapter 4 is that underlying voiced segments can surface as voiceless as a result of phonetic failure to produce voicing. That is, if an underlying voiced segment is realized as a surface voiceless segment, it does not follow that the underlyingly voiced segment went through a phonological process of devoicing, that is, through a phonological process which changes its categorical representation from voiced to voiceless. Rather, the case at issue is a phonetic failure to realize voicing. These post-lexical processes are phonetic in nature, and are not categorical.

I provide evidence to support these claims in chapter 5, where I show that, in Modern Hebrew, clusters such as /+v –v/, can surface as phonetically [–v][–v] in some cases and as [+v][–v] in other cases. Moreover, clusters such as /+v +v/ can also surface with a voiceless C1, a fact which cannot be accounted for by assimilation since there is no trigger for voicelessness in these types of clusters. The only way to account for these types of realizations is by assuming voicing failure.

In chapter 5 I provide both qualitative data, demonstrating each phonetic realization of clusters, as well as quantitative data, to show both sporadic failure of voicing and systematic voicing. The data provides evidence to support the arguments and hypotheses made in chapter 4. Phonetic characteristics correlated with voicing such as closure voicing, aspiration, burst duration, fricative duration and fundamental frequency are all provided to show that there are cases where closure voicing is the only cue for voicing and other, secondary cues are not available to compensate for the loss of the primary cue. I hypothesize that in these cases listeners appeal to the paradigm to recover the voicing specification of the target segment.

Moreover, the phonetic data supports the claim that surface voiceless segments can be either underlyingly voiced or underlyingly voiceless and that phonetic voicelessness is not always an indication of underlying voicelessness. Many of the results show mis-alignment or mis-production of voicing which can only be accounted for by assuming phonetic difficulties in the production of voicing, rather than phonological categorical change. This, once again, offers empirical support for the claims and hypotheses raised throughout this work.

6.1. Topics for future research

6.1.1. Sonority reversed clusters: their phonological representation

An unexpected result of the typological survey presented in chapter 2 is the distribution of clusters in terms of the feature [sonorant]. Clusters with sonority declines are generally considered irregular clusters and are treated in the literature as having a special phonological status. According to the SSP clusters should exhibit a sonority incline towards the nucleus and a decline towards the margins. Clusters in which the second member of the cluster is less sonorous than the first member pose a challenge for the SSP as they exhibit a sonority decline towards the margins. Traditionally these clusters have received special treatment and were treated separately from clusters with sonority inclines. Sequences in which the first member of the clusters was less sonorous than the second member were considered tautosyllabic and both C1 and C2 were syllabified together. However, sequences in which the first member of the cluster was more sonorous than the second member were treated differently. The two segments of the sequence were separated and only C2 was syllabified with the nucleus. The first member of the sequence, C1, was syllabified separately and usually attached to a separate prosodic tier. In this manner, although C1 and C2 formed a sequence, they were not a cluster since they belong to separate prosodic tiers.

Curiously, results of the cross-linguistic survey show that SO clusters are part of the typology and do not behave differently. If indeed these clusters behaved differently, or if they indeed have different phonological representations from all other clusters, we would not expect them to integrate neatly into the typology. If they have a special phonological status, they should be expected to occur randomly. Thus, if they indeed have a special status, we should expect to find languages that have clusters such as *{OS and SO} or *{OS, OO and SO}. The fact that such languages do not exist, combined with the fact that in the typology SO clusters imply all other cluster types, strongly suggests that SO clusters should be treated as regular clusters. This issue is outside the scope of this work and is left for future research.

6.1.2. Voicing failure

It was claimed in this dissertation that segments phonetically realized as voiceless are not necessarily specified as such underlyingly. Several cases have been presented where I argued that the "voicelessness" of a segment was a result of voicing failure in the realization of an underlying voiced phonological target. Voicing failure of a voiced target can be established if two conditions exist, (i) the target segment surfaces as voiced in at least a certain percentage of cases and (ii) there is clear voicing of the target segment in other contexts, for example in other forms in the paradigm. However, I suggest that voicing failure could be considered in other cases as well. Voicing failure can be contemplated for languages such as German, for which the feature [spread glottis] has been proposed, as detailed in chapter 1. The feature [spread glottis] was proposed for those languages where there is no evidence for closure voicing in word initial or word final contexts. However, experimental evidence pointed out in chapter 1 suggests that despite lack of acoustic evidence of voicing, muscular activity is still present during the production of *b*, *d* and *g*. Thus, although voicing is not realized in the full sense of the word, it seems that the target may still be [+voice] despite lack of acoustic voicing. The presence of muscular activity but lack of acoustic voicing could suggest another case of failure to achieve voicing, which could stem from failure to achieve the necessary threshold for initiation of vibration of vocal folds. This hypothesis, however, is also left for further research.

