The Voice-bar After Closure of Coda Consonants in the Speech of Young Children

by

Won Ron Cho

S.B., E.E., Massachusetts Institute of Technology (2007)

Submitted to the Department of Electrical Engineering and Computer Science

in Partial Fulfillment of the Requirements for the Degree of

Master of Engineering in Electrical Engineering and Computer Science

at the Massachusetts Institute of Technology

June, 2008

© 2008 Won Ron Cho. All rights reserved.

The author hereby grants to M.I.T. permission to reproduce and
to distribute publicly paper and electronic copies of this thesis document in whole and in part in
any medium now known or hereafter created.

Author

Department of Electrical Engineering and Computer Science

May 23, 2008

Certified by

Stefanie Shatuck-Hufnagel, Ph.D.
Principal Research Scientist
Thesis Supervisor

Accepted by

Arthur C. Smith
Professor of Electrical Engineering
Chairman, Department Committee on Graduate Theses
This page is intentionally left blank.
ABSTRACT

The purpose of this study is to gain some insight into the speech acquisition process and articulator development of young children whose mother-tongue is American English. The presence of voice-bars after the closure of voiceless coda consonants in consonant-final words from early child speech is not consistent with the idea that voice-bars are important for distinguishing voiced consonants from voiceless consonants. This study examined this potential deviation from the adult speech model and asked whether children become better at producing voice-bars after the appropriate set of voiced consonants as they become older. The consonant-final words chosen for this study are bug, cup, duck, and tub, in recordings of the speech of 5 children (ages 2;6-3;2). The recordings were taken over the period of 6 months. The results show that the children do become better at producing voice-bars after the selected set of voiced consonants as they become older. This voice-bar production pattern suggests that these children at some point have realized that voice-bars are important for distinguishing voiced consonants from voiceless consonants.

Thesis Supervisor: Stefanie Shattuck-Hufnagel, Ph.D.
Title: Principal Research Scientist, MIT Research Laboratory of Electronics
Acknowledgments

I want to express my deep appreciation and thanks toward my thesis advisor, Stefanie Shattuck-Hufnagel. I have benefited from her invaluable knowledge, patience, and encouragement. Her enthusiasm and passion for speech research has truly been an inspiration. Working with her has been a pleasure and an honor. I would like to thank members of the Speech Communication Group who have helped me feel at home at the group, as well as a former Speech Communication Group member, Annika Imbrie, whose corpus of child/mother speech makes this research possible. I would also like to thank Anne Hunter, whose door was always open whenever I needed advice. I am also very grateful to my academic advisor, Professor Joel Voldman for his excellent guidance.

Finally, I would like to dedicate my thesis to my family, to whom I owe my greatest gratitude. I am indebted to them for their love, moral support, and their unconditional faith in me.
This page is intentionally left blank.
# Table of Contents

List of Figures.......................................................................................... 9
List of Tables............................................................................................... 10
1. Introduction............................................................................................ 11
2. Background............................................................................................. 13
   2.1. Phonology....................................................................................... 13
      2.1.1. Voice Onset Time (VOT).......................................................... 13
      2.1.2. Voiceless vs Voiced................................................................. 13
      2.1.3. Voiceless Consonants............................................................... 13
      2.1.4. Voiced Consonants..................................................................... 14
      2.1.5. Voice-bar.................................................................................. 14
   2.2. The Speech of Children........................................................................ 16
   2.3. The Speech of Caretakers................................................................. 16
3. Literature Review.................................................................................... 18
   3.1. Language Acquisition....................................................................... 18
      3.1.1. Phonological Developmental Order.......................................... 18
      3.1.2. Maturation vs Phonological Understanding................................ 21
      3.1.3. Parental/Environmental Influences............................................ 21
      3.1.4. Deviation from Adult Model...................................................... 22
   3.2. Laboratory induced speech and clear speech...................................... 23
   3.3. Cues to Voicing............................................................................... 23
      3.3.1. Cues in the Time Domain........................................................ 24
      3.3.2. Frequency Domain Cues........................................................ 25
      3.3.3. Lisker's Summary of Cues......................................................... 26
4. Experimental Questions......................................................................... 28
5. Predictions............................................................................................. 29
6. Methods.................................................................................................. 30
   6.1. Data................................................................................................. 30
   6.2. Tools............................................................................................... 31
   6.3. Analysis........................................................................................... 32
List of Figures

Figure 1. Spectrogram of /pa/ ............................................................................ 15
Figure 2. Spectrogram of /ba/ ............................................................................ 15
Figure 3. Vowel space “spreading” ..................................................................... 17
Figure 4. The Sander Norm ............................................................................... 19
Figure 5. A child’s utterance of the word bug .................................................. 33
Figure 6. A second example of a child’s utterance of the word bug ................... 34
Figure 7. An example of a child’s utterance of the word tub .............................. 34
Figure 8. A second example of a child’s utterance of the word tub .................... 35
Figure 9. An example of a child’s utterance of the word duck ............................ 35
Figure 10. An example of a child’s utterance of the word cup ............................ 36
Figure 11. Spectrogram of word bug ................................................................. 37
Figure 12. Waveform and Spectrogram: is a voice-bar present ......................... 37
Figure 13. Preliminary results ............................................................................ 39
Figure 14. Frequency of voice-bar over time: Bug ............................................ 41
Figure 15. Frequency of voice-bar over time: Tub ............................................ 42
Figure 16. Frequency of voice-bar over time: Duck ........................................... 42
Figure 17. Frequency of voice-bar over time: Cup .......................................... 43
List of Tables

Table 1. Summary of Voiceless and voiced English Consonants ..........................................14
Table 2. Shriberg’s Early-8, Middle-8, and Late-8.................................................................19
Table 3. The child subjects and their ages............................................................................30
Table 4. The four words examined in this study ....................................................................31
Table 5. Results from Caretaker 1.........................................................................................43
Table 6. Results from Caretaker 2.........................................................................................44
Table 7. Results from Caretaker 3.........................................................................................44
Table 8. Results from Caretaker 4.........................................................................................44
Table 9. Results from Caretaker 5.........................................................................................44
Table 10. Results from Caretaker 6......................................................................................45
Table 11. Results from Caretaker 7......................................................................................45
Table 12. Results from Caretaker 9......................................................................................45
Table 13. Results from Caretaker 10....................................................................................45
Table 14. Results from Caretaker 11...................................................................................46
1. Introduction

As children begin to walk, they have to learn to control their limbs and retain their balance. They may also develop certain gaits, as they mimic their caretakers. Child speech acquisition may be a very similar process. Children begin by babbling, testing out their articulators and learning to control their vocal tracts to produce target sounds. Lacking the experience and motor control abilities that adults have, children may have trouble controlling their vocal folds and articulators as they learn to speak.

