말소리와 음성과학 제2권 제1호 (2010)
pp. 41~50

Acoustic Variation Conditioned by Prosody in English Motherese

Choi, Hansook[1)]

ABSTRACT

The current study exploresacoustic variation induced by prosodic contexts in different speech styles,with a focus on motherese or child-directed speech (CDS). The patterns of variation in the acoustic expression of voicing contrast in English stops, and the role of prosodic factors in governing such variation are investigated in CDS. Prosody-induced acoustic strengthening reported from adult-directed speech (ADS)is examined in the speech data directed to infants at the one-word stage. The target consonants are collected from Utterance-initial and -medial positions, with or without focal accent. Overall, CDS shows that the prosodic prominence of constituents under focal accent conditions variesin the acoustic correlates of the stop laryngeal contrasts. The initial position is not found with enhanced acoustic values in the current study, which is similar to the finding from ADS (Choi, 2006 Cole et al, 2007). Individualized statistical results, however, indicate that the effect of accent on acoustic measures is not very robust, compared to the effect of accent in ADS. Enhanced distinctiveness under focal accent is observed from the limited subjects' acoustic measures in CDS. The results indicate dissimilar strategies to mark prosodic structures in different speech styles as well as the consistent prosodic effect across speech styles. The stylistic variation is discussed in relation to the listener under linguistic development in CDS.

Keywords: Speech style, acoustic variation, prosodic prominences, child-directed speech, English stop contrast

1. INTRODUCTION

The current research is on prosodically induced acoustic variation in motherese, or child-directed speech (CDS), which is extensively discussed from adult-directed speech (ADS). Speech variation due to listener's language ability is explored in terms of the phonetic variation in the caregiver's speech addressed to the infant acquiring language. Phonetic variation is conditioned by several factors such as intrinsic segmental characteristics, speech style or speaking rate, phonological contrast, and prosodic structure. This study is intended to see the interaction of two factors, prosody and speech style. Specifically, the acoustic variation under prosodic prominences is examined in a unique speech style, CDS.

1) Haskins Laboratorie,chohaskinyale.ed, Contact Author

Received: February 1, 2010.
Revision received: February 28, 2010.
Accepted: March 3, 2010.

Speaking rate and style have been discussed as key factors to induce speech variation. Faster speaking rates are generally assumed to cause more coarticulated or overlapped gestures and thus, resulted acoustic cues are less extreme or clarified (Byrd & Tan, 1996; Gay, 1978; Munhall & Lofqvist, 1992; Zsiga, 1994 for example). Speech style is also identified as an independent factor to induce acoustic variation. Moon & Lindblom (1994), for example, shows that there is a clear difference in “undershooting” patterns of vowels in dissimilar speaking styles under constant speaking rates and stress patterns. The authors suggest that the “clear" samples were louder and involved in a systematic, undershoot-compensating reorganization of the acoustic patterns as well.

Prosody is another important feature to condition the variation in segmental representation which has been discussed mainly from ADS. Specifically, prosodic prominence affects articulatory and acoustic signals of speech. It is generally confirmed from diverse studies that prosodic prominence conditions marked articulation and acoustic expressions (Beckman & Edward, 1990; de Jong et al, 1993; Jun, 1993, etc.). Further, different types of prosodic prominence have been shown to condition different patterns of acoustic and articulatory representations of segments (Cho, 2001; Cho & Keating, 2009; Choi, 2006; Cole et al., 2007; White, 2002).

Acoustic studies in adult-to-adult speech, specifically, showed that prosodically salient contexts, such as phrase boundaries or focal and pitch accents, condition enhanced distinctiveness between contrastive stops and vowels (Cho & Keating, 2009; Choi, 2006; Cole et al, 2007; Hsu & Jun, 1998). However, dissimilar contexts of prosodic salience are found to induce different acoustic strengthening effects. Specifically, the effect of focal/pitch accent on acoustic measures is generally more global and consistent with a greater distinction between the contrastive segments compared to the localized and inconsistent effect of prosodic domain boundaries (Choi, 2006; Cole et al, 2007)[2]. Prosody is not simply superimposed on segments with independent suprasegmental features, but different prosodic structures are reflected even at the segmental representation. Acoustic correlates for segmental contrast specify the prosodic structures as well. With multiple cues from different representation levels, processing of prosody will get further aided.

