ABSTRACT

Title of Document:	MATERNAL VOICE ONSET TIME IN INFANT- AND ADULT-DIRECTED SPEECH: CHARACTERISTICS AND POSSIBLE IMPACTS ON LANGUAGE DEVELOPMENT.
	Julia L. Sampson, Master of Arts, 2013
Directed By:	Dr. Nan Bernstein Ratner, Department of Hearing and Speech Sciences

Infant-directed speech (IDS) contains many unique characteristics that may facilitate language development. One acoustic cue that may differ in IDS compared to adult-directed speech (ADS) is voice onset time (VOT). The present study examines the VOT of open- and closed-class words in speech to infants at 10/11, 18, and 24 months of age, as well as in speech to adults. This study also looks at correlations between clarification of VOT in speech to infants, and language outcomes at 2 years. Results show that VOT clarification in IDS did not differ significantly at any of the ages. Overlap between voicing categories for open class words was significantly less in ADS than IDS. The overlap for closed class words at 18 months was significantly related to language outcomes, with lower overlap relating to higher outcome scores. Possible explanations are discussed.

MATERNAL VOICE ONSET TIME IN INFANT- AND ADULT-DIRECTED SPEECH: CHARACTERISTICS AND POSSIBLE IMPACTS ON LANGUAGE DEVELOPMENT.

By

Julia Lauren Sampson

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Arts
2013

Advisory Committee:
Dr. Nan Bernstein Ratner, Chair
Dr. Rochelle Newman
Dr. Yi Ting Huang

Table of Contents

List of Tables

List of Figures

Introduction

The way we speak is influenced by our audience. Different listeners elicit different changes in vocabulary choice, rate of speech, and pitch, among many other things. For example, we frequently adopt a more formal register when speaking with authority figures compared to the casual register used when speaking with peers. Another well-studied example of this phenomenon is the characteristic speaking style used when speaking to infants, sometimes referred to as “motherese”. However, it is not just mothers who alter their speech to infants. This behavior is also exhibited by fathers (Fernald et al., 1989; McRoberts & Best, 1997), children (Weppelman, Bostow, Schiffer, Elbert-Perez, & Newman, 2003), and other adults (Soderstrom, 2007). For this reason, it is more accurately referred to as “infant-directed speech” (IDS).

IDS is characterized by slower rate, exaggerated intonation and higher pitch (Fernald & Mazzie, 1991; Fernald & Simon, 1984; Inoue, Nakagawa, Kondouu, Koga, & Shinohara, 2011), longer vowel duration (Bernstein Ratner, 1984b; Bernstein Ratner & Luberoff, 1984), shorter utterances, longer pauses (Fernald et al., 1989), and increased vowel space (Bernstein Ratner, 1986). IDS is not static; mothers are sensitive to feedback produced by infants and they increase their fundamental frequency when it elicits positive feedback but not when infants are unengaged (Smith & Trainor, 2008). IDS is also different than other registers such as pet-directed speech (PDS) and foreigner-directed speech (FDS) in significant ways. While IDS and PDS are both characterized by high pitch, vowel space to infants is significantly larger than to pets (Burnham, Kitamura & Vollmer-Conna, 2002). In foreigner-directed speech (FDS) vowel space is similar to IDS, but pitch is more similar to ADS (Uther, Knoll & Burnham, 2007). This suggests that not only do adults modify their speech in ways specific to their listener, but that some of these modifications are for emotional or affective purposes (i.e., heightened pitch) whereas others may serve a linguistic purpose (i.e., vowel hyperarticulation) (Song, Demuth & Morgan, 2010).

Although IDS serves many purposes, one advantage it offers is that infants as young as 1-month old have shown a preference for listening to it over adult-directed speech (ADS) (Cooper & Aslin, 1990; Fernald, 1985). This preference is greater for younger infants (4-5.5 months) than older infants (Werker & McLeod, 1989), with some studies finding that the preference no longer exists at 9 months (Hayash, Tamekawa, & Kiritani, 2001; Newman & Hussain, 2006). The preference for IDS over ADS persists whether the IDS is spoken by a male or a female, and even when the speaker's face is kept constant (Werker & McLeod, 1989).

Aside from infants' apparent preference for IDS, it has also been postulated that the unique features of IDS may facilitate language learning in infants (Singh, Nestor, Parikh & Yull, 2009). Infants tend to learn associations between novel words and novel objects when they are presented in IDS, but not in ADS (Ma, Golinkoff, Houston, & Hirsh-Pasek, 2011). They also listen significantly longer to passages containing words that are familiarized in IDS compared to passages containing words that are familiarized in ADS (Singh, Neston, Parikh, & Yull, 2009). This may be because IDS contains distinctive prosodic patterns to highlight focused words, such as exaggerated pitch peaks (Fernald & Mazzie, 1991), increased amplitude (Messer, 1981), slower speaking rate, and hyperarticulated vowels (Song, Demuth & Morgan, 2010).

One way in which mothers modify their speech to infants is through vowel modification. Vowels are longer in speech to preverbal infants than children and adults (Liu, Tsao & Kuhl, 2009). When comparing IDS to ADS, the degree of pre-boundary vowel lengthening (the expected lengthening of vowels in the clause-final position compared to those in the clause-medial position) is twice as great in speech to preverbal children, but less pronounced to children at the one word level (Bernstein Ratner, 1986). In addition to longer vowels, vowel space is also greater in IDS than ADS (Bernstein Ratner, 1984a; Liu, Kuhl & Tsao, 2003; Liu et al., 2009; Song et al., 2010; Werker et al., 2007). This increased vowel space has been correlated with infant speech perception performance (Liu et al., 2003). Mothers use greater vowel space in content words (e.g. nouns like “ball” or “bottle”) than for function words (e.g. “the” and “and”) in their speech to preverbal listeners, but begin to increase their vowel space for function words as the children become more linguistically sophisticated (Bernstein Ratner, 1984a). Given that content words develop earlier than function words in children’s speech, it may be that mothers are adjusting their input based on the infants’ stage of language development, or it could be that function words do not appear until later in the child’s language development because mothers are not clarifying these words until later.

IDS also helps infants to segment speech, which may contribute to learning new words and syntax. Children who performed well on speech segmentation tasks from 7.5-12 months of age were found to have higher vocabularies at age 2, and higher language profiles at 4-6 years old (Newman, Bernstein Ratner, Jusczyk, Jusczyk, & Dow, 2006). In a study using artificial languages made up of nonsense words in which the only cue to word boundaries was the statistical structure (which was the same in both the IDS and

ADS conditions) infants showed a preference for words over part words only in the IDS condition, suggesting that the acoustic properties of IDS may facilitate language acquisition (Thiessen, Hill & Saffran, 2005). The prosodic exaggeration seen in IDS may help to set off major linguistic units: infants prefer to hear IDS interrupted at major clause boundaries than IDS interrupted mid-clause (Kemler Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989). Furthermore, infants as young as 2 months old have been able to remember the order of words in a sentence, but only when they heard it in a prosodic unit rather than in a fragmented condition (Mandel, Kemler Nelson & Jusczyk, 1996).

Although in most ways IDS seems to be altered in order to simplify the input or exaggerate key features, there is also evidence that some adult characteristics of speech are maintained. For example, in Japanese there is a process of vowel devoicing which occurs on high vowels that occur between voiceless consonants, or before a pause. If mothers were trying to make the vowels more salient, it might be expected that they would decrease this process in IDS. However, mothers actually maintained this pattern of vowel devoicing in IDS at the same level as found in ADS (Fais, Kajikawa, Amano, & Werker, 2010). In other words, they did not sacrifice the adult form in order to accommodate to their infants. Another example of a maintained adult pattern in IDS can be found in a study of phonological rule usage in IDS involving American mothers, in which the phonological rule of palatalization (i.e. /dId ju/ → /dIdʒu/) was found more in IDS than ADS, but other rules such as dental deletion, /ð/ deletion, and ts/s conversion were less likely to appear in IDS than ADS (Bernstein Ratner, 1984b). However, in this study mothers appeared to be using a pattern of alternation of clarified and less clarified forms in their speech. In doing so, they may have been teaching their children about these optional phonological rules (Bernstein Ratner, 1984b). It appears that the nature of the input that infants receive may be dependent on the goal the parent hopes to achieve, whether this goal be to teach the infant a new word, to have the infant imitate a production, or to repair a non-adultlike production (Bernstein Ratner, 1996, p.141).

VOT

One acoustic cue that may be clarified in IDS is voice onset time (VOT). VOT is defined as the interval between the release of stop occlusion and the onset of vibration of the vocal folds. This acoustic cue is used in many languages to differentiate voicing of stop consonants (Lisker & Abramson, 1968). Although there are other indicators of voicing, such as formant transitions (Stevens & Klatt, 1974), it appears that VOT is a sufficient cue to differentiate voiced and voiceless stops on its own (Lisker, 1975). VOT exists along a continuum that can range from before the release of the stop (negative VOT values associated with prevoicing) all the way to 100 or more milliseconds lag. There is categorical perception of VOT. This means that listeners can only discriminate changes in VOT that occur across category boundaries, but not within them. The categorical boundaries differ depending on language.

In English, there is a bimodal distribution of VOT; stop consonants with a short lag VOT are defined as voiced ([b, d, g]), and those with a long lag are defined as voiceless ([p, t, k]) (Lisker & Abramson, 1965). The perceptual boundary is defined as the point along the continuum at which 50% of responses are associated with either the voiced or the voiceless category. In English, this boundary exists around +25 ms for labial stops, +35 ms for alveolar stops, and +45 ms for velar stops (Lisker & Abramson, 1965; Zlatin, 1974).

Infants as young as 1 month are able to discriminate between the voiced/voiceless categories /b/ and /p/ (Eimas, Siqueland, Jusczyk, & Vigorito, 1971). They are also initially able to discriminate contrasts that do not exist in their native language, but by the end of the first year, their discrimination is limited to contrasts found in their native sound system and many foreign language contrasts are no longer discriminated (Werker & Tees, 1984). This typical phenomenon can be altered by training or exposure to carefully designed input: exposure to a bimodal distribution of sounds can counteract this effect by enhancing discrimination of an originally difficult contrast, and can even be projected onto other untrained contrasts which occur in a different place of articulation (Maye, Weiss & Aslin, 2008). On the other hand, exposure to a unimodal distribution has the effect of reducing discrimination (Maye, Werker & Gerken, 2002).

When interpreting results of studies demonstrating early infant speech discrimination abilities, it is important to consider the speech stimuli used and the conditions in which they are presented. For example, in the classic Eimas et al. (1971) study in which infants as young as 1-month old were able to discriminate between /b/ and /p/, the stimuli were synthetic speech sounds /ba/ and /pa/ which were presented at 75 dB. While this experiment tells us about an infant's ability to discriminate between voiced and voiceless sounds when they are presented in single syllables under ideal listening conditions, it does not tell us much about their ability to make this discrimination when the sounds are present in their mothers' speech in a naturalistic setting. In fact, there are many factors that can affect VOT production in speech which need to be considered.

Factors that Affect VOT

Many factors have been studied which may have an effect on VOT. One of the most widely accepted factors that affects VOT is place of articulation. VOT for voiceless stops increases as place of articulation movies from labial, to alveolar, to velar (Lisker & Abramson, 1964; Zlatin, 1974; Krause & Braida, 2004). As previously mentioned, the perceptual boundary in English for changes from around +25 ms for labial stops, +35 ms for alveolar stops, and +45 ms for velar stops (Lisker & Abramson, 1965; Zlatin, 1974).

Another widely accepted factor that influences VOT is stress. Stressed voiceless stops (e.g., [p], [t], [k]) tend to have a longer VOT than unstressed voiceless stops (Lisker & Abramson, 1967). This difference between VOT for stressed and unstressed stop consonants is not as great in voiced stops, which leads to greater variation between the voiced and voiceless sounds in the stressed position than in the unstressed position (Lisker & Abramson, 1967). The former often correlates with open class words (e.g., "It's a ***PI****zza."*) whereas the latter correlates with closed class words (e.g., "Let's *put* it away**.**"). Furthermore, the voicing lag continues to increase when the voiceless stop occurs in syllables bearing sentence-final stress (Lisker & Abramson, 1967). For example, the VOT for the /k/ in the initial stressed position of the word "cookie" would be expected to be greater when it occurs in sentence-final position (e.g., "Eat the *cookie*.") compared to other positions in the sentence (e.g., "The cookie tastes good.").

It is also important to consider context when looking at VOT. The difference between VOT in isolated words compared to words in sentences is about 25 milliseconds (Lisker & Abramson, 1967). Whereas VOT can clearly identify voiced vs. voiceless words in isolation, this becomes more difficult in sentences because of the increased overlap between voicing categories (Lisker & Abramson, 1967). Speaking rate can also affect VOT, with slower articulation rates associated with increased VOT (Theodore, Miller & DeSteno, 2009). This effect may only be present for voiceless stop consonants and not voiced stop consonants (Summerfield, 1981). In a study of individual talker differences in VOT, speaking rate was the strongest predictor of VOT (Allen, Miller & DeSteno, 2003). However, the effect of rate on VOT is not systematic across different talkers (Theodore, Miller & DeSteno, 2009). It is thus important to consider context and speaking rate when comparing VOT across conditions or groups.

One context whose effects on VOT are somewhat contested is vowel context. Lisker and Abramson (1967) reported that they found no correlation between vocalic environment and VOT. However, more recent studies have produced some discrepant results. In a study of the effects of place of articulation and vowel quality on VOT in English stops, Nearey and Rochet (1994) found that the vowels associated with the longest VOTs for the preceding voiceless stop consonants were the high, long vowels, /i/ and /u/. The vowels associated with the shortest VOTs for the preceding voiceless stops were the lax vowels, /I/, /ʊ/, and /ʌ /. For /p/, the longest VOT (73.2 ms) occurred before /u/ and the shortest VOT (59.8 ms) occurred before /o/. For /t/, the longest VOT (85.1 ms) occurred before /i/, and the shortest VOT (65.6 ms) occurred before /ʊ/. Finally, for /k/, the longest VOT (86.8 ms) before /i/ and the shortest VOT (71.8 ms) occurred before /ɑ/. However, these differences were not reliable across subjects and should be taken with caution due to the high subject variability. For voiced English stops, Nearey and Rochet found that VOTs in the context of /u/ were significantly longer than /ʌ, ɛ, o, æ/. Berry and Moyle (2011) found that high vowel contexts resulted in significantly longer VOT durations, but only for the stop consonant /k/, and not /t/ or /p/. This study did not look at voiced stops. More work needs to be done before the effect of vowel context on VOT can be fully understood.

The reason that most of these effects increase voiceless stops and not voiced stops is likely because of the perceptual boundary. For example, if the perceptual boundary for the bilabial place of articulation is +25 ms, then a voiced sound can only increase 25 ms before it is perceived as voiceless, whereas a voiceless sound can theoretically increase as long as the speaker is able to aspirate. Some voiceless sounds can have a VOT over 100 ms, which means voiceless sounds have the potential to increase over 75 ms from the perceptual boundary at which they are first perceived as voiceless, 50 ms more than the VOTs for voiced sounds are able to increase.

Another factor that affects not only VOT but also clarification of words in general is whether a word is considered “given” or “new” in conversation. Words produced for the first time in a conversation (“new” words) are more intelligible than these same words produced for the second time (“given” or “old” words) (Fowler & Housum, 1971). Furthermore, words presented in noise that are produced in an uninformative context are more intelligible than the same words produced in a more informative or predictable context (Lieberman, 1963). Speakers appear to use greater precision for words that are not probable or redundant so that the acoustic information alone is sufficient for the listener to identify the “new” word (Fowler & Housum, 1971).