In Annika Imbrie's child speech database (Imbrie, 2005), the utterances by the young subjects sometimes departed from the expected. An example of those acoustic-phonetic deviations manifested in the tokens from the Imbrie database is the presence of voice-bars after the closure of the voiceless coda consonant. The duration of the voice-bar (for description of voice-bar, see next chapter) has been referred to as a primary cue to voicing (Nearey, 1997). One can envision that the absence/presence of a voice-bar after a consonant’s closure also serves as a strong cue to voicing. As a result, voice-bars should be more reliably present after the closure of voiced consonants than after the closure of voiceless consonants. The type of speech behavior observed in the Imbrie’s child speech recordings is a departure from the conventional adult speech model, an interesting phenomenon that deserves further investigation for the light it may shed on the acquisition process.

One possible explanation for the presence of voice-bars during the closure of voiceless coda consonants is that the children do not have full control of their vocal folds and unconsciously produce the low-frequency voice bars after the closure of the voiceless coda consonants. On this view, as the children attempt to transition from a vowel to the closure of a consonant, their vocal folds still vibrate after closure. As these children develop better motor control over their articulators, they are expected to produce fewer voice-bars during the closure of voiceless coda consonants and more voice-bars during the closure of voiced coda consonants.

Another possible explanation for the presence of voice-bars during the closure of voiceless coda consonants is that the children have not yet become metaphonologically aware that
voicing during closure is associated with voiced consonants but not with voiceless consonants. On this view, as these children develop better motor control over their articulators, they are also expected to produce fewer voice-bars during the closure of voiceless coda consonants and more voice-bars during the closure of voiced coda consonants.

Regardless what causes the children to produce voice-bars after the closure of both voiced and voiceless coda consonants, the appearance of voice-bars after the closure of voiceless coda consonants in consonant-final words from early child speech is not consistent with the theory of voice-bars being a strong cue to voicing (as inferred from Nearey's endorsement of voice-bar duration as a primary cue to voicing, 1997). This study will also report on voice-bar appearance after the coda closure of 4 consonant-final words (bug, cup, duck, and tub) in the speech of 5 children (ages 2;6-3;2) over the period of 6 months. This study will examine the laboratory induced speech of the caretakers of these children and treat their voice-bar production pattern as the adult model. Hopefully, results from this study will provide some insight into the acquisition of the acquisition of the feature [+voice/-voice] in the coda position, and of the child's ability to produce adult-like cues to this feature contrast.
2. Background

2.1. Phonology

2.1.1. Voice Onset Time (VOT)

Voice Onset Time (VOT) is an acoustic parameter first mentioned in 1964 by Lisker and
Abramson (Lisker & Abramson, 1964). VOT is the duration of the period of time between the
release of a stop consonant and the beginning of glottal vibrations in the following vowel. This
parameter is useful in speech research since it has been proposed as a way to classify a stop
consonant as being voiced or voiceless.

2.1.2. Voiceless vs Voiced

The [+voice]/[-voice] feature of a consonant has some degree of dependence upon the state of
the glottis during the formation of the supraglottal constriction. The vocal folds can be positioned
in an array of different configurations (Russell, 1997). Some of the possible sounds produced
with these glottal states are voiced, voiceless, murmur-like, or whisper-like. When not
whispered, vowels are generally voiced while consonants can be either voiced or voiceless. In
American English, the feature [+voice] and [-voice] dictates the difference between minimal
pairs such as: bug/buck and bob/bop.

2.1.3. Voiceless Consonants

While producing a sound, as air is expelled from the lungs, if the vocal folds are lax and pulled
slightly apart (not as far apart as they would be if the speaker were taking a silent breath), the
vocal folds will not vibrate. The air being pushed from the lungs will pass straight through the
glottis without producing regular vibration via the Bernoulli effect. Consonants produced in this
manner are generally voiceless and have the feature [-voice]. A listing of voiceless consonants in
English are is given in Table 1.
Table 1. Voiceless and voiced English Consonants are summarized in this table. The consonants are also shown in their minimal pair positions.

<table>
<thead>
<tr>
<th>Place of Articulation</th>
<th>Manner of Articulation</th>
<th>Voiceless consonant</th>
<th>Voiced Consonant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilabial</td>
<td>Stop</td>
<td>[p] (pig)</td>
<td>[b] (big)</td>
</tr>
<tr>
<td>Alveolar</td>
<td>Stop</td>
<td>[t] (tent)</td>
<td>[d] (dent)</td>
</tr>
<tr>
<td>Velar</td>
<td>Stop</td>
<td>[k] (kill)</td>
<td>[g] (gill)</td>
</tr>
<tr>
<td>Postalveolar</td>
<td>Affricate</td>
<td>[ʃ] (choose)</td>
<td>[ʒ] (juice)</td>
</tr>
<tr>
<td>Labio-Dental</td>
<td>Fricative</td>
<td>[ʃ] (fan)</td>
<td>[v] (van)</td>
</tr>
<tr>
<td>Interdental</td>
<td>Fricative</td>
<td>[θ] (thigh)</td>
<td>[ð] (thy)</td>
</tr>
<tr>
<td>Alveolar</td>
<td>Fricative</td>
<td>[s] (sip)</td>
<td>[z] (zip)</td>
</tr>
<tr>
<td>Postalveolar</td>
<td>Affricate</td>
<td>[ʃ] (pressure)</td>
<td>[ʒ] (pleasure)</td>
</tr>
</tbody>
</table>

2.1.4. Voiced Consonants

When producing a sound, as air is expelled from the lungs, if the vocal folds are brought sufficiently close together and are stiff enough, they will vibrate. According to the Myoelastic-Aerodynamic Theory of Phonation (proposed by van den Berg, 1958), the vocal folds are pushed open by pressure from the lungs and forced shut via the Bernoulli Effect, as the air rushing through the inter-fold space lowers the pressure in this region. The resulting cycles of glottal opening and closing cause the vibration of the vocal folds. As a result, the produced sound is voiced. Consonants produced with this type of glottal configuration are voiced and generally have the feature [+voice]. A summary of voiced consonants in English is given in Table 1.

2.1.5. Voice-bar

The voice-bar, also known as pre-voicing, can be seen on spectrograms as bands of very low frequency at around 200 Hz for adults, much lower than the F1 of a typical male utterance of the sound [i]. The voice-bar occurs during periods when the vocal tract and velum are closed and there are no direct paths for air to be expelled from the lungs. In other words, some sort of closure is occurring in the vocal tract, such as the closure that is made to articulate a consonant. If the vocal folds are still vibrating at this point, due to articulatory manipulations that allow air flow to continue through the glottis, the voiced energy can pass through the tissue of the vocal tract walls and emanate from the speaker to the surroundings. The tissue of the vocal tract acts like a low pass filter, filtering out the high frequency components of the vocal fold vibrations.
This phenomenon is called pre-voicing because it often accompanies a voiced consonant by appearing after the consonant closure and before the release of the consonantal burst. Figures 1 and 2 show the spectrograms of the syllables /pa/ and /ba/ as uttered by the author. The recording was done on Praat with a sampling frequency of 44100 Hz. No voice-bar is seen before the burst of the /p/ in /pa/ but a voice-bar is present before the release of /b/ in /ba/. This is expected since /p/ is a voiceless consonant while /b/ is a voiced consonant.