This prosody-induced acoustic variation has been reported in different languages. However, the manner in which prosodic prominence and prosodic phrase structure are marked at the level of segmental variation is found to be language-specific. There is an interaction between the size of the phonological inventory and the corresponding acoustic variation. A simple correlation between the size of contrast and the degree of variation (Manuel, 1990; 1999), nonetheless, cannot explain the patterns in different phonological systems. More complicated effects from the segmental contrast system are reported in different languages (Choi, 2006; 2009).

The current research question is on the extent of prosodic effects on segmental representation in different speech styles. The previous findings on prosody-induced acoustic or articulatory variation were discussed exclusively in adult-to-adult speech or adult read speech. The effects from different speech styles are not discussed yet. In particular, the pattern in the speech directed to a child is of interest because of the unique patterns reported in CDS. CDS is generally agreed to feature exaggerated prosodic realizations, such as exaggerated pitch peaks and extensive stress marking (Fernald et al., 1989; Fernald & Mazzie, 1991, for example), and more clarified consonant and vowel production (Kuhl et al., 1997; Malsheen, 1980; Ratner, 1984, for example). The uniqueness in CDS predicts probable deviation from ADS in the acoustic variation induced by prosodic structures. The deviations are predicted in multiple ways as follows.

The prosodic bootstrapping hypothesis suggests even more enhanced prosodic variation in CDS (Gleitman & Wanner, 1982; Jusczyk et al., 1992; Oberecker & Friederici, 2006 for example) to signal segmentation and phrasing. According to the findings from ADS and the prosodic bootstrapping hypothesis, the distinctiveness in prosodic structure is predicted to be still more enhanced at suprasegmental representation and at the segmental representation to provide optimal acoustic cues for processing. That is, segmental realization in focused contexts may be even more distinctive, which will be contrasted with rather reduced segmental distinction conditioned by phrasal boundaries or nonprominent contexts. Strengthening and weakening of segments conditioned by prosody will be further enhanced to maximize the prosodic information. This possibility is not always supported, however, given that CDS is marked with enhanced distinctiveness in the segmental representation.

Consonants and vowels in CDS are found optimally clarified and faithful. The extreme manifestation of segments may nullify any variation in the target tokens, including the variation due to prosodic contexts such as focal accent and domain boundaries. In addition, the reported exaggerated nonsegmental cues such as pitch and pause may provide enough prosodic information, and thus, the dependence on segmental variation for prosodic information can be reduced. Overshooting of prosodic information in nonsegmental features may lead to undershooting in segmental information. Another possibility in production of CDS is, therefore, that segments are expressed very distinctively in all contexts with little variation and do not reflect the macro-prosodic structure. The prosodic structure may be signaled more by nonsegmental acoustic cues than by segmental variation in this prediction.

The competition between prosodic conditioning and segmental faithfulness is expected greatest in the speech directed to single-word-staged infants. Motherese or CDS is a very listener-sensitive speech style. If caregivers take account of their children's linguistic development in their CDS, the infants at different linguistic stages will be addressed with rather varied speech patterns of CDS. This hypothesis is confirmed in a longitudinal study by Roy et al. (2009), where the authors reported a

2) Cho & Keating (2009) reports on acoustic strengthening of Cs and Vs in the U-initial with greater VOTs and greater vocal amplitude and argues against the localized effect of the domain boundaries. Still, the effects of boundaries and accents are found under different strengthening effects.

significant correlation between the child's Mean Length of Utterance (MLU) trajectory and the MLU trajectory directed to the child. The reported correlation is between the infant's expressive linguistic ability and the speech input provided by caregivers. In other words, the caregivers consider what the infants say and address infants in a way costumed to their production level. If so, the infant's single word production at the initial verbal stage might arouse greater caregivers' attention on word presentation and the segmental representation of individual lexical items. The expectation on segmental clarification is expected highest at this stage, and then, variation in acoustic realization is predicted minimized, including prosodically induced acoustic variations. The speech is still expected to provide good prosodic information to aid more effective speech processing particularly for early language acquirers. It is expected that the prosodic information that helps phrasing and segmentation is always maximized in CDS. The question here is whether prosody induces acoustic variation in segmental realization under this special context. It has not been discussed whether and how the macro-structure of prosodic information is depicted in the micro-structure of segmental representation under great attention on segments, or simply, how the caregivers maximize segmental and/or prosodic information in acoustic realization of segments.

The current study explores the acoustic expression of the English stop laryngeal contrast in motherese directed to one-word producing infants. The primary goal is to observe how prosody conditions variation in acoustic realization of the English stops in CDS. To observe stylistic variations due to the unique features in CDS, the acoustic patterns are discussed in comparison to the findings from the ADS style. The research questions mentioned above are tested and discussed with the following method.