BIBLIOGRAPHY

Abramson, A. S. and Lisker L. (1973). Voice-timing perception in Spanish word initial stops. *Journal of Phonetics, 1*, 1-18.

Adam, G. (1996). The fate of spirantization in Modern Hebrew: Free variation in Optimality Theory. Ms., Tel-Aviv University, Tel Aviv.

Agard, F. B. (1958). Structural sketch of Rumanian. *Language*, *34*(1), 7-21.

Aloni-Feinberg, Y. (1978). Language planning in the revival of Hebrew: Principles and ways. *Hebrew Computation Linguistics*, *13*, 13-27.

Ambrazas, V. (1997). *Lithuanian Grammar*. Vilnius: Baltos lankos.

Andrews, H. (1949). Phonemes and morphophonemes of Temoayan Otomi. *International Journal of American Linguistics*, *15*, 213-22.

Aschmann, H. P. (1946). Totonaco phonemes. *International Journal of American Linguistics*, *12*, 34-43.

Asinovskii, A. S. (1991). *Konsonantizm Chukotskogo Jazyka*. Leningrad: Nauka. (In Russian).

Awbery, G. M. (1984). Phonotactic constraints in Welsh. In Martin J. Ball and Glyn E. Jones (Ed.), *Welsh Phonology, Selected Readings* (65-104). Cardiff: University of Wales Press.

Ball, M., and Fife, J. (Eds.). (1993). *The Celtic Languages*. London: Routledge.

Barkaï, M., and J. Horvath. (1978). Voicing assimilation and the sonority hierarchy: Evidence from Russian, Hebrew and Hungarian. *Linguistics*, *212*, 77-88.

Barker, M. A. K. (1964). Klamath Grammar. University of California Publications in Linguistics 32. University of California Press.

Bat-El, O. (1989). *Phonology and Word Structure in Modern Hebrew*. Ph.D. dissertation, UCLA.

Bat-El, O. (1994). Stem modification and cluster transfer in Modern Hebrew. *Natural Language and Linguistic Theory*, *12*, 571-596.

Bat-El, O. (1998). Vocalic alternations and stress in Hebrew verbs. Unpublished Ms., Tel – Aviv University, Tel Aviv.

Bat-El, O. (2002). Hebrew reduplication the interpretation of forms with identical consonants. Presented at UCSC 2002.

Bat-El, O. (2004). The emergence of the binary trochaic foot in Hebrew hypocoristics. *Proceedings of IATL 20*. Bar-Ilan University.

Becker, M. (2003). Hebrew stress: Can't you hear those trochees? In Kaiser and Arunachalam (Eds.), *Proceedings of PLC*, *26* (9.1), 45-48.

Bergsland, K. (1997). *Aleut Grammar: Unangam Tunuganaan Achixaasix*. Fairbanks: Alaska Native Language Center.

Berman, R. (1997). Modern Hebrew. In Robert Hetzron (Ed.), *The Semitic Languages*, 312-333. New York: Routledge.

Blanc, H. (1964). Israeli Hebrew Texts. In Haiim B. Rosen (Ed.), *Studies in Egyptology and Linguistics in Honor of H. J. Polotsky* (132-152). Jerusalem: The Israel Exploration Society.

Blevins, J. (2003). The independent nature of phonotactic constraints: An alternative to syllable-based approaches. In C. Féry and R. van de Vijver (Eds.), *The Syllable in Optimality Theor*y (375-404). Cambridge: Cambridge UP.

Boersma, P. and Weenink, D. (2005). Praat: doing phonetics by computer (Version 4.4) [Computer program]. Retrieved December 27, 2005, from http://www.praat.org/

Bogoras, W. (1922). Chukchee. In Franz Boas (Ed.), *Handbook of American Indian Languages: Part 2*. Washington: Smithsonian.