Figure 1. Spectrogram of /pa/. No voice-bar is seen before the release of the consonant /p/.

Figure 2. Spectrogram of /ba/. A voice-bar is seen before the release of the consonant /b/.

The presence of a voice bar has been suggested as a cue, perhaps even a necessary cue, to the perception of the feature [+voice] in a stop consonant in adult American English speech. Our
focus in this study is the degree to which it occurs in child speech, and in the speech of the adult caretakers of the same children.

2.2. The Speech of Children

Language acquisition is one of the most important aspects of children’s development. Like learning to walk during which children explore the use of their legs and try to maintain their balance, learning to speak is a gradual process in which the children begin by testing out their articulators and exploring the types of sounds they can produce through babbling.

Babbling allows children to try out the various components of their vocal tracts and begin to gain control over them. Children will have to learn to produce the consonants that are present in their first language by modifying their vocal tracts at different points of articulation. They will also have to tune the stiffness of their vocal folds to alter their glottal states when changing from voiced to voiceless sounds.

2.3. The Speech of Caretakers

There exist theories suggesting that child-directed speech may have developed for evolutionary reasons (Falk, 2003). Like modern human children, early hominid infants developed important motor skills much later in life when compared to other primates. While their bipedal mothers foraged for food, these hominid infants could not cling onto their mothers like baby chimpanzees could. As a result, these bipedal primate mothers would often have to put their offspring on the ground while gathering food. Child-directed vocalization would be a means for these mothers to reassure their offspring and control their behavior.

Recent studies have shown that child-directed speech is more effective than adult-to-adult speech at attracting the attention of children, who seem to prefer the sound of child-directed speech (Reschke, 2002). Perhaps child-directed speech not only developed as a means to help bipedal mothers control their offspring but also as a means for parents to direct children’s attention to speech, facilitating the language acquisition process.
Child-directed speech or “motherese” (Ferguson, 1964) is usually delivered with exaggerated pitch contours and song-like qualities that are not found in normal adult-to-adult speech (Fernald, 1989). This type of slow and deliberate speech production often contains carefully enunciated consonants and spread out vowel space (Kuhl et al., 1997) illustrated in Figure 3.

Figure 3. This figure (adapted from Kuhl et al., 1997) illustrates how the vowel space in child-directed speech differs from the vowel space in adult-directed speech. The solid squares mark the vowel space of child-directed speech while the circles mark the vowel space of the adult-directed speech. There appears to be a “spreading” of the vowel space.
3. Literature Review

3.1. Language Acquisition

Research in the area of child speech is primarily concerned with the language acquisition process of children. Two main issues that are widely investigated in child speech studies are: 1) the developmental process of the acquisition itself and 2) the factors that influence progression of acquisition. The study of the developmental process deals with phonological, morphological, and syntactic maturation. The study of the influences on acquisition is concerned with how factors such as the children’s environment (i.e. parental interaction) affect their language development. This study focuses on a different aspect of language development: children’s acquisition of cues to feature contrasts, specifically voicing in coda stop consonants. The results may provide some groundwork for future studies on the effects of child-directed speech on the acquisition of cues to voicing.

3.1.1. Phonological Developmental Order

When acquiring language, children learn step-by-step, gradually building on what they have acquired as they try to approach adult proficiency. The research on phonological development is crucial to understanding how children acquire language.

Certain phonemes manifest themselves in babbling of children earlier than other phonemes. In 1983, J. Locke summarized three studies on 131 children in English-speaking environments and found that fricatives, affricates, and liquids were relatively rare in the infants’ babbling (Locke, 1983). The Sander Norm (Sander, 1972), which documents the typical order in which English phonemes are acquired, is often used to judge whether or not a child is acquiring language at the typical rate. The Sander Norm is shown in Figure 4. In a more recent publication, Shriberg divides the consonantal phonemes of English into 3 categories: the early-8, the middle-8, and the late-8 (Shriberg, 1993). The general trend observed by Shriberg is that labials stops tend to be
acquired first and affricates tend to be acquired later. The Shriberg categorization is illustrated in Table 2.

Figure 4. The Sander Norm shows that labials like /p/ and /m/ are acquired early on while affricates like /ʃ/ are acquired much later.

Table 2. The 24 English consonantal phonemes are divided into 3 categories according to the order in which they are acquired.

<table>
<thead>
<tr>
<th></th>
<th>Early 8</th>
<th>Middle 8</th>
<th>Late 8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mbjrwdp</td>
<td>tkgndjld</td>
<td>sdlzr</td>
</tr>
</tbody>
</table>

The term acquisition can be ambiguous. A child who can perceive certain phoneme but cannot yet produce that phoneme has in some sense acquired the phoneme. A child who can produce the sound at will has also in some sense acquired that phoneme.
While most published findings on developmental orders organize the phonemes by place and manner of articulation, there have been some studies addressing the acquisition of voicing in consonants. A study by Kager and colleagues shows evidence that English speaking children tend to make more devoicing errors than voicing errors when trying to produce target consonants (Kager et al, 2005). This may indicate that for children, producing voiced consonants is more difficult than producing voiceless consonants.

The gradual development of babbling and its transition into words in children’s vocabulary is described in the work of Levitt and colleagues, who report results from a study of French and English learning children (1992). Physiological maturation of the muscles of the jaw, tongue, and those in the larynx plays a significant role in babbling production. The sounds in children’s babbling repertoire will later develop into parts of words in the children’s vocabulary.

In 1991 Vihman and de Boysson-Bardies describes a series of experiments they performed that mapped out the developmental order of phonemes in children (Vihman & de Boysson-Bardies, 1991). They sampled babbling speech from children whose parents are speakers of several dissimilar languages (including French, Japanese, English, and Swedish). They studied children over the period from 10 months to 17 months. Vihman and de Boysson-Bardies found that earliest babbling shows little differentiation among languages. Labial stops were the most common utterances.