2. METHOD

2.1 Speech Material

English laryngeal contrast is observed from a consistent place of articulation, bilabial. Target consonants of /b/ and /p/ were analyzed from the initial CV syllable of six English words (with the target syllables underlined: <u>po</u>tter, <u>bo</u>ttle, <u>pe</u>ter, <u>bea</u>ter, <u>pe</u>tter, <u>be</u>ttor)[3]. All the target syllables bear the lexical stress in the present study, and thus there is no variation of prosodic prominence at the lexical level. The words were collected from different prosodic contexts, utterance-initial and -medial positions of minimally structured speech addressed to infants. There was no prepared control on the sentence or focus. Target words were given as names for toys, and mothers had a playtime with their infants using the target toys. The conversation between the mother and the infant was recorded as a sound corpus. From the sound data, the target CV sequences were extracted, a total of 1055 tokens, and identified for the prosodic condition individually. The boundaries of utterance were identified based on the obvious big pause and the context of the speech. Focus was identified based on the content of the speech as well as the acoustic patterns, such as pitch variation and loudness. The focus can be contrastive narrow focus and a broad focus (Frascarelli, 2000; Ladd, 1996) with extra intensity and pitch-accent as well. Both types of focus are counted focus in the current study. Nonfocused tokens were accompanied with other obviously focused tokens within the same utterance.

2.2 Subjects

Speech data were collected from 5 mothers who have 12 to 18 months old infants in the current study. The five participants' target infants (E1, E2, E3, E4, and E5) were at the very initial period in the MLU stage 1 and were producing very limited one-word utterances mostly. The mothers are all born in US and are native speakers of American English who are staying in the New England area, US. Both of the parents are native speakers in American English and use English to their infants as the major language of communication. The subjects were agreed to participate in child language studies at Haskins laboratories previously, and were contacted later to conduct the current study. No hearing or speaking impairments were reported from both the mothers and their children. In the acoustic analyses, however, one mother (E4)'s speech tokens were excluded due to the inconsistent settings and failure in interacting with the infants consistently. All of the subjects were paid for their participation.

2.3 Procedure

The participating mother interacted with her infant and shared conversation with the infant in a sound attenuated room at Haskins Laboratories, Yale, New Haven. The room was filled with toys, books, and a very comfortable couch in the middle so that the participating mother and child might hardly feel that they were under an experiment.

The Target CV sequences were prepared as names of toys, and each toy was guided with a flash card containing its details. For

3) Current work reports only the variation in consonants. However, the patterns in vowels are also observed and will be reported soon.

example, 'peter', a name of a stuffed animal, was printed in a bold letter and attached to the toy. In addition, flash cards were provided with information such as Peter's age, its favorite food and toys, best friends, and its family. Participating mothers were instructed to play with the children using the toys and the cards. They were asked to behave as usual as possible, and no explanation of the research goal and design was explained before the experiment. In order not to distract the infants, each toy was introduced separately as a new toy, and the mothers named the toy and described its characteristics using the prepared lists of descriptions. After finishing one, a new toy was introduced replacing the old one. No sentences were prepared to be cited by the mothers in the introduction. The participating mothers used their own sentences in their own styles and in their comfortable rates to give information about the toys.

At the end of the interaction, the mothers were asked to give lists of questions to the children. However, mothers, themselves, had to give answers to those questions or react to very few answers from their infants. This session was prepared to observe rather obvious patterns of focal accent variation. For example, for a question asking "*is Peter a bear?*", the answer, "*No, Peter is a cat*", provides a token of 'Peter' without the contrastive focal accent. On the other hand, in the question and answer of "*is Peter a bear? No, Bettor is a bear*", 'peter' is a token with a contrastive focal accent.

Each mother and child pair was observed separately at different time slots and thus, there was no interaction between participating mothers at all. Optional long breaks were provided upon the subjects' request. The interaction between mother and infant involved unrestricted short breaks frequently, and recording continued through the breaks.

The minimally structured speech of caregivers to their children was recorded using a wireless microphone and a TASCAM HD-P2 recorder at the sampling rate of 44100Hz. The recorded sounds were transferred to a PC and analyzed using Praat at the sampling rate of 44100Hz for the acoustic analysis and the SPSS for statistical results. In total, 1055 CV sequences were collected from 4 speakers (E1: 92 tokens, E2: 129 tokens, E3: 430 tokens, E5: 404 tokens) as viable tokens after getting rid of invalid tokens such as noisy or devoiced tokens. 375 tokens (35.5%) out of 1055 were from initial position, and 795 (75.4%) were with the focal accent.