Bolozky, S. (1972). *Categorical Limitations on Rules in the Phonology of Modern Hebrew*. Ph.D. dissertation, University of Illinois.

Bolozky, S. (1978). Word formation processes in the Hebrew verb system: Denominative verbs. *Afroasiatic Linguistics*, *5*, 1-26.

Bolozky, S. (1980). Paradigm coherence in Hebrew. *Afroasiatic Linguistics*, *7*(4), 103 -126.

Booij, G. (1995). *The Phonology of Dutch*. Oxford: Oxford UP.

Bowden, J. (2001). *Taba: Description of a South Halmahera language*. Pacific Linguistics 521. Canberra, Australian National University.

Briggs, E. (1961). *Mitla Zapotec Grammar*. Mexico: Instituto Lingüístico de Verano and Centro de Investigaciones Antropológicas de México.

Browman, C., and Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook 3*, 219-252.

Browman, C., and Goldstein, L. (1990). Gestural specification using dynamically defined articulatory structures. *Journal of Phonetics*, *18*, 299-320.

Browman, C., and Goldstein, L. (1992). Articulatory Phonology: An Overview. *Phonetica*, *49*, 155-180.

Butskhrikidze, M. (2002. *The Consonant Phonotactics of Georgian*. Utrecht: LOT.

Chavarria-Aguilar, O. L. (1951). The phonemes of Costa Rican Spanish. *Language*, *27*(3), 248-253.

Chayen, M. J. (1972). The accent of Israeli Hebrew. *Lenshonenu*, *36*, 212-219, 287-300.

Chayen, M. J. (1973). *The Phonetics of Modern Hebrew*. The Hague: Mouton.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, *22*, 129-159.

Chitoran, I. (1998). Georgian harmonic clusters: Phonetic cues to phonological representation. *Phonology*, *15*(2), 121-141.

Chitoran, I. (1999). Accounting for sonority violations: The case of Georgian consonant sequencing. *Proceedings of the 14th International Congress of Phonetic Sciences* (101-104). San Francisco, August 1999.

Chitoran, I., Goldstein, L. and Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. In C. Gussenhoven and N. Warner (Eds.), *Laboratory Phonology 7* (419-447). Berlin, New York: Mouton de Gruyter.

Cho, Y.-M. Y. (1990). A typology of voicing assimilation. *West Coast Conference on Formal Linguistics*, *9*, 141–155.

Chomsky, N. and Halle, M. (1968). *Sound Pattern of English*. New York: Harper and Row.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston and M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech* (282-333). Cambridge: Cambridge UP.

Coates, W. A. and da Silva, M.W.S. (1960). The segmental phonemes of Sinhalese. *University of Ceylon Review*, *18*, 163-175.

Cohen, A., Ebeling, C.L., Fokkema, K. and Holk, A.G. F. van. (1961). *Fonologie van het Nederlands en het Fries. Inleiding tot de Moderne Klankleer.* The Hague: Martinus Nijhoff's Gravenhage.

Davies, H. J. (1980). *Kobon phonology.* Pacific Linguistics B, 68. Canberra: Australian National University.

Dean, J. (1955). The phonemes of Bilaan. *Philippine Journal of Science*, *84*(3), 311 -322.

Dell, F. (1995). Consonant clusters and phonological syllables in French. *Lingua*, *95*, 5-26.

Derbyshire D. (1985). *Hixkaryana and Linguistic Typology*. SIL Publications.

Devens, M. (1980). Oriental Israeli Hebrew: A study in phonetics. *Afroasiatic Linguistics*, *7*, 127-142.

Diderichsen, P. (1964). *Essentials of Danish Grammar*. Copenhagen: Akademisk Forlag.

Dirks, S. (1953). Campa (Arawak) phonemes. *International Journal of American Linguistics*, *19*, 302-304.

Einaudi, P. F. (1976). *A Grammar of Biloxi*. New York: Garland.

Eleftheriades, O. (1985). *Modern Greek: A Contemporary Grammar*. Palo Alto: Pacific Books Publishers.

Elson, B. (1947). Sierra Popoluca syllable structure. *International Journal of American Linguistics*, *13*(1), 13-17.

Faber, A. (1997). Genetic subgrouping of the Semitic languages. In Robert Hetzron (Ed.), *The Semitic Languages* (3-15). New York: Routledge.