One proposed reason behind the frequency of labial stop consonants is formulated by MacNeilage (1998). His theory is often dubbed as the jaw oscillation theory. The single action of oscillating the mandible closes and opens the vocal tract. As long as the tongue remains in a neutral position while the jaw oscillates, children can create bilabial consonants. If the tongue tip is in the front of the mouth while the jaw oscillates, children can easily produce alveolar consonants. MacNeilage’s proposal offers an explanation for the pattern of babbling found by Vihman, suggesting that the patterns of babbling are a result of the degree to which children have control over their articulators. Children first produce labial stops without having to configure the articulators in any complex positions. Later, they move onto other types of consonants that require the vocal tract to be configured in non-neutral positions.
3.1.2. Maturation vs Phonological Understanding

The existence of a reliable phonological development order in children, as discussed in section 3.1.1., suggests that physiological maturation of the muscles of the vocal system plays a vital role in linguistic maturation. However, other aspects of this maturation may not align perfectly in time with the development of motor control. For example, children’s understanding of language and their ability to perceive phonemic contrasts develops well in advance of their ability to produce them. One classic example of this is the “/fis phenomenon” described by Berko and Brown in their 1960 book. Berko and Brown found that one child in their study could perceive the difference between the phonemes /s/ and /ʃ/ but was unable to produce the latter. When referring to his plastic toy fish, the child uses the utterance /fis/. The child understood the difference between the two phonemes but was not capable of controlling his articulators to produce /ʃ/ or was not able to use /s/ as a coda.

A more contemporary example (of children’s ability to perceive phonemic contrast developing well in advance of their ability to produce them) is presented in a yet to be published paper by Song and Demuth (to appear). Song and Demuth studied English-speaking children between the ages of 1;1-2;6. In their findings, the children lengthened vowels to compensate for coda consonants that they were unable to produce, showing that they recognize that CV utterances are different from CVCs.

3.1.3. Parental/Environmental Influences

There have been a substantial number of studies detailing the environmental factors that influence children’s language development process. While early babbling has been found to show little differentiation among languages (Vihman & de Boysson-Bardies, 1991), the influences of parents, caretakers, and the environment become more apparent later in the acquisition process.

In their aforementioned 1992 paper, Levitt and colleagues found evidence that later babbling becomes more influenced by the environment. The frequency of the phonemes appearing in the children’s babbling becomes more language-specific. The initial similarities across languages slowly begin to change as the vowel spaces of French and English children begin to shift towards the vowel spaces of adults who speak their respective languages.
Marilyn Vihman (1996) also noted that later on, in a similar fashion, consonants such as [h] that do not appear in French manifest themselves less frequently in French babbling and are subsequently dropped.

Researchers have inferred that the environment's influence "tunes" children to the phonemes of the language to which they have been exposed. For example, 4-month old Japanese infants have been shown to have the ability to distinguish between the sounds /r/ and /l/ (Werker and Tees, 1984; Werker, 1991; Kuhl et al., 1992, 1997; Kuhl, 2004). However, by the time these children are 1 year old, they can no longer perceive the distinction between these two sounds. Since the Japanese language does not distinguish between these two sounds, the children's mechanisms for perception have been tuned instead to phonemes prevalent in their environment.

3.1.4. Deviation from Adult Model

While practice makes perfect, Annika Imbrie's studies have shown that even children at the age of 2-3 still do not have articulatory patterns that match those in the adult model (2005). In her thesis, Annika studied the development of word onset stop consonants in children. Her results showed that in this word position the children's positioning of the primary articulators seems to be correct, but they have yet to learn more fine grained aspects of tongue body adjustment during stop consonant production. She also found that the children still needed to perfect their consonant release mechanism. Another of Imbrie's findings was that the children seem to have trouble controlling their vocal fold stiffness and achieving the correct vocal tract configuration before a burst release. She further noted that for both voiced and voiceless stop consonant production, the children use a high subglottal pressure, which indicated that they have are still learning to control the respiratory component of speech production.

These deviations from the adult model may be a result of exposure to child-directed speech. Child-directed speech is characterized by intonation patterns that are not characteristic of normal adult-to-adult speech. To produce this type of speech, the speaker is careful with pronunciation. As a result, this type of speech may be richer in cues to segmental feature contrasts. If infants are trying to approach any model during their early stages of speech acquisition, they would most
likely be trying to approach what they hear the most frequently (child-directed speech) or whatever type of speech most strongly attracts their attention.

### 3.2. Laboratory induced speech and clear speech

The adult speech used in this study is laboratory induced speech. Speech induced in a laboratory setting is also characterized by clear pronunciation. The experimenters would like to collect data that is as free from noise as possible. Therefore, unless the speakers are requested to produce conversational speech, the speakers are usually asked to speak clearly into the microphone.

Clear speech is characterized by a decrease in speaking rate (Picheny et al, 1986) and is thus very similar to laboratory induced speech. Unlike in conversational speech, in clear speech, vowels are less likely to be reduced and almost all coda-consonants are released. The stop consonants in clear speech also have higher RMS intensities than their conversational speech counterparts. It is possible that the adult samples in the Imbrie Corpus examined in this study reflect to some degree the child-directed speech that the child speakers were accustomed to hearing, but full exploration of this question must await further investigation.

### 3.3. Cues to Voicing

One of the distinctive features that contrast segments in language is [+voice/-voice]. The cues to voicing of consonants usually fall into two broad categories. Cues that are related to durations, signal decay time constants, offset times, and onset times are cues in the time domain. Cues that are found in the frequency components and in the harmonics are cues in the frequency domain. Results from different studies find evidence suggesting some cues have a greater perceptual effect than others.

The list of potential cues to voicing is a dynamic pool in which cues are constantly being added and removed as new experimental results emerge. While there are plenty of disputes among researchers about which acoustic phenomena are perceptual cues to voicing used by listeners and which are merely aspects of the speech signal, they generally agree that voicing is characterized by the vibration of the vocal folds during some portion of the speech signal associated with the articulations of the sounds. This generalization is not always true, since the voiced/voiceless
distinction is apparent in whispered speech, a scenario in which vocal fold vibration does not occur.

The conclusion is that in different contexts (whispered speech vs. normal speech), voicing seems to be indicated by different sets of cues. Some cues seem to be more reliably present than others. Other cues are not always present, but when they do appear may further help to disambiguate the [+voice/-voice] aspect of the consonant sound, for example in noisy circumstances.

3.3.1. Cues in the Time Domain

In 1956, Liberman of the Haskins Laboratory published a paper in which he presented what he believed to be the most important cues to consonant voicing. He stated that the timing or tempo of frequency changes in a consonant signal help listeners identify the consonant. The closure duration of the consonant (not the voice onset time) seems to be a strong cue that distinguishes consonant classes from each other, in that voiceless consonant closure durations tend to be shorter than their voiced counterparts.