2.4 Measurements

The recorded sounds were transferred to a PC at a sampling rate of 44100Hz, and 1055 tokens of designated CV sequences were extracted and analyzed with the Praat program (Version 5.0.09, Boersma & Weenink, 2000). Acoustic measurements of VOT and F0 at the onset of the following vowel were employed as acoustic correlates of stop voicing contrast. VOT is widely accepted as a feature to mark a laryngeal contrast of stop consonants in most languages including English (Lisker & Abramson, 1964), such that English voiceless stops show greater VOT values than their voiced counterparts. F0 at the following vowel onset is also reported as a cue to English stop voicing. Whalen et al. (1993) reports from their perception study that F0 at the vowel onset assists voiced versus voiceless distinction of the preceding stops even with unambiguous VOT values.

The duration from the stop release to the onset of the second formant in the following vowel was measured as VOT. F0 values at the onset of the following vowel were estimated from the mean period over the first three periodic cycles of the wave forms after the stop release through visual inspection of the waveform and the vertical striations corresponding to glottal pulses in the spectrogram. The autocorrelation pitch analysis function in Praat using a time step of 0.005s was additionally employed to get the values for comparison. The automatically measured values showed highly similar values but reported some missing or misleading values that could be analyzed with the manual measurement[4].

2.5 Statistical Analysis

The influence of each prosodic factor was evaluated based on Linear Mixed models and Analyses of Variance (ANOVAs) of the Univariate General Linear Model (GLM). Analyses of Linear Mixed models were selected to detect overall significant effects across less balanced sets of tokens based on the flexibility of the model (Hoffman & Rovine, 2007). The tokens in this study were selected from unstructured conversational settings without intended repetition and thus, the number of tokens is not well-balanced as in repeated read speech experiments. Current analyses of Linear Mixed models selected speakers as random factors, and Position (i.e., Initial, Medial), Focus (i.e., Focused, Nonfocused), and Cs (Voiceless, Voiced) as fixed factors. The individual factor effects and their interactions were tested on the dependent variables of VOTs and F0s at the onset of the following vowels.

---

4) The same method of measurement is used in the current study that was used in Choi (2006, 2009) and Cole et al. (2007) to analyze the stop contrast in males and females' speech directed to adults.

Univariate GLM analyses were additionally conducted to observe more detailed effects within individual subjects. The Linear Mixed model analyses report the overall effects of the factors while the individual patterns are not clearly reported. The unit of the statistical analysis was limited to acoustic measures of one subject in order to prevent possible idiosyncratic effects from a specific subject when all subjects' data are pooled. The Univariate ANOVAs were performed for each acoustic measure as a dependent variable individually, and the obtained results were compared across speakers. The same factors, Position, Focus, and Cs, were employed for the ANOVA analysis in the present setting as independent variables. The SPSS statistical package (SPSS for windows, Standard Version, Release 11.0.1, 15 Nov 2001, SPSS Inc.) was used for the statistical measurement.

In addition to statistical comparisons, distributions of individual subjects' tokens in each condition were compared using diverse charts and graphs.

3. RESULTS AND DISCUSSION

The effect of prosody is reported in English CDS as significant according to the statistical analysis. Linear Mixed model analyses report significant effects of Focal Accent (F(1,1047)=12.994, p<.0001) and the interaction of Focal Accent*Cs (F(1,1047) = 19.039, p<.0001) in the distribution of VOTs, and a significant effect of Focal Accent (F(1,1047)=24.181, p<.0001) in the distribution of F0s. A marginally significant interaction of Focal Accent*Cs is detected in F0 (F(1,1047) = 4.254, p<.05). However, the positional effects or any interactive effects with Position are not found significant both in VOTs and F0s (VOT: F(1,1047)=2.167, p>.05; F0: F(1,1047)=0.03, p>.05). This result is similar to the patterns in ADS. Previous studies on prosodically conditioned acoustic variation in English stops, observed in ADS reported significant effects of prosodic prominences, mainly from focal accents. An asymmetry is reported between focal effects and boundary effect in that the focal effects are more consistent and obvious the effects due to domain boundaries. In general, no consistent strengthening is found in the utterance initial Cs but not always with enhanced acoustic values. However, more divided comparisons of the tokens within individual report deviation of CDS from ADS with a lot reduced prosodic effects in CDS.