Clearly there are many factors at play that can affect VOT, especially in connected speech. Although infants are able to discriminate synthetic speech sounds presented in single syllables under ideal experimental conditions, this advantage may not help them discriminate voicing in the speech they are actually exposed to. For this reason, an interesting question to ask is whether or not mothers attempt to exaggerate the voicing distinction in their speech to infants compared to other adults.

Previous VOT Studies

To date, there has not been a great deal of research regarding differences in VOT in IDS compared to ADS. The research that has been done has revealed some contradictory results. Baran, Laufer and Daniloff (1977) conducted the first study comparing VOT in infant-directed vs. adult-directed speech. This study examined the speech of 3 English-speaking mothers with firstborn infants who attended the infant laboratory every 2 weeks from birth to 2 years to participate in naturalistic play interaction. The session chosen for VOT analysis was that which preceded the emergence of the infant's first word and occurred within 20 days of the first birthday for all infants. Samples were also taken of ADS and oral reading. The authors found no difference in VOT for voiceless stops across the conditions, but when looking at a subsample consisting of 20 exemplars for each stop, they found that voiceless stops had significantly shorter VOTs in the IDS condition than in the ADS condition. This would be expected to lead to more overlap in the IDS condition between the two voicing categories and thus less clarity; however, the authors did not report on this factor.

In another study, Moslin (1979) (later writing as Malsheen (1980)) collected speech samples from six mother-child dyads over a period of 6 months. All of the mothers were middle-class, native speakers of English. The children were classified according to their stage of language development, with the youngest children (aged 6 and 8 months) babbling, the middle children (aged 15 and 16 months) speaking primarily one-word utterances, and the oldest children (aged 2;5 and 5;2) speaking relatively syntactically complex sentences with the highest MLUs of the group. Speech samples were collected in the homes of the participants, and mothers were given no instructions other than to play with their child. They were also under the impression that the study was focusing on the child's speech, and not their own. All words that contained initial, singleton stop consonants were selected for VOT analysis. This sample included both open- and closed-class items with initial stops occurring in all vowel contexts and in both stressed and unstressed syllables. Malsheen found that the only significant difference between IDS and ADS occurred in the mothers' speech to the middle children who were producing single words, in which there was a significant reduction in overlap of VOT values between voiced and voiceless categories, with significantly higher mean VOT values for initial voiceless stops. This phonetic alteration was not present in the mothers' speech to the youngest children who had not yet begun producing their first words, or to the oldest children who had been producing relatively complex syntactic sentences. These results differ from those found by Baran, in which voiceless stops were shorter in IDS than in ADS. However, in Baran's study the IDS tokens were addressed to children who had not yet produced their first word, whereas the middle children in Malsheen's study were already at the one-word stage. Malsheen suggested that these findings might be explained by the Independent Phonetic Clarification (IPC) hypothesis.

The IPC hypothesis states that mothers may modify the phonetic component of their speech, in ways such as a VOT increase, independently of other modifications on different linguistic levels. For example, although prosodic changes in IDS are present long before infants begin to say their first words, phonetic changes may not appear until later on. This hypothesis suggests that mothers may be sensitive to their children's attempts to produce phonetic distinctions in voicing, and they attempt to assist their children by providing maximally distinct phonemic input. This would explain why the least overlap in VOT was present in IDS to infants who were at the one-word stage of development, and not to preverbal or more linguistically advanced children.

Another possible explanation for the difference in VOT in IDS and ADS might be that there are more function words in ADS than in IDS, and function words tend to be unstressed. In order to account for this possible effect, Malsheen compared the mean VOT values of one mother's production of word-initial /t/ in the IDS and ADS conditions for closed class items only. She found that even for unstressed closed class words, mean VOT duration was greater in IDS. These results help to reject the claim that VOT clarification in IDS is a by-product of syntactic differences such as fewer function words.

Kubaska (1982) investigated the IPC hypothesis by following two mother-infant dyads longitudinally. The mothers' spontaneous speech to their children was taped biweekly and covered the children's babbling, early word, and later acquisition stages. Contrary to the IPC hypothesis, no significant differences in VOT between the pre-word and post-word stages were found for the first mother. For the second mother, Kubaska found that VOT for /t/ and /k/ was significantly longer in the pre-word stage rather than the post-word stage. For the first mother there were no significant differences between the ADS and IDS samples. There was one significant difference between ADS and IDS for the second mom. VOT for /d/ in ADS was significantly shorter than pre-word IDS. Kubaska suggested that there is likely individual variation in VOT adjustments in IDS among mothers.

Sundberg and Lacerda (1999) investigated VOT characteristics in the IDS and ADS of six Swedish mothers and their 3-month-old infants (3 girls and 3 boys). The recording sessions took place in a sound-treated studio and mothers were instructed to play as they normally would at home. After each play session, the investigator talked informally with the mother in order to obtain an adult-directed speech sample. All prevocalic initial and medial stop consonants were included in this analysis. In this analysis, VOT was significantly shorter in both voiced and voiceless stops to the infant than to the adult listener. The authors suggested, similarly to Malsheen, that the mothers could be placing different weights on different properties of their speech depending on their children's linguistic needs. In order to test this hypothesis, Sundberg (2001) looked at IDS directed to 11-14 month old infants compared to ADS. Preliminary results were in agreement with Malsheen's findings, and included significantly longer VOT values in IDS relative to those in ADS.

In an additional VOT study by Englund (2005), six Norwegian-speaking mothers were recorded 10 times over 6 months while they were changing their infants' diapers. This situation was chosen for the recording sessions in order to encourage a natural interaction. After the mother-infant interaction, the mother was recorded speaking to an adult who asked about the general health of mother and child, and whether she remembered any words spoken to the infant during the diaper change. All word initial voiced and voiceless stop consonants that occurred in primary and secondary stressed syllables of both content and function words were included in the analysis. The results of this analysis indicated that VOTs were generally longer in IDS than ADS over the first 6 months of life, but there were no clear difference in voicing contrast between these two conditions. This finding contradicted the results found by Sundberg and Lacerda, in which VOTs were shorter in IDS than ADS; however, it remains in general agreement with Malsheen as it shows there is no significant difference in voicing contrast in speech to infants before their first word.

In a more recent study VOT study, Synnestvedt (2009) compared VOT in speech addressed to 7.5-month-old infants and 11-month-old infants to VOT in the ADS of the same speakers. The study involved 15 mother-infant dyads. The mothers were all native speakers of English, and the infants had no previous diagnosis of developmental problems. Unlike the previous studies, this study included only identical words that could be found in both the IDS and ADS conditions. Speech samples were collected in a sound-treated room where mothers were instructed to play with their infants as they would at home. After each play session, an experimenter spoke with the mother to obtain an ADS sample. Receptive language scores were also collected before each session by asking the mothers to fill out the MCDI form. Results indicated no difference in VOT duration between IDS and ADS voiceless stops at both 7.5 and 11 months. However, mean VOT duration for voiced stops was significantly longer in IDS at 7.5 and 11 months than in ADS. Mothers’ VOTs were more distinguishable in ADS than IDS at both ages. Furthermore, there was no correlation between degree of VOT clarification and MCDI scores. This is the only one of the studies discussed that only found a difference in VOT duration for voiced stops. It also contradicts the results found on VOT clarification by Malsheen (1980), in which VOT clarification was only greater to children who were producing their first words, and in no cases greater to the adult. However, these two studies used different measures of clarification, and the infants were at different ages.

Many of the infants in the study had not yet begun to produce their first words. Thus, if they were followed at later ages, the mothers' patterns of VOT clarification may change.

Present Study

Previous studies of VOT in maternal speech to infants and adults have produced conflicting results. Malsheen (1980) found that there was significant reduction in VOT overlap in the mothers' speech to infants who were speaking single words, but not to preverbal infants or children with greater MLUs. Similarly, when looking at 11-14 month infants who were presumably around the one-word stage, Sundberg found VOT to be significantly longer in IDS than in ADS, confirming that mothers may make the most phonetic adjustments when infants are learning their first words. However, when Baran et al. looked at VOT of voiceless stops to children just preceding their first words, VOT of voiceless stops was significantly shorter, presumably leading to more overlap. When looking at VOT in speech to preverbal infants, Sundberg and Lacerda found that VOT for both voiced and voiceless stops were significantly shorter in IDS in speech to preverbal 3-month old infants than in ADS. Contrary to these findings, Englund found that there was no clear difference in voicing contrast between IDS and ADS over the first 6 months of life, but VOTs were generally longer in IDS than ADS. The results from Synnestvedt (2009) bring even more contradictory results, as it was the only study to find that VOT of voiced stops was actually longer in speech to both 7.5 month and 11 month infants compared to adults.

In addition to the discrepant findings, none of these studies have investigated the impacts of VOT characteristics on infants' language development. Whether or not mothers do clarify their VOT in IDS compared to ADS is not of much importance if it does not help the infant in some way. Although the IPC hypothesis claims that mothers are making phonetic clarifications in order to provide an advantage to their infants, what this advantage may be is unclear without looking at infant language outcomes. The present study addresses this question by looking for correlations between VOT clarification and infant language outcomes at 2 years old. These language outcomes include scores on the *Peabody Picture Vocabulary Test* (PPVT), *Expressive One Word Vocabulary Test* (EOWVT) and the *MacArthur Communicative Development Inventory* (MCDI).

The present study also seeks to gain a better understanding of VOT clarification in IDS compared to ADS. This study includes 17 mother-infant dyads followed longitudinally. Recordings for analysis of VOT in IDS were taken from mother-infant play sessions held when the infants were 10-11 months of age, 18 months of age, and 24 months of age. These recordings were compared to the mothers' speech to an unfamiliar adult. These ages were chosen for analysis because they should correspond to different stages of language development. Infants at 10-11 months may be producing jargon or beginning to enter the one-word stage. From 18-24 months, infants should be using single words and have an MLU of 1.0-2.0. By looking at both ages, we may be able to capture differences that occur as the child's MLU increases.

This study includes separate analyses of open- and closed-class words. Other studies have chosen to only look at open class words or have combined the two. It is important to look at both, as they both constitute the input to infants at the ages to be investigated. However, as these words are often characterized by different stress patterns, which is a well known factor affecting VOT, combining them may affect the results.

Rather than only analyzing identical words that occur in both IDS and ADS, this study matched tokens based on contexts that have been shown to affect VOT. This includes stress, place of articulation, and position in a sentence.

Hypotheses

The first hypothesis of this study is that VOT clarification in IDS will be positively correlated with the infants' language outcomes at 2-years old. VOT clarification will be measured in a variety of ways, including percent overlap, $d_{(a)}$, and Discreteness of Voicing Category (DOVC), in order to be able to make direct comparisons to past studies. These measures are re-defined below. This phonetic clarification may give infants an advantage, as it helps to emphasize the voicing contrast between words in connected speech. Normally VOT in connected speech is characterized by overlapping values, and is a less reliable indicator of voicing categories (Lisker & Abramson, 1967). While more linguistically sophisticated listeners have other cues available to them to help detect voicing differences, such as semantic knowledge and context from the sentence, these cues are not readily available to infants. So while infants are able to distinguish between voiced and voiceless synthetic speech sounds occurring in single syllable words, it may be more difficult for them to detect this contrast in the natural input they receive everyday. In order to help infants detect this voicing contrast in running speech, mothers may adjust their VOT in a way that makes voiced and voiceless sounds more distinguishable. This may help infants to detect new words, or influence their own attempts of production, which would be expected to have a positive effect on language outcome scores. Furthermore, VOT clarification may be considered a proxy for a larger clarification phenomenon. If mothers are adjusting the acoustic quality of their phonetic input in IDS, it is likely that they are making additional modifications as well. While this study does not investigate other types of phonetic clarifications, if this hypothesis is confirmed, mothers could be instructed to use clear speech to help foster language development.

A second hypothesis of this study is that VOT clarification of open class words will be greater in speech to infants who are 18 month-olds compared to infants at 10/11 months, 24 months, and adults. This hypothesis is in accordance with the IPC hypothesis suggested by Moslin (later writing as Malsheen). Whereas infants at 10/11 months may not have entered the one-word stage, it is expected that the infants at 18 months should be using single words and beginning to expand their vocabularies. It is expected that VOT clarification will be greater in speech to infants at 18 months than to infants at 24 months, because by this age the infants may be beginning to combine words and increase their MLU, and the phonetic clarification may not be weighted as heavily in mothers' adjustment in IDS. This is what Malsheen found to be the case with her oldest age group in the 1980 study.

The final hypothesis is that VOT clarification of closed class words may be the greatest in speech to the 24 month-old infants compared to in speech to the 10/11 and 18 month-old infants. Given the findings that vowel space for function words begins to increase as children become more linguistically sophisticated (Bernstein Ratner, 1984a), it may be reasoned that similar clarification will occur with VOT. Thus, as infants begin to use more closed class words in their own speech, mothers may be sensitive to this development and begin to modify their input to phonetically highlight these words. This again follows the IPC hypothesis, and indicates that the adjustments in IDS are not random, but are directed towards the child's stage of development. It is expected that the measures of VOT clarification will indicate the greatest degree of overlap in function words to the youngest age group, and the greatest degree of separation to the oldest age group, who are beginning to use some closed class words in their own speech.

Methods

Participants

Participants were mother-infant dyads who were part of a larger longitudinal study at the University of Maryland (NSF grant BCS 074512, Rochelle Newman, PI). All of the mothers were native English-speakers and the infants were learning English as their native language. The infants were all born within 2 weeks of their due dates, and were normally developing with no previously diagnosed developmental conditions or hearing loss. As part of the longitudinal study, each dyad came in for visits at 7, 10, 11, 18, and 24 months, with 2 weeks variation allowed around each age visit. At the end of each visit, the infants were rewarded with a small prize for their participation. The data for this study was collected from the 10/11, 18, and 24-month visits. Initially, 29 mother-infant dyads were chosen for this study because they had completed all of their visits, and at least 2 out of 3 of the visits had been transcribed. From this initial pool of 29, 10 dyads were excluded after the data selection procedure because they did not meet the requirements for the minimum number of matched tokens. After the acoustic analysis procedure, an additional 2 dyads were excluded due to excessive noise. After the exclusions, there were 17 mother-infant dyads included in this study, including 11 male infants and 6 female infants.

IDS and ADS Speech Samples

Upon arrival at the University, participants were escorted to a sound-treated therapy room where the play session and interview would take place. The mothers were given an Audio-Technica ATR-35S lavaliere microphone to wear to record their speech. The speech samples were recorded as uncompressed WAV files using a Marantz PMD660 Professional Portable Digital Recorder at a sampling rate of 44.1 kHz. The sessions were also video-recorded using a flip camera. The experimenter brought in a cart of toys containing play food, baby dolls, books and stuffed animals for the mother and infant to choose from. Once the room was set up, the mother was instructed to play with her infant as she normally would at home. The participants in the study were not told that their speech was being examined in order to avoid a possible bias. Instead they were told that this was a study examining the infant's play behavior. For the adult-directed speech condition, an undergraduate student played with the child while the experimenter interviewed the mother about how her child usually plays at home. They play sessions and interviews lasted approximately 15 minutes. At the conclusion of the final play session or interview, participants were informed of the true aim of the study and given a consent form to fill out, or the opportunity to drop out if they did not want their data to be included.