Two years later, Liberman and colleagues identified several other cues to voicing (1958). One interesting observation they made was that in whispered speech, there were perceptible differences between voiced and voiceless consonants. Because the vocal folds do not vibrate regularly during whispered speech, the presence of vocal fold vibration is not the only cue for the voicing of consonants. In this new publication, Liberman (et al.) added voice onset time (VOT) as a cue to voicing in consonants. They discovered that the voice onset time of the first formant frequency for a voiced consonant was shorter than the VOT of the first formant frequency for a voiceless consonant.

In support of VOT as a cue to voicing, Abramson and Lisker (Abramson & Lisker, 1964) proposed that VOT is the most reliable feature that distinguishes voiced phonemes from their voiceless counterpart. In English, voiced phonemes are characterized by a short VOT while the voiceless ones are characterized by a longer VOT, often accompanied by aspiration noise, particularly for stops in word-onset or pre-stressed-vowel positions.

Besides closure duration, durations of other components of the speech signal have also been found to serve as cues to voicing. Lawrence J. Raphael performed a series of categorical
perception experiments on the perception of consonants in word final positions (1975). The variable in this experiment was the duration of the nasal before the coda consonant. Raphael found that the nasal durations were greater before voiced consonants like /d/ than in voiceless consonants like /t/. He also found that the listener's perception of voiced/voiceless shifted rapidly as the duration of the nasal changed. As a result, he concluded that nasal duration before a consonant is a strong cue to voicing.

Five years later, Raphael identified vowel duration as a cue to voicing for coda consonant (1980). He found that the listener's perception of voiced/voiceless shifted rapidly as the duration of the vowel before the coda consonant changed. Therefore, he concluded that vowel duration before a consonant is also a strong cue to voicing. In both of these studies, the cue to the value of the [voice] feature of the target segment was found in a part of the speech signal more strongly associated with the preceding phonemic segment. This observation is consistent with the view that feature cues are not only variable in their selection, but also variable in their location.

3.3.2. Frequency Domain Cues

Frequency domain cues manifest themselves in the spectrogram of the speech signals of interest. For example, the transitions of the lowest formant frequency (F1) as the articulation moves from the consonant into the following vowel are often cited as a cue to voicing in consonants, and the presence of aspiration may indicate lack of voicing in the consonant.

In the 1958 Liberman publication, aspiration is mentioned as a cue indicating the lack of voicing in English. While in some languages such as Hindi, voiced consonants such as /b/ can be followed by aspiration, in English, aspiration is present in voiceless consonants (at least in some contexts) but not in voiced consonants.

Liberman further noted that the behavior of F1 is a strong indicator of whether or not a sound is voiced. F1 must start extremely low at the beginning of a vowel after a consonant in order for the consonant to be perceived as voiced.

The voice-bar or prevoicing has also been identified as a cue to voicing in a consonant. Some proponents of the voice-bar believe that the presence of a voice-bar between the closure and the
release of a consonant deems the consonant voiced. As a result, many classify prevoicing as a strong cue to voicing in consonants. Terrance Nearey (1997) endorse the idea that the duration of the voice-bar before a consonant is the primary cue to voicing. In other words, voice-bar duration is a primary cue to voicing. Nearey defined primary cues as "those that are expected to have a large effect on the probability of responses to the category in question" (Nearey, 1997). One can infer that the absence/presence of a voice-bar after a consonant’s closure also serves as a strong cue to voicing.

3.3.3. Lisker’s Summary of Cues

In a 1986 paper, Leigh Lisker presented a summary of all potential cues to voicing for English stop consonants in trochees (a stressed syllable followed by an unstressed one). The minimal pair that he focused on was {/b/ and /p/}. He gave a summary of the following cues to voicing:

1) Before Closure
   -vowel duration
   -duration of F1 transition
   -F1 offset frequency
   -F1 transition offset time
   -timing of voice offset
   -F0 (fundamental frequency)
   -time constant of signal (decay time)

2) During Closure
   -closure duration
   -duration of glottal signal
   -glottal signal intensity

3) After Closure
   -release burst intensity
   -Voice onset time (VOT)
   -Onset of F1 transition
   -F1 onset frequency
   -F1 transition duration
-F0 (fundamental frequency) contour

Lisker also noted that it is generally agreed that vocal fold vibration characterizes a speech signal as voiced. This large category of vocal fold vibration perhaps includes voice-bars. Similarly, Lisker stated that a laryngeal buzz is present during closure (during which the vocal cords are held quite close together and the flow of air through the mouth is reduced) is another generally agreed upon cue for voicing.
4. Experimental Questions

This study focuses on the reasons behind the appearance of voice-bars after the closure of voiceless coda consonants in CVC words in the speech of young children. The appearance of voice-bars after the closure of both voiced and voiceless coda consonants in the speech of these children indicates a deviation from the adult model. To gain insight into why this deviation occurs, this study examines the following sets of questions:

1) Is the appearance of voice-bars in the speech of children random? Do the voice-bars manifest themselves more often after the closure of voiced coda consonants and less often after the closure of voiceless coda consonants?

2) Is there a pattern to appearance of voice-bars that develops as time progresses? In other words, do more practice and the maturation of the articulators play a part?

3) What sort of patterns (of the manifestation of voice-bars) exists in the speech of the primary caretakers of these children? Do they also consistently produce voice-bars after the closure of voiced coda consonants and not after the closure of voiceless coda consonants?
5. Predictions

The predictions addressing each of the experimental questions are detailed in this section. Findings from literature and preliminary studies provide the reasoning behind these predictions.

1) **The appearance of voice-bars in the speech of young children will not be random.** Voice-bars will manifest themselves more often after the closure of voiced coda consonants and less often after the closure of voiceless coda consonants. Children as young as 1 to 6 months can perceive the phonemic contrast between /p/ and /b/ (Eimas, 1971). Children of this age have been found to systematically lengthen vowels to compensate for coda consonants they are unable to produce (Song & Demuth, to appear). By inference, the appearance of voice-bars in the speech of young children is also systematic. The possibility that the voice-bar is an important voicing cue suggests that voice-bars should appear more reliably after the closure of voiced coda consonants. Preliminary analysis (see Results) of the speech of one child from the Imbrie corpus seem to indicate that voice-bars appear more often after the closure of voiced coda consonants and less often after the closure of voiceless coda consonants.

2) **There will be a pattern in the appearance of voice-bars that changes as the sessions progress.** If the children are producing voice-bars after the closure of voiceless consonants as a result of their inability to control their glottal state, then they are equally likely to produce voice-bars after the closure of both voiceless and voiced consonants. As they become older and more experienced, the voice-bars will become less frequent after the closure of voiceless consonants and more frequent after the closure of voiced consonants.