Results from 3-way ANOVAs (Position ◇ Focus ◇ Cs) for each subject's VOTs and F0s are provided in <Table 1>. The F-values and P-values show that Focal Accent is a significant factor for certain English mothers' VOT and F0 measurements (E2's F0, E3's VOTs and F0s, E5's VOTs). It shows a contrast with the findings on ADS in that the focal effects are found very consistent and highly salient across speakers in ADS (Choi, 2006; 2009; Cole et al., 2007). E3 and E5 also show an interactive effect of Focus and Cs in VOTs (E3: F(1,429)=24.088, p<.0001; E5: F(1,403)=12.110, p<.005) by marking greater VOTs for focused voiceless tokens without obvious increase in voiced ones. On the contrary, Position is never reported with a significant effect in the observed acoustic correlates in the current study. A marginal interaction of Position*Cs is detected in E2 (F(1,128)=4.728, p<.05) such that the medial voiceless tokens are marked with greater VOTs. U-initial in the current study is not reported with significant strengthening (Cho & Keating, 2009).

Table 1. Statistical results by 3-way ANOVA on VOT and F0 of English Motherese (*** p<.0001, **p<.01,* p<.05)

	E1	E2	E3	E5
VOT				
Focal Accent	**F(1, 91) =.2.375**	**F(1, 128) = 0.017**	**F(1,429)**=16.436***	F(1, 403) = 7.666**
Position	**F(1, 91) =1.357**	**F(1, 128) = 0.235**	**F(1,429) =1.455**	**F(1, 403)** =0.000
F0 at the onset				
Focal Accent	**F(1, 91) =2.407**	**F(1, 128)** = 5.855*	**F(1,429)** = 21.829***	F(1, 403) = 1.647
Position	**F(1,91) =0.727**	**F(1, 128) =0.714**	**F(2,429) = 1.259**	**F(1,403) = 2.132**

Statistical results on the interaction of the factors, Focal Accent and Consonant types, suggest a paradigmatic influence of accent. It means that the voiced and voiceless stops show different acoustic variation in the focused contexts. The detailed acoustic variations are discussed with the distributional charts for the mean VOTs as an acoustic correlate of English stop voicing and 95% confidence intervals as is depicted in <Figure 1> for focal variation and <Figure 2> for positional variation.

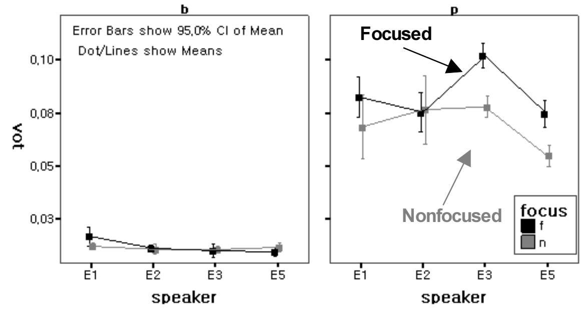

Figure 1. Effects of Focal Accent on VOT values of English voiced and voiceless stops in motherese.[5]

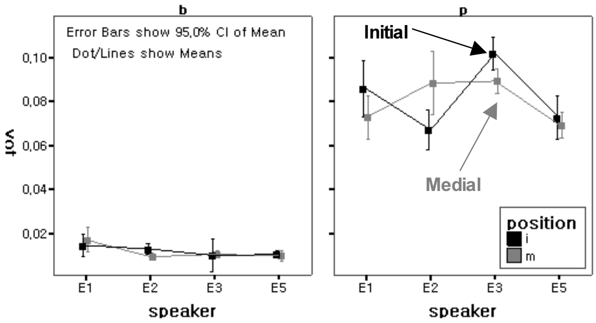

Figure 2. Positional variation in VOT values of English voiced and voiceless stops in motherese.

The acoustic distribution of motherese in <Figure 1> reveals an increase in the focused token, which is counted as strengthening.[6] The increase is, however, not uniform in all the subjects, but in a bit inconsistent way. E3 and E5 show a significant increase of the focused tokens, particularly in voiceless tokens. The paradigmatic change under focus is revealed as a clear increase of VOTs in voiceless tokens with no remarkable changes in voiced tokens. Any systematic patterns conditioned by positional variation are, on the other hand, not found in the distributional view of VOTs in <Figure 2>. The initial and medial values involve certain changes in the voiceless tokens, but in a very dissimilar way. The utterance initial position, a prosodically salient context, is not specified with a robust change of acoustic measures compared to the patterns in the less salient medial contexts. Acoustic variation in CDS reveals the effect of prosody mainly from the focal condition, which is similar to the prosodic effects in ADS.