Outcome Measures

At the 24 month session, children were administered a variety of standardized language assessments in order to collect language outcome data. The *Peabody Picture Vocabulary Test 4* (PPVT-4) is a receptive vocabulary test that measures understanding of single words. The norms for this test begin at 2 years, 6 months old, so raw scores were used for this measure. In order to assess use of single words, the *Expressive One Word Vocabulary Test* (EOWVT) was administered. Standardized scores and percentile ranks were calculated for this measure. Mothers were also given the *MacArthur Communicative Development Inventories* (MCDI) form to fill out at this visit. The MCDI asks mothers to check off words that they believe their infants can understand. Raw scores were calculated for this measure by adding the number of checked items.

Data Selection Procedure

The WAV files of the mother-child play session and interview were uploaded as one raw file to a PC. These files were then split into two files using Audacity, a free sound recording and editing program (http://audacity.sourceforge.net/). One file contained only the mother-child play session and the other contained only the interview. Play sessions and interviews were transcribed using the Computerized Language ANalysis (CLAN) program (MacWhinney, 2009). This is a program which is available for free download online as part of the Child Language Data Exchange System (CHILDES). This program allows the audio file to be split into smaller segments that can then be replayed while typing for a more accurate transcription.

Once files were transcribed, the experimenter used the CLAN utility FREQ with the +u option to obtain a frequency count of all words spoken by the mothers to either the child or adult (freq +t*MOT +u). The experimenter identified all tokens with an initial singleton stop consonant (/t, d, p, b, k, g/) and transferred them to one of two .cut files; one file contained open class words and the other contained closed class words files (openclasstargets.cut and closedclasstargets.cut). The experimenter then ran another FREQ command on the files using the search option (+s). This option looks only for words that are listed in the subsequently named .cut file and provides frequency counts for each one. The +d0 option was used, which includes the line numbers in the output as well as the text of each line containing a target word. This command (freq +t*MOT +s@openclasstargets.cut +d0; freq +t*MOT +s@closedclasstargets.cut) was run on all IDS and ADS conditions for each mother.

The outputs were then compared in order to find “matched” tokens across conditions. Tokens were considered matched if they had the same initial phoneme, same word class (open or closed), and occurred in the same position of the sentence. This method was chosen rather than only matching identical words in order to obtain a larger sample to analyze. In a naturalistic setting in which the words the participants say are not controlled, identical words cannot always be obtained. Preliminary analyses revealed that even when attempting to match word families (i.e. cook, cooking, cookies would all be considered a family), there were no participants that met the criteria of at least 6 matched tokens for each stop consonant. Furthermore, while the effects of vowel context on VOT are inconclusive, place of articulation, stress, and context in sentence (whether it occurs in isolation, in the sentence-final position, or other positions of the sentence) are all factors that are widely accepted to affect VOT. Thus, it seems more appropriate to match tokens based on these variables than to limit the analyses to identical words. Although this may increase the variability of the sample, it also allows for a more accurate sample of the words heard in IDS compared to in ADS.

The frequency of occurrence of phonemes in English is not evenly distributed. One study examining 26 interviews of English speakers found that nearly two-thirds of the consonants were voiced, and the majority of consonants were articulated at the front area of the mouth (Mines, Hanson, & Shoup, 1978). It is well accepted that VOT varies with place of articulation, so in order to select which places of articulation should be included in the analysis, the frequency of each stop consonant was calculated for the current data set. In the current study, out of all possible open class tokens beginning with a singleton voiced or voiceless stop consonant, 49.2% were voiced and 50.8% were voiceless. The distribution of phonemes in the closed class targets was more similar to the findings of Mines et al. (1978) with 59.8% voiced and 40.2% voiceless. For place of articulation, the most frequently occurring phonemes for the open class group were the bilabial stop consonants /b/ and /p/ (51.8%), and the most frequently occurring phonemes for the closed class group were the alveolar stop consonants /t/ and /d/ (45.5%), followed by the velar stop consonants /k/ and /g/ (33.4%). Preliminary investigation of the closed class /t/ and /d/ tokens revealed that these tokens were often palatalized or so reduced that measurement of VOT was not feasible. Given the distribution of the current data and the preliminary investigation, the experimenter chose to analyze the bilabial voiced-voiceless pair ([b, p]) for the open class group and the velar voiced-voiceless pair ([k, g]) for the closed class group. Participants were only included if after the matching process they had at least six tokens for each of the target stop consonants occurring in the word-initial position. This resulted in a total of 18 mother-infant dyads for the open class group, and 10 mother-infant dyads for the closed class group. Of the participants in the closed class group, nine of them were also included in the open class group, and one additional dyad met criteria only for the closed class group.

Matched tokens were recorded in a separate Excel workbook for each mother. Each workbook included either 4 or 8 spreadsheets, depending on whether the mother met the criteria for both word classes (open and closed) or only one. There was one spreadsheet with either open or closed class tokens for each age (10/11, 18, and 24 months), and one spreadsheet with tokens addressed to an adult that matched the tokens addressed to infants at all ages. For each word, the experimenter listed the file from which it was taken, the line on which it appears in the file, the target phoneme for analysis, whether the phoneme is voiced or voiceless, and where it occurs in a sentence (initial position, medial position, final position, or single word). During the acoustic analysis, VOT was recorded in milliseconds for each token. Additionally, in order to make sure that each of the tokens were matched for stress, an auditory review was conducted by the experimenter and when it was clear that a target stop occurred in an unstressed syllable, the token was removed from the spreadsheet. In order to avoid any selection bias, tokens were analyzed in the order they appeared in the transcripts.

Acoustic Analysis

Each target was acoustically analyzed using PRAAT (Boersma & Weenink, 2009). CLAN allows media to be sent directly to PRAAT, where it can be viewed in a spectrogram and waveform. Using the spectrogram and acoustic signal, the target word was identified and VOT measurements were taken from the release of the stop consonant to the onset of low frequency periodic signal corresponding to the voicing of the following vowel. VOTs were recorded in the spreadsheets to the nearest millisecond. Figure 1 shows how VOT was measured. Previous studies have excluded tokens which were pre-voiced or for which there was no observable release. In the current study, a floor of 0 milliseconds VOT was set. There were a number of tokens in which the stop burst was masked by the vocalic environment. However, because these tokens made up a large proportion of the input to the child, excluding them would not provide an accurate depiction of VOT characteristics in IDS. By setting the floor at 0, the experimenter was able to make reliable judgments across files.

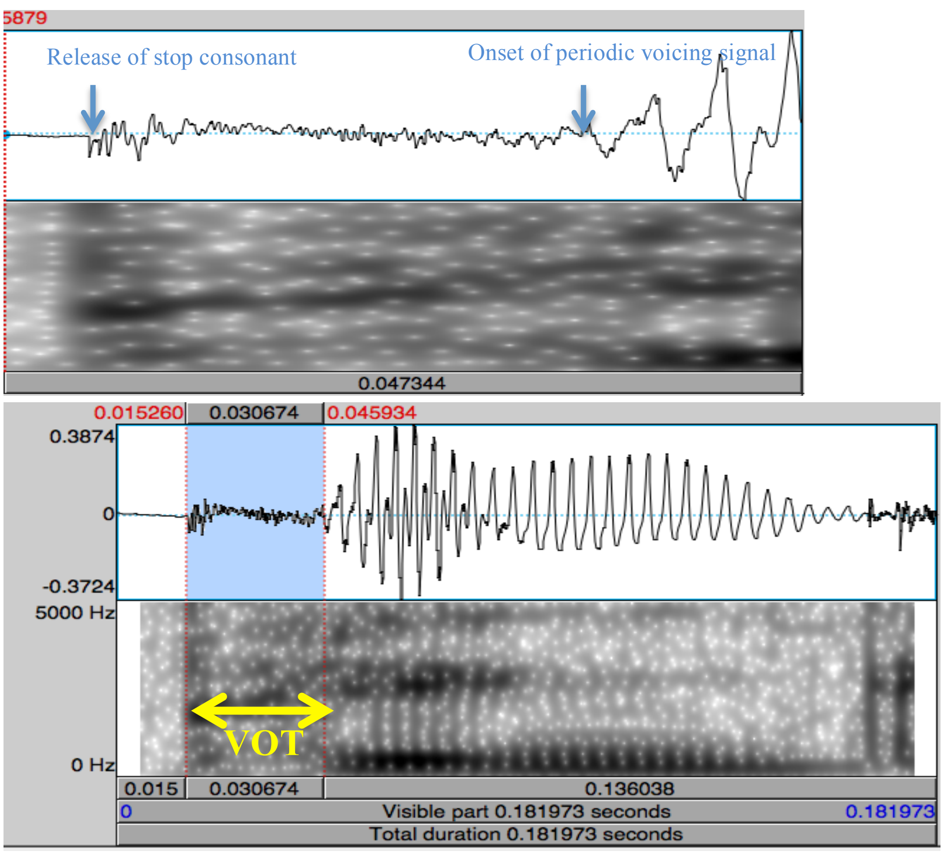

Figure 1. Measuring VOT for the word "pink" using the PRAAT display

Tokens were excluded from analysis if they were whispered, occurred in the presence of noise or an overlapping voice, if the stop was not actually present (e.g., “y’unna” instead of “you gonna”) or was fricated, and if the token was a sound effect or the mother was laughing or yawning. For the open class group tokens were also excluded if the stop occurred in an unstressed syllable. For the closed class group, tokens were also excluded when the vowel following the initial stop consonant was deleted. Figure 2 shows the number of open- and closed-class tokens which were excluded for each of these reasons. After the acoustic analysis was complete, two mothers from the open class group were dropped from further analyses because they no longer met the criteria after tokens had been excluded due to noise. A total of 2,264 tokens were included in the final analysis: 1,020 voiced bilabial stops, 508 voiceless bilabial stops, 416 voiced velar stops, and 320 voiceless velar stops. This represented 32.1% of the total number of open class tokens beginning with a bilabial stop consonant, and 30.7% of the total number of closed class tokens beginning with a velar stop consonant.

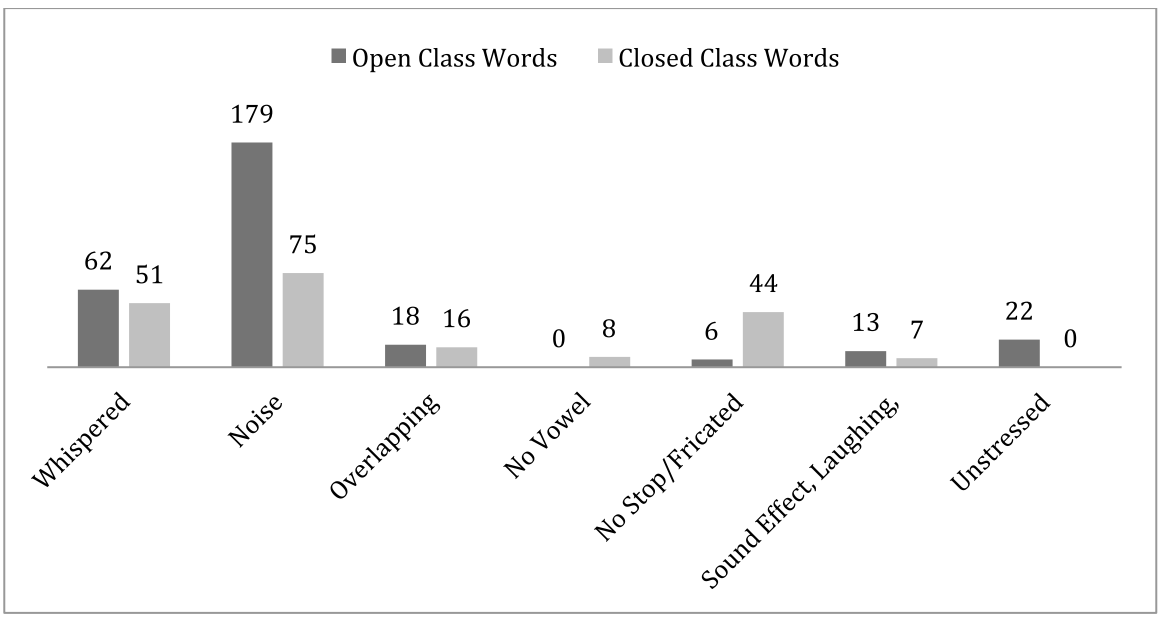

Figure 2. Number of open- and closed-class tokens excluded for each of the exclusion categories

Reliability

To estimate interrater reliability, two mothers were randomly chosen and VOT for all of the IDS and ADS tokens for those mothers were re-measured by another researcher. This represented approximately 10 percent of the total sample. Differences in VOT measures between the two raters were expressed in absolute numbers, and an average disagreement value was computed, to compare to any VOT changes seen across conditions. The mean difference between measurements was 3.5 milliseconds between the open class measurements, and 3.9 milliseconds between the closed class measurements. Pearson correlations were also computed between the two sets of measurements. Pearson's *r* between the open class tokens was 0.98, and between the closed class tokens was 0.95. Figures 3 and 4 show a Tukey Mean-Difference Plot for the open- and closed-class reliability tokens.

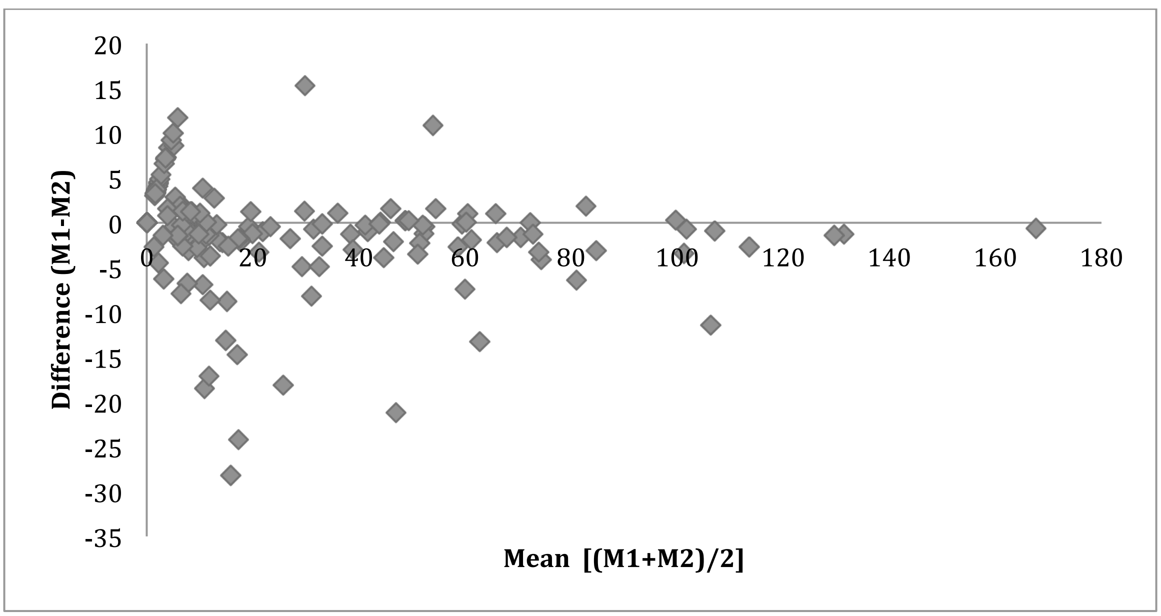

Figure 3. Tukey Mean-Difference Plot for open class reliability tokens.