3) **The primary caretakers will produce more voice-bars after the closure of voiced coda consonants and will produce fewer voice-bars after the closure of voiceless coda consonants.** This would be consistent with the voice-bar being a strong cue to voicing.
6. Methods

6.1. Data

For this study, two types of speech are needed: 1) speech of children going through the language acquisition process and 2) clear speech/laboratory induced speech of the caretakers of such children. Annika Imbrie’s (2005) child speech corpus contains both the speech of young children as well as laboratory induced speech produced by adults. Imbrie’s collection of speech contains recordings of words uttered by 10 young children whose speech development was documented over the course of six months. Roughly one recording session was conducted each month for each child. Imbrie first read stories to the children, eliciting target words from them after asking them questions like (“She drinks from . . .”) or (“What’s the girl’s name?”). She then played with the children, asking them to name toys before she gave the toys to the children. After the play sessions, the children were prompted to name the toys before putting them away. The corpus also contains laboratory induced speech of the caretakers of 10 of these children. These adults were shown a set of pictures and prompted to speak the same set of target words. These words are available as *.wav sound files.

The speech of 5 of these 10 children and the speech of all 10 of the caretakers was used in this analysis. Motherese or child-directed speech was not available. However, since laboratory induced speech is often produced clearly, and thus may be very similar to child-directed speech, the laboratory induced speech produced by the caretakers of the children was used instead. Information about the 5 children is summarized in Table 1, as adapted from Imbrie (2005).

Table 3. The 5 chosen children will be referred to as C01, C09, C10, C02, and C07. Since Imbrie’s study took the course of 6 months, almost every child has 6 sessions, one for each month. However, C01 and C02 were only available for 5 sessions.

<table>
<thead>
<tr>
<th>Child</th>
<th>Date of Birth</th>
<th>Date of Session 1</th>
<th>Age at Session 1</th>
<th>No. of Sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>C01</td>
<td>09/20/2000</td>
<td>04/09/2003</td>
<td>2;6</td>
<td>5</td>
</tr>
<tr>
<td>C09</td>
<td>12/18/1999</td>
<td>02/21/2003</td>
<td>3;2</td>
<td>6</td>
</tr>
<tr>
<td>C10</td>
<td>12/04/1999</td>
<td>02/25/2003</td>
<td>3;2</td>
<td>6</td>
</tr>
<tr>
<td>C02</td>
<td>07/20/2000</td>
<td>02/22/2003</td>
<td>2;7</td>
<td>5</td>
</tr>
<tr>
<td>C07</td>
<td>02/22/2000</td>
<td>03/10/2003</td>
<td>3;0</td>
<td>6</td>
</tr>
</tbody>
</table>
The words from the Imbrie database that were examined in this study are *cup, tub, duck, and bug*. All four words are consonant-final and have the same nucleus vowel, the tense [A] sound. The properties of their coda consonants are summarized in the table below.

Table 4. Properties of the coda consonants of the words *cup, tub, duck, and bug*. For each of the 5 children and the 9 adults, the aforementioned words were examined.

<table>
<thead>
<tr>
<th>Word</th>
<th>Coda Consonant</th>
<th>Place of Articulation</th>
<th>Manner of Articulation</th>
<th>Voiced/Voiceless</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cup</td>
<td>/p/</td>
<td>Bilabial</td>
<td>Stop</td>
<td>Voiceless</td>
</tr>
<tr>
<td>Tub</td>
<td>/b/</td>
<td>Bilabial</td>
<td>Stop</td>
<td>Voiced</td>
</tr>
<tr>
<td>Duck</td>
<td>/k/</td>
<td>Velar</td>
<td>Stop</td>
<td>Voiceless</td>
</tr>
<tr>
<td>Bug</td>
<td>/g/</td>
<td>Velar</td>
<td>Stop</td>
<td>Voiced</td>
</tr>
</tbody>
</table>

**6.2. Tools**

Praat was the primary software tool employed in this study. Praat is a phonetics analysis software developed by Paul Boersma and David Weenink of the University of Amsterdam (2005). This tool can read a *.wav file, and calculate and display its spectrogram. Praat also allows the user to play selected sections of a spectrogram, allowing the user to simultaneously visually analyze the frequency components of a sound segment and also analyze the segment by ear.

Statistical analysis of the data and the production of the figures in the Results chapter were carried out using Microsoft Excel.
6.3. Analysis

The type of analysis made needs to reflect the experimental questions asked in the fourth section. The analysis determines the plausibility of each of the predictions made in the fifth section.

6.3.1. Analysis: Child Speech

For each of the sessions recorded and from all of the tokens produced by each child, 4 target words (bug, cup, duck, tub) were examined. On average, in each session, a child produced 15-20 utterances of each word. For the analysis, the *.wav file of each utterance is loaded into Praat, displaying the corresponding waveforms and spectrogram. The region between the closure and release of the coda consonant is the segment of interest. If a voice-bar exists, it will manifest itself on the spectrogram in the form of low-frequency noise resembling F1 (see Figure 5). Even though there is not a path for air to exit through the vocal tract or for sound to radiate from the lips, the slack walls of the vocal tract allow enough air to flow through the glottis for continuation of vibration. The energy escaping into the surroundings by the means of skin vibration is of low frequency because the low pass filtering effects of the tissues of the vocal tract. For each word in each session for each child, the number of tokens for which a voice-bar appears (after the closure of the coda consonant) is compiled. Dividing the number of appearances by the total number of utterances for that particular word gives an estimate of the frequency of voice-bar appearance (for the given word in a specific session).

6.3.2. Analysis: Speech of Caretakers

Only one session was available for each of the 10 caretakers. For each caretaker, the same 4 words (bug, cup, duck, tub) were analyzed. On average, a caretaker produced 7-9 utterances of each word. The *.wav file of each word was read into Praat to display the corresponding
waveform and spectrogram. Similar to the analysis of child speech, the period between the closure and release of the coda consonant is the segment of interest. If a voice-bar is present, it can be seen on the spectrogram in the form of low-frequency noise resembling F1. For each word spoken by the caretakers, the number tokens which showed a voice-bar (after the closure of the coda consonant) is compiled. Dividing the number of appearances by the total number of utterances for that particular word gives an estimate of the frequency of voice-bar appearance (for the given word spoken by each caretaker).

6.3.3. Criteria for voice-bar

When there is not path for air to exit through the vocal tract, air flow into the expandable vocal tract is reduced and the resulting vocal fold vibration is of low amplitude. In addition, energy is escaping by the means of low frequency skin vibration, a result of the low pass filtering effects of vocal tract tissue. These low frequency vibrations constitute the voice-bar and can usually be seen on a spectrogram as a very low amplitude F1. In the absence of noise, a voice-bar can easily be identified visually. Examples spectrograms of words with post coda closure voice-bars and words without post coda closure voice-bars are shown in Figures 5 through 10.