Fast, P. W. (1953). Amuesha (Arawak) phonemes. *International Journal of American Linguistics*, *19*, 191-194.

Ferguson, C. and Chowdhury, M. (1960). The Phonemes of Bengali. *Language*, *36*(1), 22-59.

Flege, J. E. (1982). Laryngeal timing and phonation onset in utterance-initial stops. *Journal of Phonetics*, *10*, 177-192.

Fortescue, M. D. (1984). *West Greenlandic*. London, Dover, N.H.: Croom Helm.

Fujimora, O. and Sawashima, M. (1971). Consonant sequences and laryngeal control. *Annual Bulletin of Research Institute of Logopedics and Phoniatrics*, *5*, 1-13.

Garvin, P. L. (1948). Kutenai I: Phonemics. *International Journal of American Linguistics*, *14*, 37-42.

Graf, D. (2002). A Study on Nominal Reduplication in Modern Hebrew. *Proceedings of the 18th conferences of the Israeli Association of Theoretical Linguistics*, Ramat-Gan, Israel.

Graf, D. and Ussishkin, A. (2003). Emergent iambs: Stress in Modern Hebrew. *Lingua*, *113*, 239-270.

Gralow, F. L. (1976). Fonología del Chamí. *Sistemas Fonológicos de Idiomas Colombianos 3*, 29-42. Bogotá: Ministerio de Gobierno and Instituto Lingüístico de Verano.

Green, A. (1997). *The Prosodic Structure of Irish, Scots Gaelic, and Manx*. Ph.D. dissertation, Cornell University.

Green, A. (2004). Opacity in Tiberian Hebrew: Morphology, not phonology. *ZAS Papers in Linguistics*, *37*, 37-70. (ROA 7030105)

Greenberg, J. (1965). Some generalizations concerning initial and final consonant sequences. *Linguistics*, *18*, 5-34. (reprinted as Greenberg 1978).

Greenberg, J. (1978). Some generalizations concerning initial and final consonant clusters. In J.H. Greenberg (Ed.), *Universals of Human Language, vol. 2: Phonology*. Stanford, CA.: Stanford UP.

Gumperz, J. (1958). Phonological differences in three Hindi dialects. *Language*, *34*, 212-224.

Gvarjaladze T. and Gvarjaladze, I. (1974). *English-Georgian Dictionary*. Tbilisi: State Publication House.

Hajek, J. and Bowden, J. (1999). Taba and Roma: clusters and geminates in two Austronesian languages. In *Proceedings of the XIVth Congress of Phonetic Sciences: San Francisco, 1-7August* (1033-1036). American Institute of Physics.

Hall, T. A. (1993). The phonology of German /R/. *Phonology*, *10*(1), 83-105.

Halle, M. and Stevens, K. N. (1971). A note on laryngeal features. *MIT Research Laboratory of Electronics Quarterly Progress Report 101* (198 –213).

Halpern, A. M. (1946). Yuma I: Phonemics. *International Journal of American Linguistics*, *12*(1), 25-33.

Ham, W. H. (1998). *Phonetic and Phonological Aspects of Geminate Timing*, Ph.D. dissertation, Cornell University.

Hansen, A. (1967). *Moderne Dansk*. København: Forlag. Harley.

Henderson, E. (1991). Khasi clusters and Greenberg's universals. *Mon Khmer Studies*, *18*(19), 61-6.

Hirose, H. (1997). Investigating laryngeal physiology. In W. J. Hardcastle and J. Laver (Ed.), *The Handbook of Phonetic Sciences*. Oxford: Blackwell.

Hoberman, R.D. (1995). Current issues in Semitic phonology. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory*. Oxford: Blackwell.

Hodge, C. T. (1946). Serbo-Croatian phonemes. *Language*, *22*,112-20.

Hoff, B. J. (1968). *The Carib Language: Phonology, Morphonology,Morphology, Texts and Word Index*. The Hague: Martinus Nijhoff.

Honda, K., Hirai, H., and Kusakawa, N. (1993). Modeling vocal tract organs based on MRI and EMG observations and its implication on brain function. *Annual Bullatin of the Research Institute of Logopedics and Phoniatrics*, *27*, 37-50.