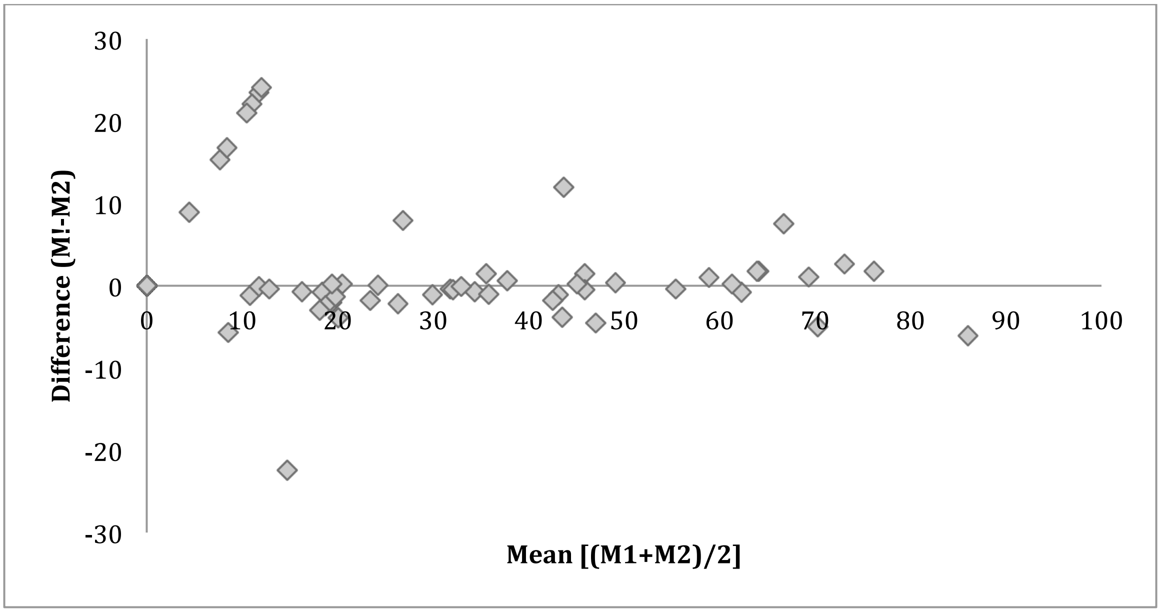

Figure 4. Tukey Mean-Difference Plot for closed class reliability tokens.

Data Analysis

In order to determine whether overall VOT differs between IDS and ADS at the different ages, a mean for voiced and voiceless items was calculated for each mother in each condition. Comparisons were made using a repeated measures ANOVA, with the within-subject variables of addressee (10/11 months, 18 months, 24 months, and Adult) and voicing (voiced or voiceless). This comparison allowed results from this study to be compared to previous VOT studies that did not use one of the calculations of VOT overlap. This test, as well as the following analyses, was all conduced in the statistical software package SPSS 20.

After mean VOT values were calculated, a variety of measures were used to quantify the degree of clarification between voicing categories in IDS and ADS. The first measure that was calculated was the acoustic measure of Discreteness of Voicing Category (DOVC). DOVC is defined as the difference in milliseconds between the longest VOT for a voiced stop and the shortest VOT for a voiceless stop across several productions of each (Zlatin, 1974). Negative values represent overlap between the categories, and positive values represent the degree of separation. DOVC values were calculated for each mother in each condition.

The percentage of overlap between voiced and voiceless sounds was also calculated. Moslin used this measure in her study (1979). The percentage overlap is defined as the proportion of voiceless stops that fall within plus or minus 2 standard deviations of the mean for voiced stops. This percentage was calculated for each mother in each condition. Percentages were then converted to arcsin values so that they could be used in parametric tests.

In addition to DOVC and percent overlap, $d_{(a)}$ was calculated for each mother in each condition. This measure was used in the study by Synnestvedt (2009), and was included in order to make direct comparisons between outcomes of the studies. This value is calculated by taking the difference in the means of the two categories times the square root of 2, divided by the square root of the sum of the variances. After all three of the clarification measures were calculated, correlations were calculated between the measures in order to see how well they agree, as they should be measuring the same thing.

To address the first hypothesis regarding the relationship between VOT clarification and infant language outcomes at 2 years, we calculated Pearson's *r* between each measure of clarification at 10/11, 18 and 24 months, and the scores on the *Peabody Picture Vocabulary Test* (PPVT), *Expressive One Word Vocabulary Test* (EOWVT), and the *MacArthur Communicative Development Inventories* (MCDI). These correlations were run separately between each measure of clarification and each of the assessments.

To see if VOT clarification of open class words varied by addressee, we performed repeated-measures ANOVAs separately for each of the clarification measures. The within-subject factor was addressee, with the 4 levels of 10/11-month child, 18-month child, 24-month child, or adult. We used Mauchly's Test of Sphericity to see if the assumption of sphericity was met for each of the clarification measures. When the significance of this test is <.05, the assumption of sphericity is violated, suggesting that using an uncorrected ANOVA would likely result in an inflation of Type I errors. When the assumption of sphericity was not met we applied the Greenhouse-Geisser correction. This correction raises the critical F value needed to reject the null hypothesis, and thus reduces the likelihood of Type I errors. This was chosen over using a multivariate approach due to the relatively small sample size. Significant within-subjects main effects were investigated using post-hoc Fisher's Protected *t* Tests. The same method was used for the closed class group in order to see if VOT clarification of closed class words varied by addressee.

Results

VOT Duration-Open Class Words

A repeated-measures ANOVA of the open class group with the within-subject factors of addressee and voicing revealed no difference in VOT duration between the different addressees ($F(3, 45)=.314$, $p=.815$), and no significant interaction between addressee and voicing ($F(3, 45)=2.063$, $p=.119$). As expected, there was a significant main effect of voicing ($F(1, 45)=326.548$, $p<.001$). VOT duration of voiceless sounds was longer than voiced sounds at all ages. The mean VOT durations for open class words to each addressee are depicted in Figure 5.

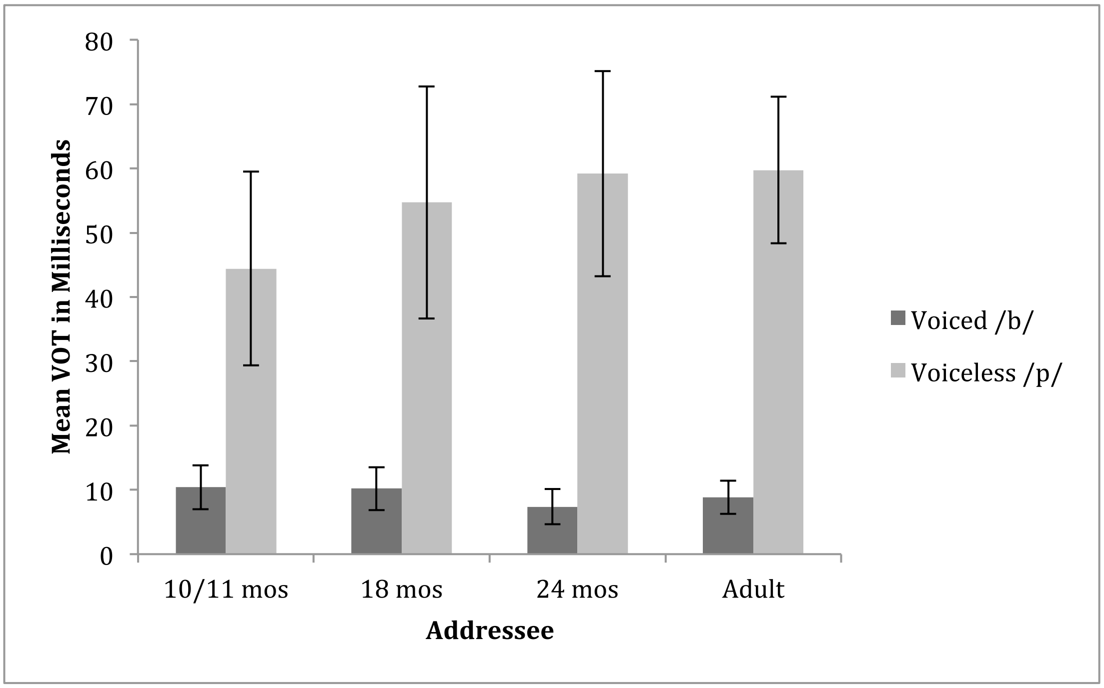

Figure 5. Mean VOT Values for open class words. Error bars represent plus or minus 1 standard deviation

VOT Duration-Closed Class Words

A repeated-measures ANOVA with the within-subject factors of addressee and voicing was also computed for the closed class group. There was no significant effect of addressee ($F(3, 27)=1.992$, $p=.139$), and no interaction between addressee and voicing ($F(3, 27)=.438$, $p=.727$). There was a significant main effect of voicing $F(1, 9)=253.079$, $p<.001$), with voiceless stops having a significantly longer VOT than voiced stops. The mean VOT durations for the initial voiced or voiceless velar stops of closed class words to each addressee can be seen in Figure 6.

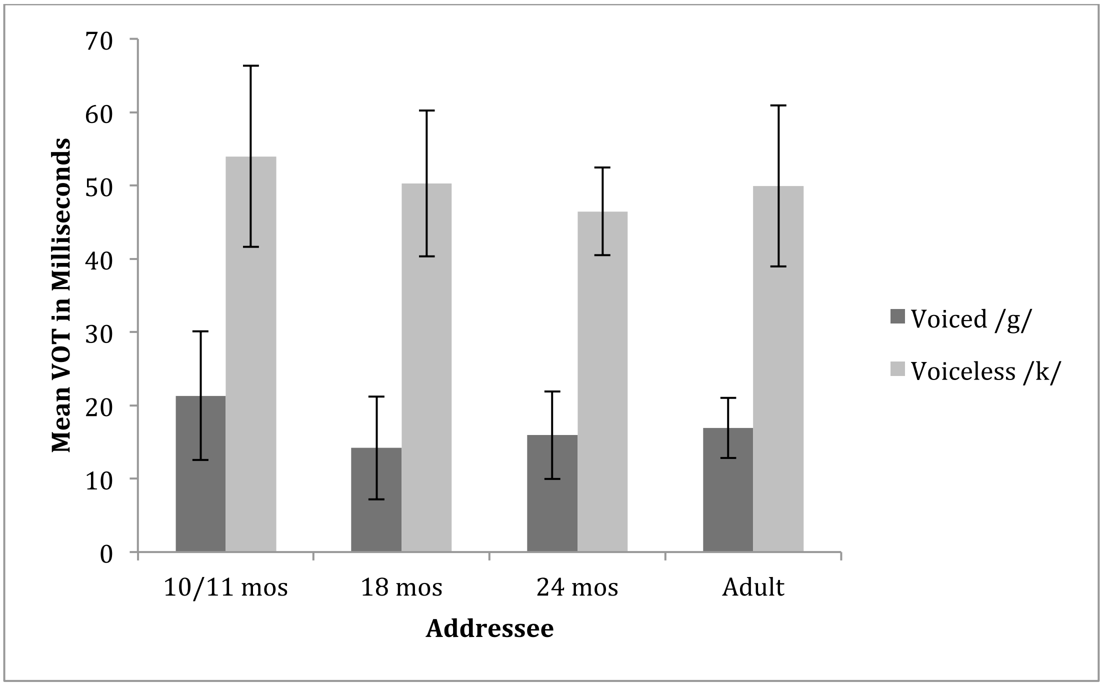

Figure 6. Mean VOT values for closed class words. Error bars represent plus or minus 1 standard deviation.

Agreement Between VOT Clarification Measures

Before any analyses were computed using the VOT clarification measures, we looked at the agreement between the measures at each age. These comparisons were made using Pearson product-moment correlations (see Table 1). The measures of DOVC and $d_{(a)}$ were significantly correlated at all ages for both open- and closed-class groups. These two measures were most strongly related in the closed class group at 18 months (r = -.747, p<.05). Percent Overlap (converted to arc sin values) was significantly correlated with DOVC for both open- and closed-class groups when infants were 10/11 months and 18 months old. In speech to the infants at 24 months, the relationship between these two measures was not significant for the closed class words (r = .558, p >.05), and in speech to the adult this relationship was not significant for the open class words (r = .453, p >.05). The weakest relationship was between the measures of $d_{(a)}$ and percent overlap. These measures were not significantly related for open or closed class groups in ADS. In speech to the 10/11 month-old infants, these measures were significantly related in the closed class group (r = -.900, p<.01), but not in the open class group (r = -.388, p >.05). The same relationship between these measures was found in speech to the 24 month-old children, with a significant correlation in the closed class group (r = -.859, p<.01), but not in the open class group (r = -.388, p >.05).

Table 1

Correlations Between Measures to 10/11 Month-old Children

	DOVC	$d_{(a)}$	Percent Overlap
DOVC			
Open Class	_____	-.597 (.015)*	.689 (.003)**
Closed Class	_____	-.747 (.013)*	.686 (.028)*
$d_{(a)}$			
Open Class		_____	-.468 (.068)
Closed Class		_____	-.900 (<.001)**

Correlations Between Measures to 18 Month-old Children

	DOVC	$d_{(a)}$	Percent Overlap
DOVC			
Open Class	_____	-.718 (.002)**	.740 (.001)**
Closed Class	_____	-.856 (.002)**	.851 (.002)**
$d_{(a)}$			
Open Class		_____	-.655 (.006)**
Closed Class		_____	-.954 (<.001)**

Correlations Between Measures to 24 Month-old Children

	DOVC	$d_{(a)}$	Percent Overlap
DOVC			
Open Class	--	-.743 (.001)**	.639 (.008)**
Closed Class	--	-.750 (.012)*	.558 (.094)
$d_{(a)}$			
Open Class		_____	-.388 (.138)
Closed Class		_____	-.859 (.001)**

Correlations Between Measures to Adult

	DOVC	$d_{(a)}$	Percent Overlap
DOVC			
Open Class	_____	-.663 (.005)**	.453 (.078)
Closed Class	_____	-.756 (.011)*	.946 (>.001)**
$d_{(a)}$			
Open Class		_____	-.284 (.286)
Closed Class		_____	-.605 (.064)

Note. *p < .05, **p < .01. N =16 for the open class group and N =10 for the closed class group

Comparison of VOT Clarification of Open Class words by Addressee

Comparisons of VOT clarification of open class words by addressee were made separately for each of the clarification measures. The descriptive statistics used to make these comparisons are summarized in Table 2. The assumption of sphericity was violated for all measures (DOVC W=.258, p<.05; $d_{(a)}$ W=.252, p<.05; Percent Overlap W=.415, p<.05). In order to correct for this violation, the Greenhouse-Geisser adjustment was used for all tests of within-subject effects. The results of these tests are summarized in Table 3.

There were no significant differences in VOT clarification of open class words across the addressees for the measures of DOVC (F(1.772, 26.585)=2.565, p=.101) and $d_{(a)}$ (F(1.606, 24.084)=1.023, p=.359). For DOVC, there was more overlap to the children at 10/11 months old and 18 months old, and less overlap to the 24 month olds and Adults. The measure of $d_{(a)}$ also increased as the infants got older; $d_{(a)}$ in IDS was the least to the youngest infants at 10/11 months (M=2.606) and the greatest to the 24 month olds (M=3.135). Again, these were non-significant differences. These trends are compared in Figure 7. Although there were significant differences in rate of speech across the addressees ($F(3, 45)=81.694$, $p<.001$) with speech rate increasing at each age, rate of speech was not significantly correlated with percent overlap at any age.