With the overall similarity of prosodic effects in the two different speech styles, the degree of consistent effects is still suggested as a discrepancy between CDS and ADS. In <Figure 1> for CDS, the increase under focus is obvious only for voiceless tokens of certain speakers, whereas the differences under two focal contexts are not clearly marked in voiceless tokens. This focal effect in ADS is, however, reported with a consistent increase for voiced and voiceless tokens across speakers as in <Figure 3> for example.

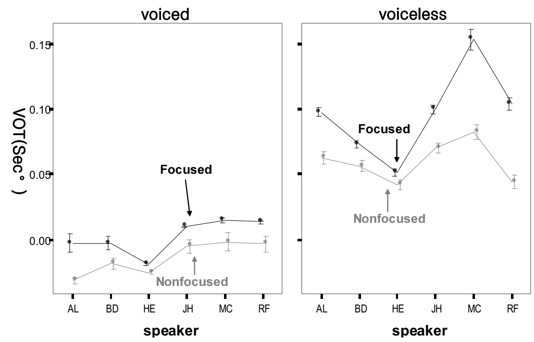

Figure 3. Variation by Focal Accent on VOT values of English voiced and voiceless stops in ADS.

<Figure 3> shows results from an adult-to-adult speech using same prosodic contexts and target words (Choi, 2006), which is reintroduced here for a comparison. VOTs in <Figure 3> show an increase under focus. VOT values are greater for all of the contrastive stops across speakers, and at the same time, the increase is even greater only for the voiceless tokens. In other words, the focal accent induces syntagmatic and paradigmatic effects on the segmental variations in ADS very consistently. Not only voiceless tokens but also voiced tokens depict prosodic variation in ADS.

It is noteworthy that the focal accent-induced variation in CDS in <Figure 1> does not replicate a syntagmatic increase for all the focused stop tokens. Focus in CDS is revealed with a paradigmatic increase only in VOTs of voiceless stop tokens.

The distribution of F0 at the onset of following vowels in <Figure 4> is marked with enhanced values for the focused tokens. The variance of the distribution is, nevertheless, large and thus significant separation between focused and nonfocused is not reported from the statistics except for E2 and E3. The increased

5) In <Figure 1, 2, & 3>, boxes depict means, and each error bar corresponds to 95% confidence interval of mean, which are marked for individual speakers'voiced (left chart) and voiceless (right chart) tokens. The darker boxes correspond to the Mean value of focused ones in both voiced and voiceless stops in <Figure 1 & 3>, whereas the lighter ones demarcate the counterparts without focal accent. The darker one in <Figure 2> is for the initial condition.

6) Keating (1984) and Pierrehumbert & Talkin (1992) provide evidence for the enhancement at the onset consonant of the accented syllable such that the onset consonants generally involve greater VOT under accent. Greater VOT is considered as a more obstruent-like pattern which can be an acoustic strengthening of consonantal features under accent. Syntagmatic changes involve whole Cs as a group whereas paradigmatic changes treat individual contrastive Cs in different ways (Cho & Keating, 2001; Cho & Jun, 2000; Hsu & Jun, 1998).

F0 is not very surprising, given that the focused tokens are marked with higher F0 and intensity. The vowels with the marked pitch accent and greatest intensity in the utterance may result in higher F0 at the onset. The syntagmatic increase of F0 under focus, therefore, seems a by-product of global prosody, and is hardly identified as a segmental increase of the stop contrast itself. Positional variation does not show any clear variation in F0s.

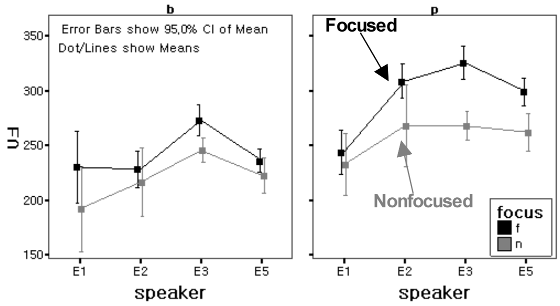

Figure 4. Focal variation in F0 values at the voicing onset following voiced and voiceless stops in motherese.

Overall, statistical results and distributions of individual acoustic measures in motherese show a similar but reduced effect of prosody in segmental representation compared to the prosody-induced acoustic variation in ADS. Prosodic effects are found under focus, but the effects are not as significant as in CDS. Finally, visualized comparison of the contrastive stop tokens also confirms the reported patterns, which is depicted in the 2-dimensional acoustic space in terms of VOT and F0 in <Figure 5> and <Figure 6>. The mean values for the tokens in the given contexts are depicted in <Figure 5> and <Figure 6>, and the distance between matching symbols, namely 'B' and 'P' in the same shade, is identified as distinctiveness between the contrastive segments. For example, the darker B and P are mean voiced and voiceless stops under focus, while lighter B and P are mean voiced and voiceless stops without focus in <Figure 5>. The distance between the darker B and P is counted as distinctiveness between the contrastive stops under focal accent. Greater distance indicating increased distinctiveness for the focused tokens in darker symbols is clearly detected from E3 and E5 in <Figure 5>. E1 and E2 do not provide a salient increase under focus. The focused tokens reveal enhanced distinctiveness in some subjects.