Honda, K., Hirai, H., Masaki, S., and Shimada, Y. (1999). Role of vertical larynx movement and cervical lordosis in F0 control. *Language and Speech*, *42*, 401-411.

Hongkai, S. (1986). Notes on Tibeto-Burman consonant clusters. *Linguistics of the Tibeto-Burman Area*. Beijing: Chinese Academy of Seoul Sciences.

Howard, L. (1967). Camsa phonology. In Viola G. Waterhouse (Ed.), *Phonemic Systems of Colombian Languages* (73-87). Summer Institute of Linguistics Publications in Linguistics and Related Fields, 14. Norman, Summer Institute of Linguistics of the University of Oklahoma.

Hualde, J. I. (1991). *Basque Phonology*. London, New York: Routledge.

Huffman, F. E. (1990). Burmese Mon, Thai Mon, and Nyah Kur: A synchronic comparison. *Mon Khmer Studies*, *16-17*, 31-84.

Hsin, T.-H. (2000). Consonant clusters in Tsou and their theoretical implications. *The Proceedings of the 18th West Coast Conference on Formal Linguistics*, Cascadilla Press.

Idsardi, W. J. (1998). Tiberian Hebrew spirantization and phonological derivations. *Linguistic Inquiry*, *29*, 37-73.

Iverson, G. and Salmons, J. (1995). Aspiration and laryngeal representation in Germanic. *Phonology*, *12*, 369-396.

Jacobs, N. G. (2005). *Yiddish: A Linguistic Introduction*. Cambridge: Cambridge UP.

Jakobson R., Fant, C.G.M. and Halle, M. (1952). *Preliminaries to Speech Analysis*. MIT Acoustic Laboratory Technical Report, 13.

Jessen, M. (2001). Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants. In T.Alan Hall (Ed.), *Distinctive Feature Theory* (237-294). Berlin, New York: Mouton de Gruyter.

Jessen, M. and Ringen, C. O. (2002). Laryngeal features in German. *Phonology*, 19, 189-218.

Joseph, B. D. and Philippaki-Warburton, I. (1987). *Modern Greek*. London: Croom Helm.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, *60*, 286-319.

Kenstowicz, M. (1981). The phonology of Chukchee consonants. In B. Comrie (Ed.), *Studies in the Languages of the USSR*. Carbondale: Linguistic Research Inc.

Kenstowicz, M. (1993). *Phonology in Generative Grammar*. Oxford: Blackwell.

Kingston, J. and Diehl, R. L. (1994). Phonetic knowledge. *Language*, *70*(3), 419-454.

Kreitman, R. (2003). Diminutive reduplication in Modern Hebrew. *Working Papers of the Cornell Phonetics Laboratory*, *15*, 101-129.

Kreitman, R. and Bolozky, S. (2007). Uvulars in Modern Hebrew – Their phonetic and phonological status. Talk presented at the conference of the National Association of Professors of Hebrew, Sydney, Australia.

Kučera, H. (1961). *The Phonology of Czech*. The Hague: Mouton and Company.

Kučera, H. and Monroe, G. K. (1968). *A Comparative Quantitative Phonology of Russian, Czech and German*. New York: American Elsevier Publication.

de Lacy, P. (2002). *The Formal Expression of Markedness*. Ph.D. dissertation, University of Massachusetts, Amherst.

Ladefoged, P. and Maddieson, I. (1996). *The Sounds of the World's Languages*. Oxford: Blackwell.

Laufer, A. (1983). Introduction to Acoustic Phonetics. *Torat ha-hegeh shel ha-Ivrit ha-ḥadashah*. Unit 5. Tel-Aviv: The Open University. (In Hebrew).

Laufer, A. (1984). Acoustic description of sounds. *Torat ha-hegeh shel ha-Ivrit ha-ḥadashah*. Units 6 & 7. Tel-Aviv: The Open University. (In Hebrew).

Laufer, A. (1994). Voicing in contemporary Hebrew. *Leshonenu*, *57*(4), 299 -342 (in Hebrew.)

Laufer, A. (1995). Voice/voiceless against lenis/fortis. *Hebrew Linguistics*, *39*, 41-65. (In Hebrew).

Laufer, A. (1998). Voicing in contemporary Hebrew in comparison with other languages. *Hebrew Studies*, *39*, 143-179.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Mass: MIT Press.