Table 2

Descriptive Statistics of VOT Clarification Measures for Open Class Words

	Mean	*Std. Deviation*	*N*
DOVC			
10/11months	-4.475	19.2530	16
18 months	-4.950	12.0833	16
24 months	-13.513	12.1453	16
Adult	-13.406	10.9595	16
$d_{(a)}$			
10/11months	2.606	1.0923	16
18 months	2.977	.9705	16
24 months	3.135	1.9926	16
Adult	3.322	.9293	16
Percent Overlap			
10/11months	12.165	15.7864	16
18 months	10.416	12.4650	16
24 months	6.175	9.58239	16
Adult	1.506	6.0244	16

Table 3

Within-Subject Effects for Clarification Measure of Open Class Words

Measure	*df*	*F*	*p value*
DOVC	1.772	2.565	.101
$d_{(a)}$	1.606	1.023	.391
Percent Overlap	2.149	3.2881	.047*

Note. *$p<.05$, Greenhouse-Geisser adjustment applied

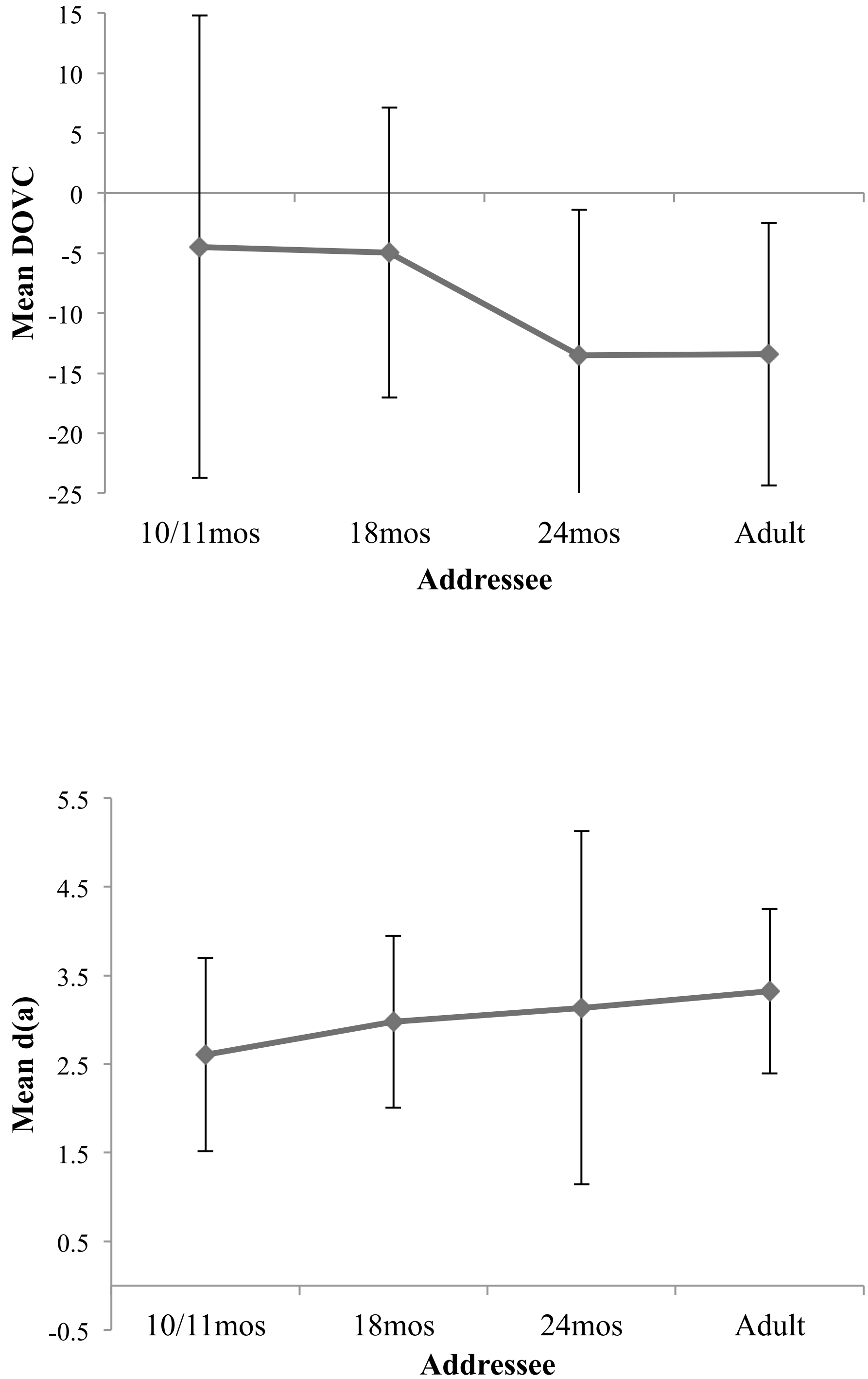

Figure 7. Means of DOVC (top) and $d_{(a)}$ (bottom) for open class words

Percent overlap (converted to arcsin units) was also compared between open class words to the different addressees. A statistically significant difference was found in percent overlap of open class words across addressees (F(2.149, 32.238)=3.288, p=.047). This effect can be seen in Figure 8. Post-hoc analysis with Fisher's Protected t Test was conducted. There was a significant difference in the VOT overlap of open class words to the child at each age when compared to the adult. The amount of overlap to the adults (M=1.506, SD=6.024) was significantly less than to the child at 10/11 months old (M=12.165, SD= 15.786; t(15)=2.367, p=.032), 18 months old (M=10.416, SD=6.1754; t(15)=2.375, p=.031), and 24 months old (M=6.175, SD=1.5061; t(15)=2.224, p=.042). These findings are summarized in Table 4.

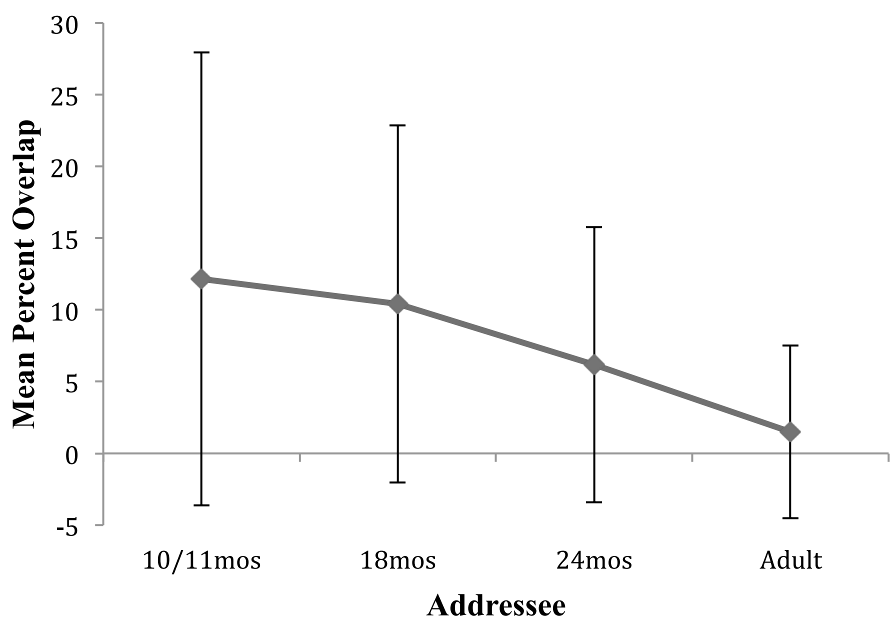

Figure 8. Means of Percent Overlap (converted to Arcsin units) for open class words

Table 4

Fisher's Protected t Test for Overlap of Open Class Words

	t	*df*	*Sig. (2-tailed)*
10/11 mos – 18 mos	.425	15	.677
10/11 mos – 24 mos	1.390	15	.185
10/11 mos – Adult	2.367	15	.032*
18 mos – 24 mos	1.467	15	.163
18 mos – Adult	2.375	15	.031*
24 mos – Adult	2.224	15	.042*

*p<.05

Comparison of VOT Clarification of Closed Class words by Addressee

In order to address whether VOT of closed class words was clarified more to different addressees, repeated measure ANOVAs were computed for each of the clarification measures, with the within-subject factor of addressee. The means and standard deviations of the clarification measures are reported in Table 5. For the measures of DOVC and percent overlap the assumption of sphericity was met (DOVC W=.270, p>.05; percent overlap W=.430, p>.05) therefore no corrections were applied. For the measure of $d_{(a)}$, the assumption of sphericity was violated (W=.180, p<.05), so the Greenhouse-Geisser adjustment was used. There were no significant differences found for any of the clarification measures of closed class words across the different addressees (DOVC $F(3, 27)=2.115$, p=.122; $d_{(a)}$ $F(1.829, 16.461)=1.009$, p=.379; Percent Overlap $F(3, 27)=1.455$, p=.249). The results of these tests are summarized in Table 6. As with the open class words, there were no significant relationships between rate of speech and VOT clarification of closed class words.

Table 5

Descriptive Statistics for VOT Clarification Measures of Closed Class Words

	Mean	*Std. Deviation*	*N*
DOVC			
10/11months	13.100	11.6545	10
18 months	3.360	13.8152	10
24 months	6.420	14.6383	10
Adult	.910	12.2952	10
$d_{(a)}$			
10/11months	1.968	.6019	10
18 months	2.706	1.0198	10
24 months	2.652	.9703	10
Adult	2.707	1.6049	10
Percent Overlap			
10/11months	35.131	18.7678	10
18 months	23.630	23.1639	10
24 months	31.328	19.8140	10
Adult	17.946	17.6098	10

Table 6

Within-Subject Effects for Clarification Measure of Closed Class Words

Measure	*df*	*F*	*p value*
DOVC	3	2.115	.122
$d_{(a)}$ [a]	1.829	1.009	.379
Percent Overlap	3	1.455	.249

Note. [a]The Greenhouse-Geisser Adjustment was applied for $d_{(a)}$

For all clarification measures, the greatest VOT clarification of closed class words in IDS occurred at 18 months, followed by 24 months. The least clarification occurred at 10/11 months. VOT clarification of closed class words was the greatest in speech to the adults. Figures 9 through 11 show this trend for each of these measures.

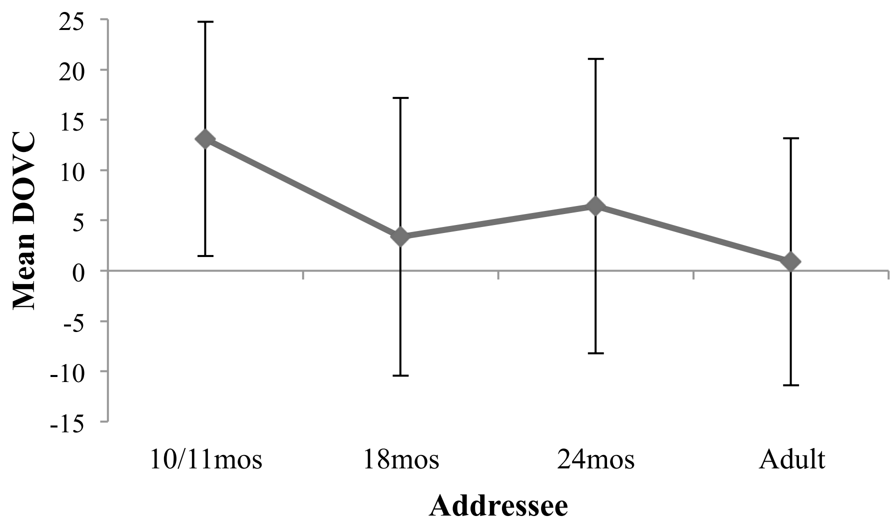

Figure 9. Mean DOVC for closed class words

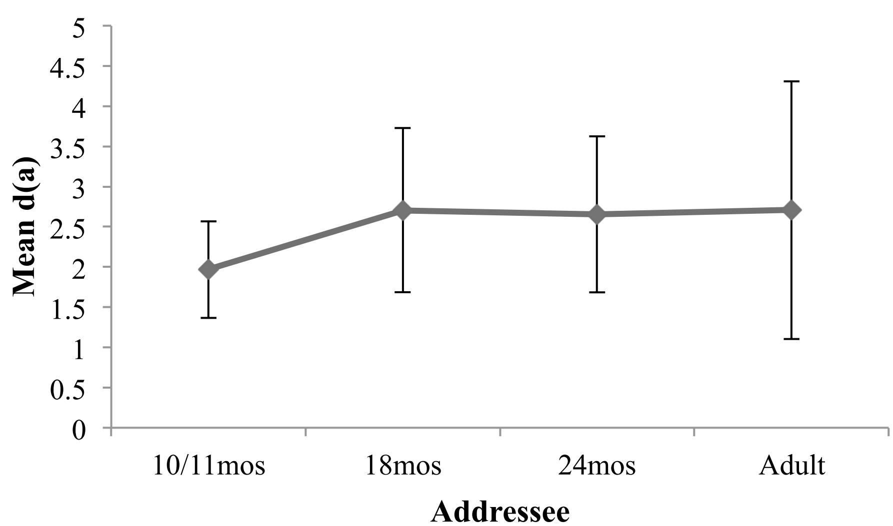

Figure 10. Mean $d_{(a)}$ for closed class words

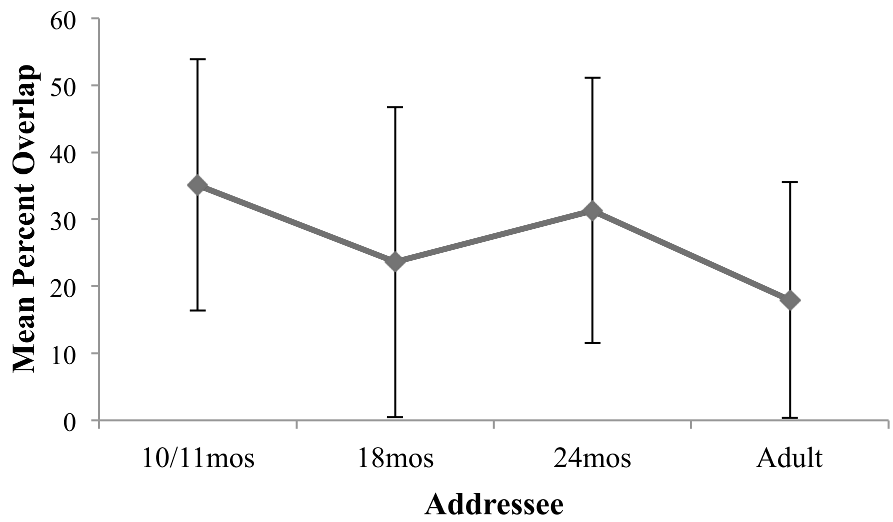

Figure 11. Mean percent overlap for closed class words

VOT Clarification of Open Class Words and Infant Language Outcomes

Each measure of VOT clarification was compared with language outcomes at 2 years using a Pearson product-moment correlation. No significant correlations were found between the clarification measures and infant language outcomes at any age. The results of these correlations for open class words are summarized in Table 7. Non-significant negative correlations were found between DOVC and language outcomes at 10/11 months, with less overlap associated with higher scores on all of the language tests (PPVT r = -.028, p=.919; EOWVT r = -.199, p=.460; MCDI r = -.170, p=.528). This trend did not continue at 18 months (PPVT r = .202, p=.452; EOWVT r = .356, p=.176; MCDI r = .045, p=.869) or 24 months (PPVT r = .206, p=.443; EOWVT r = .366, p=.164; MCDI $r(14)$=.104, p=.701).