![Figure 5. A child’s utterance of the word bug. The voice-bar can clearly be identified in both the spectrogram panel (low frequency vibration) and in the waveform panel (regular time varying function). The voice-bar in this particular case is of rather high amplitude. An epenthetic vowel can be seen after the release of the coda consonant. Epenthesis, the addition of a syllable or sound, has been postulated to facilitate the production of certain speech sounds (Burton, 2007).](image-url)
Figure 6. Another example of a child’s utterance of the word *bug*. The voice-bar can clearly be identified in both the spectrogram panel (low frequency vibration) and in the waveform panel (regular time varying function). In this particular example, the release of the coda consonant is not followed by an epenthetic vowel, something that might facilitate the production of the /g/ sound.

Figure 7. An example of a child’s utterance of the word *tub*. The voice-bar can clearly be identified in both the spectrogram panel (low frequency vibration) and in the waveform panel (regular time varying function).
Figure 8. A second example of a child's utterance of the word *tub*. The voice-bar can clearly be identified in both the spectrogram panel (low frequency vibration) and in the waveform panel (regular time varying function). Note that in this token the voice bar decreases substantially in amplitude during the closure.

Figure 9. An example of a child’s utterance of the word *duck*. The voice-bar is not seen in either the spectrogram panel (no low frequency vibration) or in the waveform panel (lack of regular time varying function). The anticipation of a voiceless velar stop following the vowel may be the cause of the large amount of noise at the end of the vowel. Perhaps this could be a cue to the feature [-voice] for the coda consonant (Shattuck-Hufnagel et al., 2007).
6.3.4. Troubleshooting

The process of deciding whether or not a voice-bar is present during the closure of a consonant is not always straightforward. Whether or not the examples in Figure 5 through Figure 10 have voice-bars in the closure of the coda consonant can be simply decided by examining the region on the spectrogram between the closure of the coda consonant and the release. However, there are situations in which telling if a voice-bar is present may be difficult. Most of the ambiguous cases occur in the word *bug* in which the coda consonant is a voiced velar stop. Figure 11 shows the spectrogram of one such instance of *bug*. 

Figure 10. An example of a child’s utterance of the word *cup*. The voice-bar is not seen in either the spectrogram panel (no low frequency vibration) or in the waveform panel (lack of regular time varying function).
One way to decide if a voice-bar is present is to examine just the region between closure and release by zooming in on the region using Praat. Both the waveform and the spectrogram can be used to determine if a voice-bar is present. When looking at the waveform, if there is no regular time-varying oscillation after the closure of the consonant, the voice-bar is not present. When looking at the spectrogram, if there is low amplitude F1 after the closure of the consonant, the voice-bar is not present. This method is illustrated in Figure 12.

If the waveform shows signs of a voice-bar but the spectrogram does not (or vice versa), the human ear becomes the judge. By listening to the *.wav file on headphones at high volume, the
human judge can decide if a voice-bar is present. If a voice-bar is present, it will sound like low frequency creaking/humming.
7. Results and Discussion

This section will first present preliminary results of a short study of the target words that appear in Session 1 of C01. Next, the final results will be presented to show how they answer experimental questions 1 through 3 posed in Section 4. The implications of these results and how they support or discredit each prediction from Section 5 will be also be discussed. Concluding remarks and possible future work will end this report.

7.1. Preliminary Results

Preliminary results of a short study of the target words that appear in Session 1 of C01 (see Figure below) seem to indicate that voice-bars appear more often after the closure of voiced coda consonants and less often after the closure of voiceless coda consonants. This session took place when the subject was 2 years and 6 months old. C01’s production of voice-bars does not appear to be random but occurs more frequently in the closure of voiced coda consonants. In C01’s utterances of the word bug, only 2 do not have a voice-bar. Neither cup nor duck have any tokens that have voice-bars. If voice-bar presence/absence is an important cue to voicing, the occurrence of a few tokens with no voice bar for [+voice] codas may reflect the results of Kager’s study (Kager et al, 2005): children tend to make more devoicing errors than voicing errors when trying to produce target consonants, suggesting that producing voiced consonants is more difficult than producing voiceless consonants.

Figure 13. Preliminary results of a short study of the target words that appear in Session 1 of C01.
These pilot results provide support for the view that children in this age range have mastered some, but not all, of the presumed adult pattern of cues to the voicing feature of coda consonants in CVC words. Encouraged by these observations, we continued our analysis of 4 additional children and the caretaker adults. Final results for these analyses are presented in the following section.

7.2. Final Results

Figures 14 through 17 below show the plots (for each target word) of the progression of the frequency of voice-bars appearing after the closure of coda consonants over the course of 5-6 sessions for each child. **Experimental question 1 asks whether or not voice-bars appear randomly.** Do voice-bars manifest themselves more often after the closure of voiced coda consonants and less often after the closure of voiceless coda consonants? As seen in Figures 14 through 17, voice-bar appearance is seemingly systematic and occurs more frequently after the coda closure of voiced consonants /g/ and /b/.

**Experimental question 2 asks if the pattern of the appearance of voice-bars changes over time.** In Figure 14, for the cases of C01 and C07, post coda closure voice-bars appear in the word *bug* more frequently in later sessions. In C02’s session 1, only 3 out of 10 utterances of *bug* contained post coda closure voice-bar. In the later sessions for C02, the post coda closure voice-bars appear more frequently. In session 1, 2, 3, 5, and 6, C09 produced post closure coda voice-bars for every utterance of the word *bug*. In session 4, C09 only produced 9 utterances of the word *bug*, only one of which did not contain a post coda closure voice-bar, causing a sharp dip in what would otherwise be a constant trend. A general rising trend is seen in C10’s data. However, in his last session, he only produced a post coda closure voice-bar out in 1 out of 7 utterance of the word *bug*.

In Figure 15, there appears to be a general rising trend in C10 and C02’s post coda closure voice-bar frequencies. The *tub* data of the other 3 children remain relatively constant over the sessions. It may be that for these children, the developmental period during which voice-bars increase in reliability for voiced codas occurred at an earlier age.
In Figure 16 and 17, no obvious trend on the appearance frequency of the post coda closure voice-bar is observed. However, while the other four children seem to produce small numbers of voice-bars after the coda closure of these voiceless consonants, C10 produces post coda closure voice-bars for the word *cup* in two sessions. In C10’s session 3, he produced post coda closure voice-bars 6 times out of 13 utterances. In his session 5, C10 produced these voice-bars 5 times out of 13 utterances.

![Frequency of voice-bar over time: Bug](image.png)

Figure 14. Frequency of voice-bar over time: Bug.
Figure 15. Frequency of voice-bar over time: Tub.