Levi, S. V. (2004). *The Representation of Underlying Glides*. Ph.D. dissertation, University of Washington, Seattle.

Levin, J. (1985). *A Metrical Theory of Syllabicity.* Ph.D. dissertation, MIT.

Levinsohn, S. H. (1979). Fonología del inga. In Marilyn E. Cathcart *et al*. (Eds.), *Sistemas fonológicos de idiomas colombianos 4*, 65-85. Bogota: Ministerio de Gobierno.

Liberman, P. (1970). Towards a unified phonetic theory. *Linguistic Inquiry*, *1*, 307 -322.

Liberman, P. (1977). *Speech Physiology and Acoustic Phonetics*. New York: Macmillan.

Lindblöm, B. (1983). Economy of speech gestures. In P. MacNeilage (Ed.), *Speech Production* (217-246). New York: Springer-Verlag.

Lindblöm B. and Maddieson, I. (1988). Phonetic universals in consonant systems. In L. M. Hyman and C. N. Li (Eds.), *Language, Speech and Mind*: 62-78.

Lindqvist, J. (1972). Laryngeal articulation studies in Swedish subjects. *Stockholm Royal Institute of Technology Speech Transmission Laboratory, Quarterly Progress and Status Report*, *2-3*, 10-27.

Lisker, L. (1963). On Hultzen's voiceless lenis stops in prevocalic clusters. *Word*, *19*, 376-387.

Lisker, L. (1975). Is it VOT or a first-formant transition detector? *Journal of the Acoustical Society of America*, *57*, 1547-1551.

Lisker, L., and Abramson, A. S. (1964). A cross language study of voicing in initial stops: Acoustical measurements. *Word*, *20*, 384-422.

Löfqvist, A., and Yoshioka, H. (1980a). Laryngeal activity in Swedish obstruent clusters. *Journal of the Acoustical Society of America*, *68*, 792-801.

Löfqvist, A., and Yoshioka, H. (1980b). Laryngeal activity in Icelandic obstruent production. *Haskins Laboratories: Status Report on Speech Research*, *SR-63/64*, 275–292.

Lombardi, L. (1991). *Laryngeal Features and Laryngeal Neutralization*. Ph.D. dissertation. University of Massachusetts, Amherst.

Lombardi, L. (1995a). Laryngeal features and privativity. *The Linguistic Review*, *12*, 35-59.

Lombardi, L. (1995b). Laryngeal neutralization and syllable wellformedness. *Natural Language and Linguistic Theory*, *13*, 39-74.

Lombardi, L. (1999). Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language and Linguistic Theory*, *17*, 267-302.

MacKay, C. J. (1994). A sketch of Misantla Totonac phonology. *International Journal of American Linguistics*, *60*(4), 369-419.

Maddieson, I. and Ladefoged, P. (1993). Phonetics of partially nasal consonants. In M. K. Huffman and R. Krakow, (Eds.), *Nasals, Nasalization and the Velum* (251-301). San Diego, Academic Press.

Mallinson, G. (1986). *Rumanian*. London: Croom Helm.

Marantz, A. (1982). Re Reduplication. *Linguistic Inquiry*, *13*, 483-545.

Marlett, S. A. and Pickett, V. B. (1987). The syllable structure and aspect morphology of Isthmus Zapotec. *International Journal of American Linguistics*, *53*, 398 -422.

Marlett, S. A. (1988). The syllable structure of Seri. *International Journal of American Linguistics*, *54*, 245-278.

Matthews, H. (1955). A phonemic analysis of a Dakota dialect. *International Journal of American Linguistics*, *21*, 56-59.

McArthur, H. S. and McArthur, L. E. (1956). Aguacatec (Mayan) phonemes within the stress group. *International Journal of American Linguistics*, *22*, 72-76.

McCarthy, J. (1979). *Formal Problems in Semitic Phonology and Morphology*. Ph.D. dissertation, MIT.

McCarthy, J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry*, *12*, 373-418.

McCarthy, J. and Prince, A. (1990a). Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory*, *8*, 209-283.

McCarthy, J. and Prince, A. (1990b). Prosodic Morphology and templatic morphology. In M. Eid, and J. McCarthy (Eds.), *Perspectives on Arabic Linguistics 2: Papers from the Second Annual Symposium on Arabic Linguistics* (1-54). Amsterdam: John Benjamins.