Although they were non-significant, the relationships between the measure of $d_{(a)}$ and the language outcomes were the opposite of what we expected. As $d_{(a)}$ increased and

VOT became more clarified, language outcome scores decreased. This was true of all relationships except for that between $d_{(a)}$ at 10/11 months and the PPVT raw score, in which there was a positive correlation (r = .116, p=.668).

The final measure of VOT clarification of open class words that we looked at was percent overlap. Percentages were converted into arcsin values in order to meet the assumptions underlying parametric tests. No significant correlations were found between percent overlap in open class words and infant language outcomes at any age.

Table 7

Correlations between VOT Clarification Measures of Open Class Words and Language Outcomes

	PPVT	*EOWVT*	*MCDI*
DOVC			
10/11 months	-.028 (.919)	-.199 (.460)	-.170 (.528)
18 months	.202 (.452)	.356 (.176)	.045 (.869)
24 months	.206 (.443)	.366 (.164)	.104 (.701)
$d_{(a)}$			
10/11 months	.116 (.668)	-.163 (.545)	-.051 (.850)
18 months	-.111 (.681)	-.115 (.672)	-.002 (.995)
24 months	-.235 (.381)	-.069 (.799)	-.011 (.968)
Percent Overlap			
10/11 months	.155 (.566)	.219 (.415)	.198 (.463)
18 months	.007 (.980)	.055 (.840)	-.195 (.469)
24 months	-.267 (.318)	.074 (.785)	-.251 (.349)

Note. No measures significant at p<.05. N=16 for all analyses.

VOT Clarification of Closed Class Words and Infant Language Outcomes

We also examined whether VOT clarification of closed class words was related to infant language outcomes. As with the open class group, a Pearson product-moment correlation was calculated for each addressee between each measure of VOT clarification and each language outcome score. The results of these analyses are summarized in Table 8. The first VOT clarification measure we investigated was DOVC. No significant correlations were found between this measure and any of the language outcomes. However, the direction of the non-significant relationship between DOVC and language outcome scores was as expected, with less overlap associated with higher scores.

Table 8

Correlations Between VOT Clarification Measures of Closed Class Words and Language Outcomes

	PPVT	*EOWVT*	*MCDI*
DOVC			
10/11 months	-.455 (.187)	-.185 (.610)	-.580 (.079)
18 months	-.431 (.213)	-.305 (.391)	-.598 (.068)
24 months	-.383 (.275)	-.071 (.845)	-.364 (.301)
$d_{(a)}$			
10/11 months	.450 (.191)	.262 (.465)	.573 (.084)
18 months	.605 (.064)	.548 (.548)	.604 (.065)
24 months	.191 (.597)	-.035 (.924)	.202 (.576)
Percent Overlap			
10/11 months	-.360 (.307)	-.214 (.552)	-.630 (.051)
18 months	-.593 (.071)	-.583 (.077)	-.687 (.028)*
24 months	.035 (.924)	.065 (.859)	-.050 (.891)

Note. **p*<.05. *N*=10 for all analyses.

No significant relationships were found between $d_{(a)}$ of closed class words at any age, and the infant language outcomes at 2 years. However, at 18 months, the correlation between $d_{(a)}$ and the PPVT and MCDI approached significance. As mothers increased the degree of separation between voiced and voiceless stops, their children's receptive vocabulary skills appeared to improve (PPVT r = .605, p=.064; MCDI r = .604, p=.065).

Prior to running Pearson product-moment correlations for the last VOT clarification measure, percentage overlap, percentages were converted to arcsin values. No significant correlations were found between the arcsin values at 10/11 months and the language outcomes, although the correlation between the arcsin values at 10/11 months and the MCDI raw score did approach significance. As percent overlap decreased and VOT became more clarified at 10/11 months, MCDI scores improved (r = -.630, p=.051). The relationship between percent overlap at 10/11 months and the other language measures did not reach significance, although they did follow the same trend (PPVT r = -.360, p=.307; EOWVT r = -.214, p=.552). A significant negative correlation was found between percent overlap at 18 months and the MCDI raw score. At 18 months less overlap was associated with higher MCDI scores (r = -.762, p=.010). A scatterplot of MCDI scores against percent overlap can be found in Figure 12. The same trend was found with the other language outcome scores, although they did not reach levels of significance (PPVT r = -.593, p=.071; EOWVT r = -.583, p=.077). The percent overlap at 24 months was not significantly correlated with language outcomes (PPVT r = .035, p=.924; EOWVT r = .065 p=.859; MCDI r = -.050, p=.891).

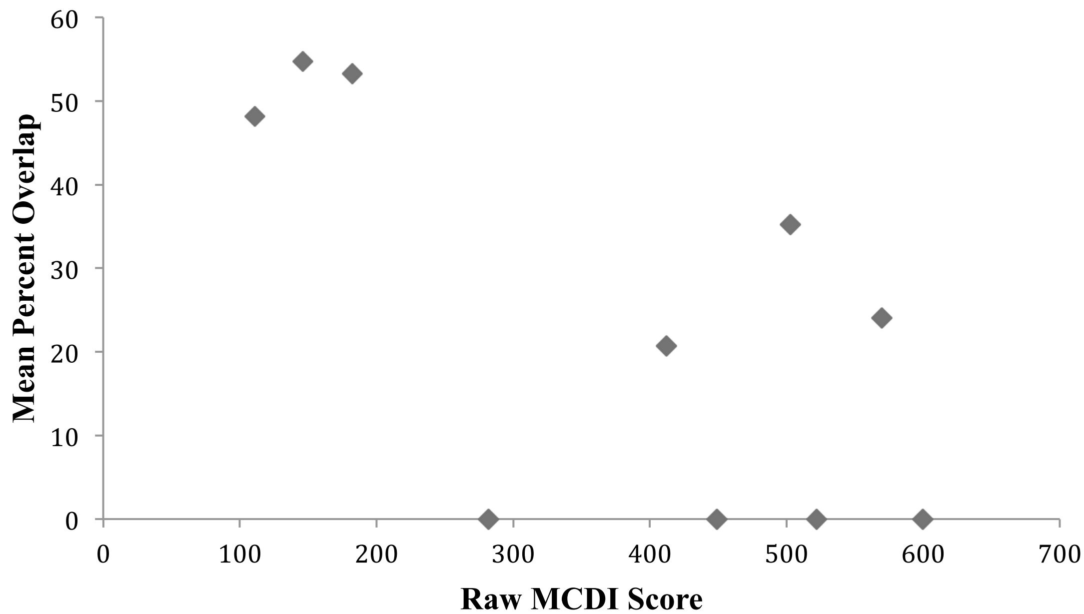

Figure 12. Percent overlap of closed class words at 18 months plotted against MCDI raw scores

Group Comparison of Language Outcomes

As a follow-up to the correlations, mothers were split into two groups at each age-those with more VOT clarification, and those with less VOT clarification. Mothers were included in the group with more VOT clarification if they fell into the half of the distribution associated with more clarification (lower 50% for DOVC and percent overlap, or upper 50% for $d_{(a)}$) for at least 2/3 of the clarification measures. The remaining mothers made up the group with less VOT clarification. Outcome scores for the 2 groups of participants were compared using paired *t* tests. Average outcome measures for each group and *t*-test results are shown in Table 9. For the open class groups, there was no significant difference in the language outcome scores between the two clarification groups. The More VOT Clarification Group had higher mean MCDI scores than the Less VOT Clarification group at all ages, but the differences were non-significant. At 10/11 months, the More VOT Clarification Group also performed better on the PPVT ($t(7)$=.511) and EOWVT ($t(7)$=.152), but again the differences were not significant.

Table 9

Open Class Group Comparisons, 10/11 Months

	More VOT Clarification Group	*Less VOT Clarification Group*	*t scores*
PPVT mean raw scores	38.875	30.125	t=.511 p=.625
EOWVT mean Std. Score	96.5	95.125	t=.152 p=.883
MCDI mean raw scores	414	405	t=.088 p=.932

Open Class Group Comparisons, 18 Months

	More VOT Clarification Group	*Less VOT Clarification Group*	*t scores*
PPVT mean raw scores	27.375	36.6250	t=-1.240 p=.255
EOWVT mean Std. Score	92.5	99.125	t=-.665 p=.527
MCDI mean raw scores	418.5	400.5	t=.167 p=.872

Open Class Group Comparisons, 24 months

	More VOT Clarification Group	*Less VOT Clarification Group*	*t scores*
PPVT mean raw scores	30.375	33.625	t=-.588 p=.575
EOWVT mean Std. Score	91.75	99.875	t=-1.197 p=.270
MCDI mean raw scores	433.5	385.5	t=.537 p=.608

Note. No measures significant at p<.05, N=8 for all groups

For the closed class groups, a significant difference in MCDI scores was found between the two groups at 10/11 months. The group that received more VOT clarification had a significantly higher mean MCDI score ($t(4)=2.886$, $p<.05$). The More VOT Clarification Group at 10/11 months also had higher scores than the Less VOT Clarification Group on the PPVT ($t(4)=.680$, $p>.05$) and EOWVT ($t(4)=.897$, $p>.05$) but these differences were not statistically significant. The group that received more VOT Clarification at 18 months performed better than the other group on all language outcome measures, but these differences did not reach significance (PPVT $t(4)=.985$, $p>.05$; EOWVT $t(4)=1.354$, $p>.05$; MCDI $t(4)=1.851$, $p>.05$). The results of these group comparisons are summarized in Table 10.

Table 10

Closed Class Group Comparisons, 10/11 months

	More VOT Clarification Group	*Less VOT Clarification Group*	*t scores*
PPVT mean raw scores	36	28.8	$t=.680$ $p=.534$
EOWVT mean Std. Score	98	88	$t=.897$ $p=.420$
MCDI mean raw scores	480.8	274.6	$t=2.886$ $p=.045$*

Closed Class Group Comparisons, 18 months

	More VOT Clarification Group	*Less VOT Clarification Group*	*t scores*
PPVT mean raw scores	37.8	27	$t=.985$ $p=.381$
EOWVT mean Std. Score	99	87	$t=1.354$ $p=.247$
MCDI mean raw scores	453	302.4	$t=1.851$ $p=.138$

Closed Class Group Comparisons, 24 months

	More VOT Clarification Group	*Less VOT Clarification Group*	*t scores*
PPVT mean raw scores	32.8	32	*t*=.120 *p*=.910
EOWVT mean Std. Score	88.4	97.6	*t*=-1.407 *p*=.232
MCDI mean raw scores	397.2	358.2	*t*=.689 *p*=.529

Note. *p<.05. N=5 for each group.

Comparison Between VOT Clarification of Open- and Closed-Class Words

For the nine mothers who met the criteria for both the open- and closed-class analyses, we compared VOT clarification across these word classes. There was a significant main effect of word class for DOVC (F(1, 8)=14.908, p=.005). Follow-up t tests revealed that this effect was present to the 24 month-old infants (t(8)=-2.912, p=.02) and to adults (t(8)=-4.017, p=.004). The mean DOVC values for each age are represented in Figure 13. There was also a significant main effect of word class for percent overlap (F(1, 8)=23.165, p=.001), with significantly less overlap of open class words compared to closed class words to infants at 10 months (t(8)=-2.781, p=.024), 24 months (t(8)=-3.187, p=.013) and to adults (t(8)=-3.43, p=.009). The mean values for percent overlap are shown in Figure 14. There were no significant differences in $d_{(a)}$ between the open- and closed-class words. The mean values for this clarification measure can be seen in Figure 15.