Figure 16. Frequency of voice-bar over time: Duck
Experimental questions 3 asks: if there exist voice-bars after the closure of coda consonants in the speech of the caretakers, what sort of patterns (of the manifestation of voice-bars) exists in the speech of these adults; do they consistently produce voice-bars after the closure of voiced coda consonants and not after the closure of voiceless coda consonants. As evidenced in Tables X1 through X10, the caretakers do produce post coda closure voice-bars in their speech, and do so exclusively before voiced coda consonant. No exceptions have been found in the 4 words spoken by these 10 adults.

Table 5. Results from Caretaker 1.

<table>
<thead>
<tr>
<th>Word</th>
<th>Instances with Voice-bar</th>
<th>Instances without Voice-bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>bug</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>cup</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>duck</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>tub</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 6. Results from Caretaker 2.

<table>
<thead>
<tr>
<th>Word</th>
<th>Instances with Voice-bar</th>
<th>Instances without Voice-bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>bug</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>cup</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>duck</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>tub</td>
<td>8</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 7. Results from Caretaker 3.

<table>
<thead>
<tr>
<th>Word</th>
<th>Instances with Voice-bar</th>
<th>Instances without Voice-bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>bug</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>cup</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>duck</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>tub</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 8. Results from Caretaker 4.

<table>
<thead>
<tr>
<th>Word</th>
<th>Instances with Voice-bar</th>
<th>Instances without Voice-bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>bug</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>cup</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>duck</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>tub</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 9. Results from Caretaker 5.

<table>
<thead>
<tr>
<th>Word</th>
<th>Instances with Voice-bar</th>
<th>Instances without Voice-bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>bug</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>cup</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>duck</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>tub</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

44
Table 10. Results from Caretaker 6.

<table>
<thead>
<tr>
<th>Word</th>
<th>Instances with Voice-bar</th>
<th>Instances without Voice-bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>bug</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>cup</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>duck</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>tub</td>
<td>8</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 11. Results from Caretaker 7.

<table>
<thead>
<tr>
<th>Word</th>
<th>Instances with Voice-bar</th>
<th>Instances without Voice-bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>bug</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>cup</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>duck</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>tub</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 12. Results from Caretaker 9.

<table>
<thead>
<tr>
<th>Word</th>
<th>Instances with Voice-bar</th>
<th>Instances without Voice-bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>bug</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>cup</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>duck</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>tub</td>
<td>11</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 13. Results from Caretaker 10.

<table>
<thead>
<tr>
<th>Word</th>
<th>Instances with Voice-bar</th>
<th>Instances without Voice-bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>bug</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>cup</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>duck</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>tub</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 14. Results from Caretaker 11.

<table>
<thead>
<tr>
<th></th>
<th>Instances with Voice-bar</th>
<th>Instances without Voice-bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>bug</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>cup</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>duck</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>tub</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

7.3. Discussion

7.3.1. Predictions

The results of this study produced some interesting findings. **Prediction 1** correctly predicted that the appearance of voice-bars in the speech of young children will not be random; voice-bars will appear more often after the closure of voiced coda consonants and less often after the closure of voiceless coda consonants. While the children do not produce voice-bars exclusively after the closure of voiced coda consonants, they are more likely to produce them after the closure of voiced coda consonants than after the closure of their voiceless counterparts. There seem to be very few instances in which children produce voice-bars after voiceless coda consonants, as indicated by Figure 16 for *duck* and Figure 17 for *duck*. Perhaps this reflects Kager’s interpretation (Kager et al., 2005), i.e. it may be an indication that the acquisition of the feature [+voice] in coda consonants is more difficult than the acquisition of the feature [-voice].

**Prediction 2** predicted that there will be an observable trend (increasing production of voice-bars after voiced coda consonant closures; decreased production of voice-bars after voiceless coda consonant closures). Such trends are seen in the results of the voiced coda consonants, at least for some children, but not in the results of the voiceless coda consonants. Perhaps children have reached a plateau in terms of the acquisition of the feature [-voice] in coda consonants but are still actively acquiring the feature [+voice].

**Prediction 3** correctly predicted that the primary caretakers will produce more voice-bars after the closure of voiced coda consonants and fewer voice-bars after the closure of voiceless coda consonants. The results show that the adult caretakers of the children do not
produce any voice-bars after voiceless coda consonant closures. All of the caretakers produce voice-bars after the closure of voiced coda consonants, at least in this set of elicited utterances. In the four words that were examined, there were no exceptions. If the voice-bar is one of the most important cues to voicing, the pattern in which these adult caretakers produce voice-bars supports that hypothesis.

Assuming that motherese or child-directed speech is at least as rich in cues as laboratory induced speech, these caretakers would also only produce voice-bars after the closure of voiced coda consonants when speaking to their children. As a consequence, these children would emulate this voice-bar pattern.

Since the children seem to become better at producing voice-bars after the closure of voiced-coda consonants, they have probably recognized the voice-bar as an important cue to voicing. It is reasonable to suppose that feedback from the caretakers' motherese (which is likely similar to laboratory induced speech) greatly influences the speech of children during the acquisition period. Perhaps the children, through imitation of their caretakers, have discovered that the voice-bar distinguishes voiced stops from unvoiced stops, and that it is an important cue to voicing. With practice, these children will learn to perfect their voice-bar production, producing them exclusively after the closure of voiced coda consonants, just like their caretakers produce them during laboratory induced speech (and maybe child-directed speech).

### 7.3.1. Future Work and Concluding Remarks

This study can be expanded and many future studies can be done using the results of this study. While this study only focused on the speech of 5 children, if time permitted, the speech of more children can be examined to make the study more complete. Different words, not just the four chosen, can also be studied to see if similar observations can be made. The study of the duration of voice-bars will probably yield results supporting the theory that voice-bar duration is a primary cue to voicing (Nearney, 1997). The study of voice-bars in recordings of real child-directed speech may also help shed some light on the acquisition of the [+voice/-voice] feature. In addition, an extended study of language-learning children beyond the age of the children in this study (i.e. both earlier and later) in conjunction with the speech of their caretakers would be interesting. Perhaps as the children become older, their caretakers eventually stop speaking to
them in motherese and shift to adult-directed speech. The appearance of voice-bars in the speech of adult-directed speech may affect the appearance of voice-bars in the speech of children who are in later stages of language acquisition.

The acquisition of speech is also one of the most important aspects of a child's development. Although the acquisition of language may not seem complex or effortful to adult observers, the process is actually far from simple or straightforward, considering how much difficulty adults encounter while trying to learn a second language. Children have to isolate speech sounds from their noisy surroundings, mimic them, and eventually be able to use their newly learned tool to communicate with others. The human capacity for learning to perceive and reproduce such a large repertoire of sounds in such a short amount of time is remarkable. Plenty of literature detailing the language acquisition process exists but nevertheless, the acquisition process is far from being fully understood. The work performed for the purpose of this study is just the tip of the iceberg.
This page is intentionally left blank.
8. Bibliography