McKaughan, H. P. (1954). Chatino formulas and phonemes. *International Journal of American Linguistics*, *20*, 23-27.

Morag, S. (2003). Sheva kfulot beged keferet. In M. Bar-Asher, Y. Broyer and A. Maman (Eds.), *Iyunim be-Ivrit, ba-Aramit uvi-leshonot ha-Yehudim*: 150-194. Jerusalem: Magnes, Hebrew University. (In Hebrew).

Morelli, F. (1998). Markedness relations and implicational universals in the typology of onset obstruent clusters. In *Proceedings of NELS 28*.

Morelli, F. (1999). *The Phonotactics and Phonology of Obstruent Clusters in Optimality Theory*. Ph.D. dissertation, University of Maryland at College Park.

Morelli, F. (2003). The relative harmony of /s+stop/ onsets: Obstruent clusters and the sonority sequencing principle. In C. Féry and R. van de Vijver, (Eds.), *The Syllable in Optimality Theory*: 356-371. Cambridge: Cambridge UP.

Mortensen, C. A. (1999). *A Reference Grammar of the Northern Embera Languages: Studies in the Languages of Colombia 7.* Arlington: Summer Institute of Linguistics and the University of Texas, Publications in Linguistics, 118, Dallas, Summer Institute of Linguistics and the University of Texas at Arlington.

Munhall, K. and Löfqvist, A. (1992). Gestural aggregation in speech: laryngeal gestures. *Journal of Phonetics*, *20*, 11-126.

Nacaskul, K. (1978). The syllabic and morphological structure of Cambodian words. In P. Jenner (Ed.), *Mon-Khmer Studies*, *7*, 183-200.

Nagaraja, K.S. (1990). *Khasi Phonetic Reader*. Mysore: Central Institute of Indian Languages.

Næs, O. (1965). *Norsk Grammatikk: Elementære Strukturer og Syntaks.* Forlag: Fabritius & Sønners. (In Norwegian).

O'Dochartaigh, C. (1992). The Irish language. In D. Macaulay (Ed.), *The Celtic Languages*. Cambridge: Cambridge UP.

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In Peter MacNeilage (Ed.), *The production of speech*: 189-216. New York: Springer-Verlag.

Ohala M. (1983). *Aspects of Hindi Phonology*. Delhi: Motilal Baarsidass.

Okrand, M. (1979). Metathesis in Costanoan grammar. *International Journal of American Linguistics*, *45*, 123-130.

Penzl, Herbert. (1955). *A Grammar of Pashto: A Descriptive Study of the Dialect of Kandahar, Afghanistan.* Washington, D.C.: American Council of Learned Societies.

Petrova, O. Plapp, R., Ringen, C., and Szentgyorgyi, S. (2006). Voice and aspiration: evidence from Russian, Hungarian, German, Swedish, and Turkish. *The Linguistic Review*, *23*, 1-35.

Rabel, L. (1961). *Khasi, a Language of Assam*. Baton Rouge: Louisiana State UP.

Rex, E. and Schöttelndreyer, M. (1973). Sistema fonológico del Catío. In *Sistemas Fonológicos de Idiomas Colombianos 2*, 73-85. Bogotá: Ministerio de Gobierno.

Riehl, A. (2008). *The Phonology and Phonetics of Nasal-Obstruent Sequences*. Ph.D. dissertation, Cornell University.

Riggs, V. (1949). Alternate phonemic analysis of Comanche. *International Journal of American Linguistics*, *15*, 229-231.

Rögnvaldsson, E. (1993). *Íslensk Hljoðkerfisfræði*. Reykjavík: Málvísindastofnun Háskóla Íslands. (In Icelandic).

Rosen, H. (1957). *Ivrit tova: lyunim btaħbir halašon ha"nxona"*. Jerusalem: Kiryat Hasefer. (In Hebrew).

Sapir E. (1923). The Phonetics of Haida. *International Journal of American Linguistics*, *2*(3/4), 143-158.

Sawashima, M. and Hirose, H. (1983). Laryngeal gestures in speech production. In P. MacNeilage, (Ed.), *The Production of Speech*: 11-38. New York: Springer -Verlag.

Sawicka, I. (1974). *Struktura Grup Spółgłoskowych w Jezykach Słowiańskich.* Wrocław: Zakład Narodowy im. Ossolińskich.