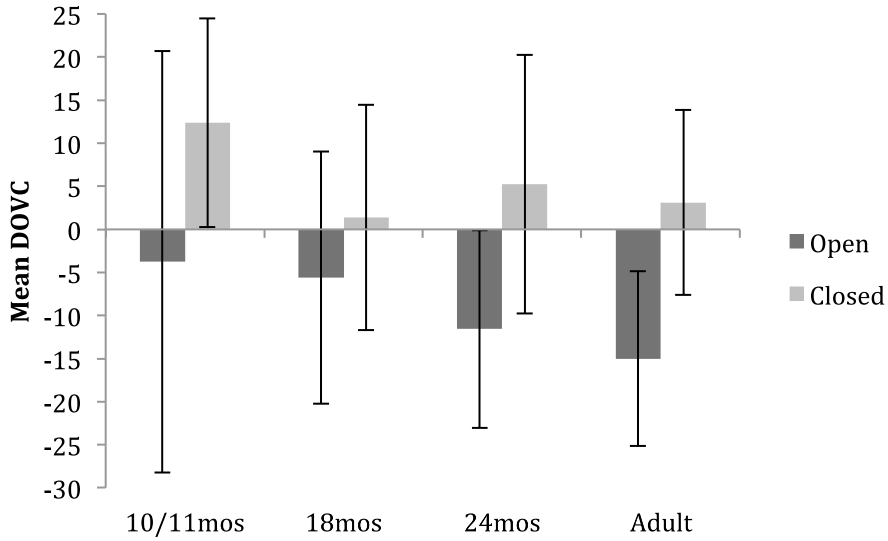

Figure 13. Comparison of mean DOVC for open- vs. closed-class words

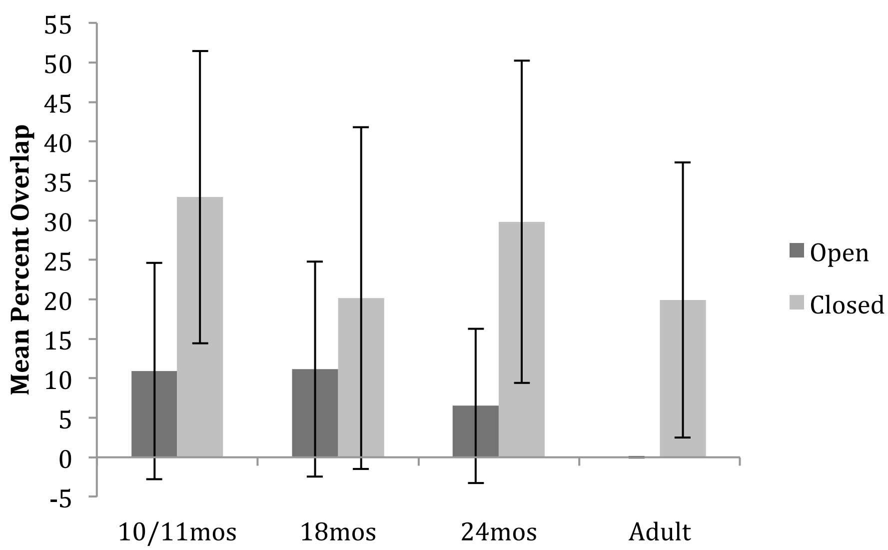

Figure 14. Comparison of mean percent overlap for open- vs. closed-class words

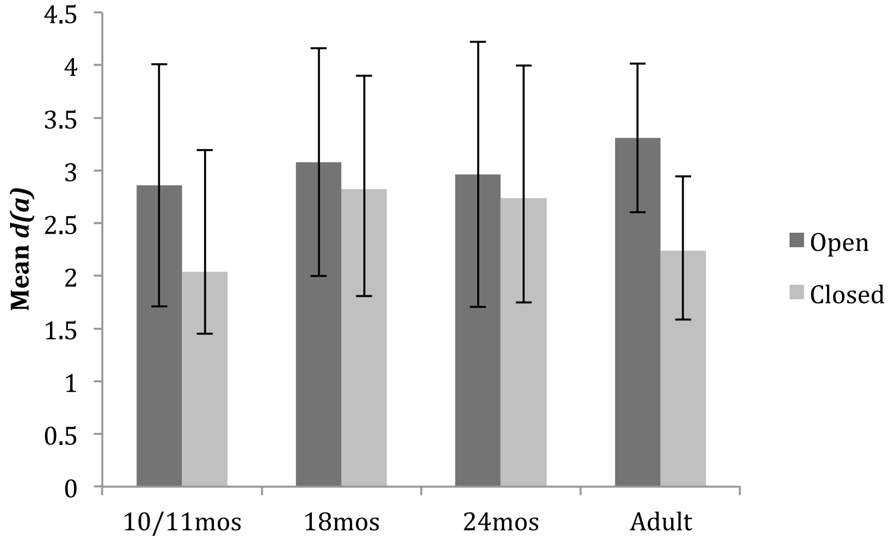

Figure 15. Comparison of mean $d_{(a)}$ for open- vs. closed-class words

Type-Token Ratio Comparisons and Effects

Another factor previously mentioned to affect clarification is whether a word is "given" or "new". Type-token ratio (TTR) is the number of different words divided by the total number of words. Thus, low TTRs may reflect fewer "new" tokens and a greater degree of repetition, whereas high TTRs reflect greater "new" words and less repetition. We used TTR as a way of roughly measuring the effect of given versus new words by first comparing TTR to the different addressees, and then looking for correlations between TTR and VOT clarification measures. The descriptive statistics for the 16 mothers in the open class group and the 10 mothers in the closed class group are reported in Table 11. A repeated-measures ANOVA with the within-subject factor of addressee was computed for each of these groups. The assumption of sphericity was not met for the open class group, therefore the Greenhouse-Geisser correction was applied. There was no significant difference in TTR across the addressees ($F(1.014, 15.215)=.985$, $p=.338$). For the closed class group, there was a significant effect of addressee on TTR ($F(3, 27)=23.791$, $p<.001$). A follow-up *t* test revealed significant differences in TTR between each addressee, with the smallest TTR to infants at 24 months ($M=.223$), then 18 months ($M=.248$), then 10 months ($M=.279$), and the highest TTR to adults ($M=.345$). The findings from these *t* tests are summarized in Table 12.

Table 11

Descriptive Statistics of TTR for Open- and Closed-Class Groups

	Mean TTR	*SD*	*N*
10/11months			
Open	.498	.79983	16
Closed	.279	.04844	10
18 months			
Open	.415	.63162	16
Closed	.248	.03560	10
24 months			
Open	.275	.17670	16
Closed	.223	.03374	10
Adult			
Open	.356	.06198	16
Closed	.345	.05460	10

Table 12

Fisher's Protected t Tests for Closed Class Group

	t	*df*	*Sig. (2-tailed)*
10/11 mos – 18 mos	2.597	9	.029*
10/11 mos – 24 mos	5.144	9	.001**
10/11 mos – Adult	-3.187	9	.011*
18 mos – 24 mos	2.257	9	.050*
18 mos – Adult	-6.960	9	<.001**
24 mos – Adult	-6.126	9	<.001**

Note. *$p<.05$, **$p<.001$

In addition to comparing TTR across addressees, we also looked for relationships between TTR and VOT clarification measures. The results of these Pearson's correlations are listed in Table 13 for open class words, and Table 14 for closed class words. There were no significant correlations between TTR and DOVC or $d_{(a)}$ of open class words at any age. There was one significant correlation between TTR to infants at 10/11 months and the VOT clarification measure of percent overlap ($r = .599$, $p=.014$), with a higher TTR associated with greater overlap of open class words. There were no significant correlations between TTR and VOT clarification of closed class words.

Table 13

Pearson's Correlations Between TTR and VOT Clarification Measures for Open Class Words

	DOVC	$d_{(a)}$	Percent Overlap
TTR 10 months	.141 (.561)	-.265 (.321)	.599 (.014)*
TTR 18 months	.175 (.516)	-.219 (.416)	.226 (.401)
TTR 24 months	.093 (.733)	-.110 (.686)	-.204 (.448)
TTR Adult	.002 (.993)	.115 (.673)	.077 (.778)

Note. *$p<.05$. N=16 for all analyses.

Table 14

Pearson's Correlations Between TTR and VOT Clarification Measures for Closed Class Words

	DOVC	$d_{(a)}$	Percent Overlap
TTR 10 months	-.378 (.281)	-.023 (.949)	-.129 (.722)
TTR 18 months	-.125 (.730)	-.041 (.911)	.135 (.709)
TTR 24 months	-.339 (.338)	-.237 (.510)	.207 (.566)
TTR Adult	-.166 (.647)	.393 (.261)	.040 (.912)

Note. No measures significant at $p<.05$. N=10 for all analyses.

Discussion

The present study was designed to evaluate the characteristics of VOT in mothers' speech to infants and adults, and to determine whether these characteristics correlated with infant language outcomes at 2 years. Specifically we wondered if mothers would clarify VOT of open- and closed-class words differently at ages corresponding to different stages of language development. In order to answer these questions we conducted separate analyses of open- and closed-class words in IDS at 10/11, 18 and 24 months, and in ADS.

Open Class VOT Characteristics

We hypothesized that VOT clarification of open class words would be greatest in speech to the 18 month-old children. Contrary to expectations, there were no statistically significant differences on any of the clarification measures in speech to the infants at 10/11, 18 or 24 months. The only difference found between the IDS and ADS groups was that the percent overlap for open class words was significantly less in speech to the adults than in speech to the children at any age. Despite the lack of significant findings, there were some interesting trends in the clarification measures for open class words that are worth noting. First of all, at this youngest age of 10/11 months, around when the infants were probably just preceding or beginning to use single words, mothers as a whole did not provide the children with distinct voicing information of open class words. The least amount of VOT clarification occurred at this age across all three measures. The amount of clarification then increased at each age, with the greatest clarification of IDS occurring at the 24-month visit. Unexpectedly, VOT clarification was the greatest to adults for the measures of $d_{(a)}$ and percent overlap, but not for DOVC.

We looked at type-token ratio (TTR) as a way to see if VOT clarification was influenced by the amount of repetition versus new words in speech to the infants at different ages. TTR is a measure of the number of different tokens over the number of total tokens produced. Thus, greater TTR represents less repetition of the same tokens, and a greater proportion of "new" words. We had expected that mothers who repeated less might have greater overall VOT clarification because previous research has shown that speakers use greater precision for words that are not redundant (Fowler & Housum, 1971). However, in the current study this was not the case. Although there were no significant differences in TTR in IDS across the different ages, we did find a significant positive correlation between TTR and percent overlap at 10/11 months with higher TTR associated with greater overlap of their VOT tokens. In other words, the higher the proportion of new words the mothers used in speech to infants at 10/11 months, the less they clarified their VOT. One possible explanation for these findings is that these two variables (TTR and VOT clarification) are both being influenced by another factor-the child's language abilities. At this young age when the infants are not yet producing single words, the mothers use a greater variety of words rather than repeating the same word multiple times in order to elicit imitation, and they also are not attempting to assist in the infant's production of words by providing phonetic clarification.

There are some interesting similarities and differences between the open class word VOT characteristics of IDS and ADS found in the present study and those found in previous VOT studies. In particular, our findings that percent overlap was actually greater in IDS than ADS at all ages directly contradict those of Malsheen (1980), who found that there was a significant *reduction* in overlap of VOT values in the mothers' speech to children producing single words. This reduction was due to significantly higher mean VOT values for initial voiceless stops in the IDS condition. In the present study, there were no significant differences in VOT values between IDS and ADS at any age. There are a number of reasons for these discrepant findings. First of all, Malsheen's study looked at the VOT characteristics of 6 mother-infant dyads, 2 for each stage of language development. The present study consisted of 16 mother-infant dyads who were followed longitudinally. The greater number of participants in the current study likely led to greater variability in VOT. Given that percent overlap is a measure that is calculated based on group means and standard deviations, greater variability would effect this calculation of VOT clarity. Furthermore, the current study looked only at the bilabial place of articulation. In Malsheen's study VOT was investigated for all stop consonants, and when these phonemes were separated by place of articulation there were actually individual trends in which phonemes the mothers clarified. So while altogether VOT overlap was less in the IDS condition for the two mothers whose children were using single words, the first mother clarified the voiced-voiceless distinction more for bilabials than alveolar stops, and did not clarify the velar cognates at all. The second mother, on the other hand, provided the *most* clarification for the velar voiced-voiceless contrast. In the current study, by looking only at one place of articulation, we may have missed the individual trends of differences in the other stop consonants. This is one limitation of the current study.

One possible explanation for the finding that VOT clarification was greater to the adults than to the children stems from the situation in which the ADS was obtained. These speech samples were collected from interviews after each of the play sessions in which the mother was asked about what her child played with and how this is similar or different to what her child likes to play with at home. Furthermore, the interviewer was often an unfamiliar adult to the mother. Although they were not intended to be formal interviews, the mothers may have used a more formal register because of the interview structure than they would have in a casual conversation with a friend. Thus, we may have obtained a somewhat biased sample of ADS with greater phonetic clarification. Future studies can address this issue by obtaining ADS samples in more casual situations between familiar adults.

Despite the differences from Malsheen's (1980) findings, the present study is not the first to find no significant difference in mean VOT values between IDS and ADS. Baran et al. (1977) and Englund (2005) reported similar findings. Synnestvedt (2009) also found that there was no significant difference in the VOT of voiceless stops between IDS and ADS, but that voiced stops were significantly longer in IDS, leading to less clarification in IDS at 7.5 and 11 months compared to ADS. However, all of these studies investigated VOT in speech to infants prior to the emergence of their first meaningful words. Baran et al. (1977) speculated that differences might be expected to emerge once children begin to produce meaningful words. By following infants at 10/11, 18 and 24 months we were able to address this speculation. Although the differences between VOT clarification measures at these ages were not significant, the trend of less clarification to infants at the youngest age and greater clarification to infants as they got older supports the idea that mothers may tailor their input based on their children's stage of language development.

Closed Class VOT Characteristics

The present study also investigated the VOT characteristics of closed class words in IDS and ADS. We hypothesized that VOT clarification of these words would be greatest in speech to the children at the oldest age, 24 months, because this is when more closed class words would be emerging in their own expressive vocabularies. We found no significant differences between overall VOT duration of closed class words at any age. Additionally, there were no significant differences in VOT clarification measures between the different ages. Visual inspection of the means (see Figures 9-11 for reference) show that although non-significant, the clarification of closed class words in IDS was actually greatest at 18 months. Again, unexpectedly but similarly to the open class words, VOT clarification of closed class words was the greatest in the ADS condition.

Examination of TTR revealed that there were significant differences in the proportion of new words to infants at different ages. In IDS, TTR was the greatest at 10/11 months and the least at 24 months. There were no significant correlations between TTR and the measures of VOT clarification at any of the ages. However, it is interesting that the age at which TTR was the greatest was also the age at which VOT clarification was the least. As with the open class group, it is likely that there is another factor influencing both TTR and VOT clarification at this age.

Previous studies have not all performed separate analyses of open- and closed-class words. This is problematic, because open class words tend to be stressed whereas closed class words are almost always unstressed. Given that stress is a factor known to increase VOT, combining these words would likely lead to greater variability.

Furthermore, open- and closed-class words should be separated because they represent different types of input to the child. While open class words consist mainly of content words and vowels, and make up the majority of the input the child receives, closed class words contain the important function words that help hold our sentences together, and are highly prevalent in ADS. It is important to know which words mothers are clarifying in their speech to infants, and when this clarification is occurring.

Malsheen performed post-hoc analyses of the word-initial alveolar stops in closed-class words for one mother in her study. She found that the mean VOT value for voiceless stop the IDS condition was significantly longer than for the ADS condition. Thus, this mother clarified both open- and closed-class words to her infant at the one word stage of development. These findings are not directly supported by the current study, given that we found no difference in overall VOT duration for closed class words between the conditions. However, it is difficult to make direct comparisons given that Malsheen's findings were drawn from only one mother and she looked only at the alveolar cognate pair, whereas this study investigated the velar cognate pair for closed class items of 10 mothers. Because the present study compared the infants longitudinally at different ages, and did not sort the children by language stages, it is also difficult to make generalizations about what stage of language development the mothers used the most clarification. It is possible that more significant group differences would have emerged if the infants had been sorted by language stage rather than age.

VOT Clarification and Language Outcomes

The main question that this study sought to answer was whether or not VOT clarification of open or closed class words was related to infant language outcomes at 2 years. We hypothesized that greater VOT clarification would be positively correlated with outcome scores on the PPVT, EOWVT and the MCDI. However, results showed that VOT clarification of open class words was not significantly correlated to language outcomes at 2 years. Furthermore, the relationships between clarification and the language outcome scores were not uniform across the different measures, making them difficult to interpret. At 10/11 months, language outcome scores increased as DOVC decreased, suggesting that the more the mothers clarified their VOT, the better their children's language outcomes. However, this relationship was not significant and there was a high probability of Type I error. The other two measures of VOT clarification at 10/11 months had the opposite relationship with language outcome scores at 2 years. As VOT clarification (as judged by $d_{(a)}$ and percent overlap) increased, language outcome scores decreased. At 18 and 24 months, greater clarification was associated with lower language outcome scores for all measures except percent overlap, which was negatively related to the raw PPVT and MCDI scores. Again, these relationships were not significant and should be interpreted with caution.

In order to gain a better understanding of the relationship between VOT clarification of open class words and language outcomes, we split the mothers into two groups at each age and compared their children's scores. The mothers were divided into a "More VOT Clarification Group" and a "Less VOT Clarification Group". Each of the measures of VOT clarification was taken into consideration when assigning the mothers to a group. We found that the children of the More VOT Clarification Group at each age had a higher mean on the MCDI than the Less VOT Clarification Group, but this difference was not significant.

For closed class words, our hypothesis was partially supported. At 18 months, VOT clarification (as measured by percent overlap) was significantly correlated with MCDI raw scores ($r(8) = -.687$, p=.028). Although there were strong correlations between the other measures of VOT clarification at this age and MCDI raw scores ($d_{(a)}$ $r(8) = .604$, p=.068; DOVC $r(8) = -.598$, p=.065), these correlations did not reach the critical value of .632 to meet significance. This a power issue due to the small sample size. Similarly, there were strong negative correlations between percent overlap and the other two language outcomes, with greater clarification relating to higher scores (PPVT $r(8) = -.593$, p=.071; EOWVT $r(8) = -.593$, p=.077) but again they did not reach the high critical value of significance. The same relationship was seen at 10/11 and 18 months, but not always at 24 months. At 24 months, greater $d_{(a)}$ (more clarification) was with lower scores on the EOWVT, and greater percent overlap (less clarification) was associated with higher scores on the PPVT and EOWVT. These relationships were not significant and had very high probabilities of type I error. Still, it is interesting to note that the age at which we initially hypothesized VOT clarification of closed class words would be the greatest was the one age at which the relationship between VOT clarification and language outcomes was actually negative. It is not surprising that we found the only significant correlation at 18 months. This was the age at which descriptive statistics showed us the VOT of closed class words tended to be the most clarified.