Comparatively, the symbols depicting initial and medial values in <Figure 6> present very inconsistent changes due to positional difference. The changes are either not very obvious as in E5 or without increased distance as in E1 and E2. A slightly increased distance is detected only in E3's initial contrast, which was not statistically significant. It indicates the discrepancy between focal effects and positional effects, the first of which is more consistent and further distinctive. Dominance of the focal effect is not very different from the patterns of prosodic variation in adult-directed speech (Choi, 2006; Cole et al., 2007). However, the focal effects in CDS are found even reduced, compared to the ones in ADS. The contrast marking in ADS shows significant increase under focus for most of the speakers.

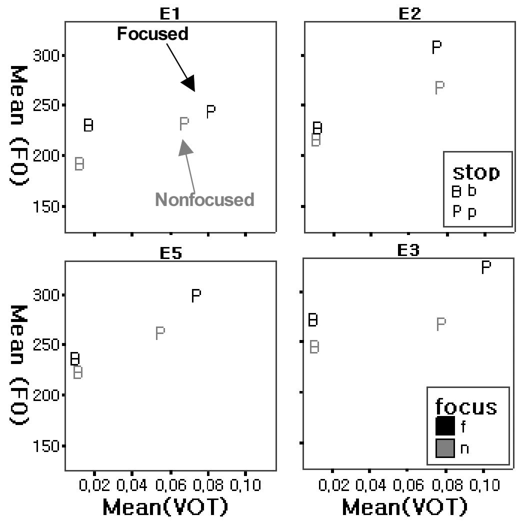

Figure 5. Variation in stop laryngeal contrast in CDS as a function of Focus in the 2-dimensional acoustic space by VOT and F0.[7]

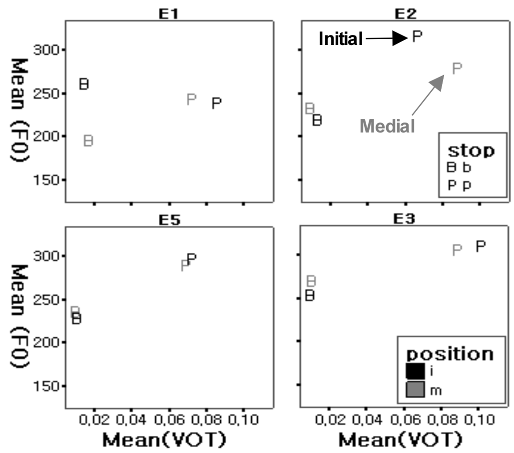

Figure 6. Variation in stop laryngeal contrast in CDS as a function of Position in the 2-dimensional acoustic space by VOT and F0.

7) The voicing contrast in stops in <Figure 5, 6, & 7> is depicted in the 2-dimensional acoustic space of VOT as X-axis and F0 as Y-axis. The mean VOT and mean F0 of contrastive stop tokens are marked for the given condition within each speaker. 'P'indicates voiceless tokens and 'B' indicates voiced counterparts. In <Figure 5 & 7>, darker ones are for focused condition and lighter ones are for nonfocused condition, whereas darker ones are for initial condition and lighter ones are for medial condition in <Figure 6>.

For example, <Figure 7>, reproduced from Choi (2006), displays an example of the acoustic variation in adult-to-adult speech. The focused contrast with darker symbols in <Figure 7> involves greater distinctiveness than the nonfocused one with lighter symbols, which is depicted as the length of lines between symbols. The darker lines are always longer than the lighter lines for all the provided speakers, and the increase entails both axes of VOTs and F0s.