Scatton, E. A. (1984). *A Reference Grammar of Modern Bulgarian.* Cambridge: Slavica Publishers.

Schaub, W. (1985). *Babungo.* London: Croom Helm.

Schwartzwald - Rodrigue, O. (1995). The components of the Modern Hebrew lexicon: The influence of Hebrew classical sources, Jewish languages and other foreign languages on Modern Hebrew. *Hebrew Linguistics*, *39*, 79-90. (In Hebrew).

Selkirk, E. (1984). On the major class features and syllable theory. In M. Aronoff and R.T. Oehrle (Eds.), *Language Sound Structures*: 107-136. Cambridge, MA: MIT Press.

Sigurd, B. (1965). *Phonotactic Structure in Swedish*. Lund: Berlingska Boktryckeriet.

Siptár, P. and Törkenczy, M. (2000). *The Phonology of Hungarian*. Oxford: Oxford UP.

Skorik, P. J. (1961). *Grammatika Čukotskogo Jazyka: Fonetika i Morfologija Imennyx Častej Reči*. Leningrad: Nauka. (In Russian).

Spencer, R. (1946). The Phonemes of Keresan. *International Journal of American Linguistics*, *12*(4), 229- 236.

Steriade, D. (1982). *Greek Prosodies and the Nature of Syllabification.* Ph.D. dissertation, MIT.

Steriade, D. (1994). Complex onsets as single segments: the Mazateco pattern. In J. Cole and Charles Kisseberth (Eds.), *Perspectives in Phonology* (203-293). Stanford, CA: CSLI Publications.

Steriade, D. (1997). Phonetics in phonology: The case of laryngeal neutralization. Ms., UCLA.

Stevens, K. N. and Blumstein, S. (1981). The search for invariant acoustic correlates of phonetic features. In P. Eimas and J. Miller (Eds.), *Perspectives on the Study of Speech*. Hillsdale, NJ: Erlman.

Stevens, K. N., Keyser, S. J. and Kawasaki, H. (1986). Toward a phonetic and phonological theory of redundant features. In Joseph S. Perkell and Dennis H. Klatt (Eds.), *Symposium on Invariance and Variability of Speech Processes*: 432–469. Hillsdale, NJ: Lawrence Erlbaum.

Stevens, K. (1998). *Acoustic Phonetics*. Cambridge: Blackwell.

Ternes, E. (1992). The Breton language. In Donald MacAulay (Ed.), *The Celtic Languages* (371-452). Cambridge: Cambridge UP.

Trubetskoy, N. (1939). *Grundzüge der Phonologie*. [Translated by C. Baltaxe (1969)]. Berkley: University of California Press.

Ussishkin, A. (2000). *The Emergence of Fixed Prosody*. Ph.D. dissertation, UCSC.

Vaux, B. (2004). The appendix. Talk presented at CUNY. The CUNY Phonology Forum.

Ventzel, T.V. (1983). *The Gypsy Language*. Moscow: Nauka Publication.

Watkins, J. (2002). *The Phonetics of Wa*. Canberra: Pacific Linguistics.

Weinberg, W. (1966). Spoken Israeli Hebrew: Trends in the departures from classical phonology. *Journal of Semitic Studies*, *11*(1), 40-68.

Werner, W. (1962). Problems of Armenian phonology III. *Language*, *38*(3), 254-262.

Westbury, J. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America*, *74*, 1322 -1336.

Westbury, J. and Keating, P. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics*, *22*, 145-166.

Wetzels, W. L. and Mascaró, J. (2001). The typology of voicing and devoicing. *Language*, *77*(2), 207-244.

Wheeler, M. W. (2005). Talk presented at the University of Manchester. The 13th Manchester Phonology Meeting.

Wiese, R. (1996). *The Phonology of German*. Oxford: Calderon Press.

Wonderly, W. L. (1951). Zoque II: Phonemes and morphophonemes. *International Journal of American Lingusitics*, *17*(2), 105-123.

Wright, R. (1996). *Consonant Clusters and Cue Preservation in Tsou*. Ph.D. dissertation. UCLA.

Yoshioka, H., Löfqvist, A. and Collier, R. (1982). Laryngeal adjustments in Dutch voiceless obstruent production. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics*, *16*, 27-35.