Just like with the open class group, we split the mothers into two groups based on their VOT clarification of closed class words. Interestingly, the only significant difference on language outcomes was found between the groups at 10/11 months on the MCDI score. The "More VOT Clarification" group received significantly higher MCDI scores than the "Less VOT Clarification group". This is interesting because our descriptive statistics indicated that VOT clarification was actually the least for all measures when the infants were 10/11 months old. One possible explanation is that although overall mothers clarified their VOT of closed class words more when their infants were 18 months old, the mothers who began to provide this clarification earlier provided a unique advantage to their infants. This was also the age at which we found TTR in IDS was the greatest, suggesting that the mothers were using a greater variety of words and less repetition. It could be that with less repetition, VOT clarification took on a greater role in helping the infants discriminate between the voicing categories in connected speech. This may have contributed to higher receptive vocabulary at 24 months. However, it could also be that the mothers' perception of their children's receptive vocabulary influenced their clarification of closed class words. We cannot infer causality in either case, but only look at the relationship.

The findings from this part of the present study bring up an important issue regarding IDS clarification and syntax. It is well known that IDS is characterized by shorter utterances (Fernald et al., 1989), leading to fewer closed class words. These closed class words do not emerge in a child's own expressive language until the child begins combining words and expanding their MLU. Given this background information, it would make sense that mothers may selectively clarify their IDS, providing greater clarification of content words early on, perhaps upon the emergence of the first word as suggested by Malsheen, and clarification of function words when the children are more linguistically advanced. Indeed, this was the pattern found for vowel modification in IDS. Bernstein Ratner (1984a) found that the vowels of function words here highly centralized in speech to pre-verbal listeners, whereas the vowels of content words were more widely dispersed. However, in speech to children at the holophrastic stage, vowel space was similarly expanded for both content and function words, and in speech to children with MLUs of 2-3.5, the vowel space for function words actually larger than for content words. Thus, as children became more linguistically advanced, the mothers began to clarify their speech in a non-selective fashion. Bernstein Ratner (1984a) suggested that

> "…while content word clarification could facilitate learning of lexical items by the pre-verbal and holophrastic stage child, it would need to be supplemented by increasingly precise articulation of function words as the children progressed to the stage of developing grammars" (p. 577)

This is similar to what we found in the current study. Although the differences were not significant, the general trend was that VOT clarification of closed class words was the least at 10/11 months when the infants were likely preceding or just beginning the one word stage, and greatest at 18 months.

In order to further investigate this, we compared the VOT clarification of open versus closed class words in the nine mothers who met the criteria for both of these word classes. Interestingly, we found that in general, mothers clarified open class words more than closed class words at all ages (see Figures 13-15 for reference). However, even though open class words received greater VOT clarification, this clarification did not significantly relate to language outcomes, whereas clarification of closed class words did relate to language outcomes. One possible explanation is that although mothers clarify open class words more, this clarification is not as necessary because the infants are receiving other cues to help them identify these words. For example, when the mother says, “The baby is eating her peas” she is also holding or pointing to the baby. Closed class words do not have this physical reference, and thus VOT clarification may play a more important role.

VOT Clarification Methodology

In the present study, we used three methods of calculating VOT clarification. Two of the methods, $d_{(a)}$ and percent overlap, were chosen in order to make comparisons to previous studies. Unfortunately, these measures were not significantly related at all ages. This brings up issues for deciding which measure to use in future studies. The measures of DOVC and $d_{(a)}$ appeared to be the most strongly and consistently related, with significant correlations at all ages for both open- and closed-class words. Percent overlap, on the other hand, was not significantly correlated with DOVC of closed class words at 24 months or with DOVC of open class words to adults, and it was not significantly correlated with $d_{(a)}$ for open or closed words to the adult, or open class words to the 10/11 and 24 month old infants. This is worrisome, because the significant findings were only present for the measure of percent overlap. Percent overlap was defined as the percentage of voiceless tokens that fell within plus or minus 2 standard deviations of the mean for voiced tokens. The problem with this measure is that if there is no overlap, it does not tell you how far apart the distributions are. Thus, two mothers could both have a percent overlap of 0, but one mother could have significantly higher VOT duration for voiceless stops than the other mother. For this reason, it is suggested that one of the other methods of VOT clarification be used for future studies because they may be more sensitive to differences.

Limitations

When interpreting the findings from this study, it is important to note that because of the matching process, the number of tokens analyzed in the open class group made up only 32.1% of the total number of tokens produced by the mothers that began with an initial singleton bilabial stop. For the closed class group, the total number of tokens analyzed made up only 30.7% of the total number of closed class tokens produced by the mothers that began with an initial singleton velar stop. This matching process was necessary, because it ensured that the tokens from each group were comparable in the many factors that may affect VOT beside addressee. However, it also limited the proportion of tokens that were included in the final analyses. For many of the mothers, only six tokens were analyzed for each phoneme. This is a small sample from which to make comparisons across groups. It is possible that the differences between groups may have been more pronounced if a greater number of tokens had been analyzed. This also limited the total number of participants included in the final analyses, which resulted in a power issue for some of the statistics.

The matching process was also limiting because it looked only at tokens that were inherently more similar between the two groups. For example, tokens occurred in single words in IDS with much greater frequency than they did in ADS, but only the words that were matched in both conditions for this context were included. So not only did the matching process limit the number of tokens included, but it also provided somewhat of an artificial representation of IDS. Thus, the greatest limitation may be not the limited proportion of tokens included, but the fact that we do not know what the majority of the input the children received looked like. This is an important factor for future IDS researchers to consider when designing their studies.

Another limitation of this study, again due in part to the matching process, was that we looked at only one place of articulation for each of the word classes. Thus, when we made the comparisons between the open- and closed-class word groups we were unable to say whether these differences were because of word class or because of place of articulation. It is also possible that if we had been able to include all places of articulation, we may have found different patterns of clarification.

Finally, this study is limited in that we attempted to compare IDS at different ages, rather than different stages of language development. Although we made assumptions as to what stage of language development the infants were at based on their age, there may have been individual variability around each of these ages. It is possible that some of the infants may have already produced their first words at 10/11 months, and there may have been variation in MLU at the older ages.

Future Directions

There appears to be an interesting relationship between VOT clarification of closed class words and infant language outcomes. However, this relationship is not conclusive based on the findings of this study. Future studies should investigate this relationship based on children’s language abilities rather than chronological age. This could be accomplished though a longitudinal design in which infants are followed from 10 months until 4 years of age. However, rather than comparing infants based on their age, they should be grouped based on receptive/expressive vocabulary and MLU. Future studies should also try to make comparisons of VOT between all of the voiced stops rather than just one cognate pair in order to look for individual trends.

Future studies should also include a greater number of participants in order to increase the power of the statistics. In order to obtain a greater number of tokens, it is recommended that rather than matching tokens at each age, all tokens should be included. This will provide a more accurate depiction of the input the infants are receiving. Tokens can then be described according to the factors known to influence VOT, such as place of articulation, and these can be included as factors to be considered in the statistical analyses. Furthermore, to gather more similar tokens across the play sessions, it is recommended that fewer choices of toys be provided. In the present study it is suspected that that large variety of toys available at each play session led to fewer similarities in the words the mothers used across the visits.

Another possibility for future research would be to investigate the effect of VOT clarification of words with an experimental design rather than a naturalistic design. It would be interesting to see if infants are able to learn new vocabulary in speech with greater VOT clarification. This could be tested using a preferential looking task in which new words are presented in IDS with maximally distinct VOT or in IDS with overlapping VOT values. Using an experimental design would also allow us to make causal inferences. Findings from these future studies could be used to instruct mothers how to best facilitate word learning.

References

Allen, J.S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America, 113*(1), 544-552.

Baran, J. A., Zlatin Laufer, M., & Daniloff, R. (1977). Phonological contrastivity in conversation: A comparative study of voice onset time. *Journal of Phonetics*, *5*, 339-350.

Bernstein Ratner, N. (1984a). Patterns of vowel modification in mother-child speech. *Journal of Child Language*, *11*(3), 557-578.

Bernstein Ratner, N. (1984b). Phonological rule usage in mother-child speech. *Journal of Phonetics*, *12*(3), 245-254.

Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother-child speech. *Journal of Phonetics*, *14*, 303-309.

Bernstein Ratner, N. (1996). From "signal to syntax": But what is the nature of the signal? In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 135-150). Mahwah, NJ: Lawrence Erlbaum Associates.

Bernstein Ratner, N., & Luberoff, A. (1984). Cues to post-vocalic voicing in mother–child speech. *Journal of Phonetics*, *12*, 285-289.

Berry, J., & Moyle, M. (2011). Covariation among vowel height effects on acoustic measures. *The Journal of the Acoustical Society of America*, *130*(5), EL365-EL371.

Bharadwaj, S. V., & Graves, A. G. (2008). Efficacy of the Discreteness of Voicing Category (DOVC) measure for characterizing voicing errors in children with cochlear implants: A Report. *Journal of Speech, Language & Hearing Research*, *51*(3), 629-635.

Boersma, P., & Weenink, D. (n.d.). *Praat*. Amsterdam, The Netherlands: University of Amsterdam. Retrieved from www.praat.org

Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science, 296*(5572), 1435.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*(5), 1584-1595.

Cooper, R. P., & Aslin, R. N. (1994). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development, 65*, 1663-1677.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171*(968), 303-306.

Englund, K. (2005). Voice onset time in infant directed speech over the first six months. *First Language*, *25*(2), 219-234.

Fais, L., Kajikawa, S., Amano, S., & Werker, J. F. (2010). Now you hear it, now you don't: Vowel devoicing in Japanese infant-directed speech. *Journal of Child Language, 37*(2), 319-340.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, *8*(2), 181-195.

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, *27*(2), 209-221.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*(1), 104-113.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(3), 477.

Fowler, C.A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listener's perception and use of the distinction. *Journal of Memory and Language, 26,* 489-504.

Hayashi, A., Tamekawa, Y., & Kiritani, S. (2001). Developmental change in auditory preference for speech stimuli in Japanese infants. *Journal of Speech, Language, and Hearing Research, 44*(6), 1189-1200.

Inoue, T., Nakagawa, R., Kondouu, M., Koga, T., & Shinohara, K. (2011). Discrimination between mothers' infant- and adult-directed speech using hidden Markov models. *Neuroscience Research, 70*, 62-70.

Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, *16*(1), 55-68.

Krause, J. C., & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America*, *115*(1), 362.

Kubaska, C.A. (1982). *A longitudinal study of mothers' speech characteristics.* (Unpublished Doctoral Thesis). Brown University, Providence, RI.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech, 6,* 172-175.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in

English stops. *Language and Speech*, *10*(1), 1-28.

Lisker, L. (1975). Is it VOT or a first-formant transition detector? *The Journal of the Acoustical Society of America*, *57*(6 Pt 2), 1547-1551.

Liu, H., Kuhl, P. K., & Tsao, F. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, *6*(3), F1-F10.

Liu, H., Tsao, F., & Kuhl, P. K. (2009). Age-related changes in acoustic modifications of Mandarin maternal speech to preverbal infants and five-year-old children: a longitudinal study. *Journal of Child Language*, *36*(4), 909-922.

Lowenstein, J. H., & Nittrouer, S. (2008). Patterns of acquisition of native voice onset time in English-learning children. *Journal of the Acoustic Society of America, 124*(2), 1180-1191.

Ma, W., Golinkoff, R. M., Houston, D. M., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development*, *7*(3), 185-201.

MacWhinney, B. (n.d.). *CLAN*. Pittsburgh, Pennsylvania: Carnegie Mellon University. Retrieved from childes.psy.cmu.edu/clan/

Malsheen, B. (1980). Two hypotheses for phonetic clarification in the speech of mothers to children. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child phonology, Volume 2* (pp. 173-174). New York: Academic Press.

Mandel, D., Kemler Nelson, D. G., & Jusczyk, P. W. (1996). Infants remember the order of words in a spoken sentence. *Cognitive Development*, *11*(2), 181-196.

Maye, J., Weiss, D. J., & Aslin, R. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science*, *11*(1), 122-134.

Maye, J., Werker, J., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101-111.

McRoberts, G. W., & Best, C. T. (1997). Accommodation in mean f0 during mother–infant and father–infant vocal interactions: a longitudinal case study. *Journal of Child Language*, *24*(3), 719-736.

Messer, D. J. (1981). The identification of names in maternal speech to infants. *Journal of Psycholinguistic Research*, *10*(1), 69-77.

Mines, M. A., Hanson, B. F., & Shoup, J. E. (1978). Frequency of occurrence of phonemes in conversational English. *Language and Speech, 21*(3), 221-241.

Moslin, B. J. (1979). *The Role of Phonetic Input in the Child's Acquisition of the Voiced-Voiceless Contrast in English Stops: A Voice Onset Time Analysis* (Unpublished Doctoral Dissertation). Brown University, Providence, RI.

Nearey, T. M., & Rochet, B. L. (1994). Effects of Place of Articulation and Vowel Context on VOT Production and Perception for French and English Stops. *Journal of the International Phonetic Association*, *24*(1), 1-18.

Newman, R., Bernstein Ratner, N., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversation speech signal predicts later language development: A retrospective analysis. *Developmental Psychology, 42*(4), 643-655.

Newman, R.S., & Hussain, I. (2006). Changes in preference for infant-directed speech in low and moderate noise by 4.5 to 13-month-olds. *Infancy, 10*(1), 61-76.

Singh, L., Nestor, S., Parikh, C., & Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy, 14*(6), 654-666.

Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy, 13*(4), 410-420.

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review, 27*, 501-532.

Song, J. Y., Demuth, K., & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *Journal of the Acoustical Society of America, 128*(1), 389-400.

Stevens, K. N., & Klatt, D. H. (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *The Journal of the Acoustical Society of America*, *55*(3), 653-659.

Sundberg, U. (2001). Consonant specification in infant-directed speech. *Working Papers*, *49*, 148-151.

Sundberg, U., & Lacerda, F. (1999). Voice onset time in speech to infants and adults. *Phonetica, 56*, 186-199.

Synnestvedt, A. (2009). *Voice onset time in infant-directed speech at two ages.* (Master's Thesis). Retrieved from http://drum.lib.umd.edu//handle/1903/10834

Theodore, R. M., Miller, J. L., & DeSteno, D. (2009). Individual talker differences in voice-onset-time: Contextual influences. *Journal of the Acoustical Society of America, 125*(6), 3974-3982.

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy*, *7*(1), 53-71.

Uther, M., Knoll, M., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, *49*(1), 2-7.

Utman, J. (1998). Effects of local speaking rate context on the perception of voice-onset time in initial stop consonants. *The Journal of the Acoustical Society of America*, *103*(3), 1640-1653.

Weppelman, T. L., Bostow, A., Schiffer, R., Elbert-Perez, E., & Newman, R. S. (2003). Children's use of the prosodic characteristics of infant-directed speech. *Language & Communication, 23*, 63-80.

Werker, J., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, *103*(1), 147-162.

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology*, *43*(2), 230-246**.**

Zlatin, M. A. (1974). Voicing contrast: Perceptual and productive voice onset time characteristics of adults. *Journal of the Acoustical Society of America*, *56*(3), 981-994.

Zlatin, M., & Koenigsknecht, R. (1976). Development of the voicing contrast: a comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research*, *19*(1), 93-111.