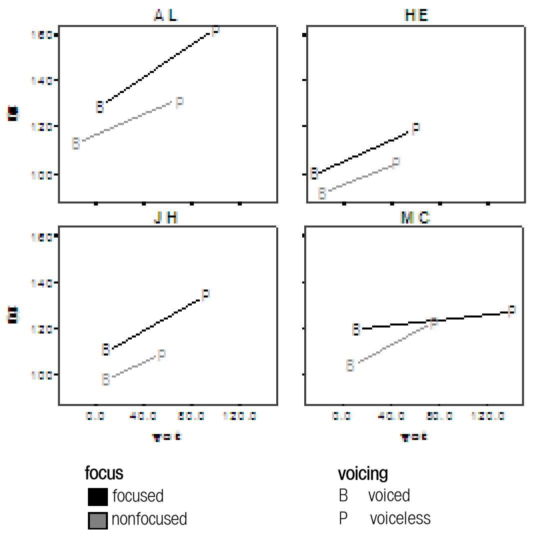

Figure 7. Variation in stop laryngeal contrast in ADS as a function of Focus in the 2-dimensional acoustic space by VOT and F0.

To sum up the results from diverse analyses of the measurements, the acoustic expression of laryngeal contrast in stops is found to be influenced by the prosodic contexts in the current study on the speech directed to one-word staged infants. The tokens under focal accent are marked with enhanced acoustic cues for the voicing contrast mainly in the voiceless tokens. The focal accent in CDS is found to induce paradigmatic acoustic strengthening effects in stops providing further distinctive acoustic cues. The difference in the prosodic contexts, focal accent and domain boundaries, is expressed in the acoustic variation in the observed motherese such that the focal accent induces the paradigmatic changes whereas the different positions within a domain do not involve clear acoustic variation. The current study does not show any acoustic strengthening in the utterance initial position.

The results also suggest that the unique speech style of CDS seems to condition the effects of prosody on segmental representation. With all the similarities between CDS and ADS, the acoustic variation conditioned by prosodic structures is found comparatively limited in the current study on the CDS to one-word producing children. Significant prosodic effects are reported only from a few subjects and in a less significance level, which shows deviation from previous findings on the very consistent and distinctive prosodic effects in ADS. The syntagmatic effects in ADS, the strengthening of stops, are not replicated in CDS either. The overall prosody-induced variation in acoustic correlates for English stops is found reduced in CDS. In other words, the acoustic correlates for stops in CDS seem to depict less systematic variation under different prosodic contexts. The prosodic information signaled by acoustic variation of stops is not enhanced but curtailed in CDS[8].

4. CONCLUSION

The primitive goal to observe the prosodic variation in different speech styles is achieved here by reporting on the prosodically induced acoustic variation in CDS, which shows similarity to the patterns in ADS. The prosodic structures are demarcated at the segmental representation in different speech styles in similar ways. The similarity between CDS and ADS suggest that caregivers' speech in the current study provides information of the utterance structures or macro-prosody though the infants are not producing phrases or sentences. Children at the one-word stage are found to process multi-word sentences (Hirsh-Pasek & Golinkoff, 1996), though they produce single-word utterances mostly. Young infants are exposed to the speech input in a rather uniform pattern from early stages of linguistic development, and are assisted with multiple acoustic signals for the micro and macro structures of the speech.

Stylistic variation is still found. Overall similarities are found between two speech styles of ADS and CDS, but details of the effects are not identical. Results in this study show that the prosodic effects in CDS are not as significant as in ADS. The syntagmatic enhancement of contrastive segments under accent

8) Another possible reason of the weakened prosodic effects in Motherese may be found in the experimental settings, ADS on read speech versus CDS on the spontaneous speech. Cutler (1994) introduces that spontaneous speech tends to be produced at a slower rate, and to have longer and more frequent prosodic demarcation and hesitations than read speech. The features of spontaneous speech are shared in the CDS according to Cutler (1994). The features are understood as enhanced prosodic features in CDS, and thus, it is not reasonable to predict any reduction of prosodic effects in spontaneous speech. In addition, Cole et al. (2007) reports a significant and consistent effect of pitch accent in BostonRadio-news sound corpus, which includes a recoding of rather spontaneous speech.

and phrase boundary in ADS is not replicated in the acoustic variations of the observed CDS. The strengthening due to prosodic prominences is marked only in some of the speakers. Overall, prosodic effects in CDS are found reduced in the acoustic correlates for segmental contrast. The diminished prosodic effects on segmental variation in CDS indicate less attention on macro prosodic structure, and comparatively greater attention on the expression of the micro structure of segments at the representation level of segmental acoustic correlates. The reduced prosodic effects at segmental representation can be understood a functional solution in CDS. The greater expectation on segmental clarity is achieved at the cost of reduced prosodic signals that can be possibly aided with enough acoustic cues from nonsegmental features. Current results do not suggest that the overall macro-prosodic information is sacrificed. It indicates that the prosodic information in the segmental representation is not enhanced. The nullified variations signaling prosody in the representation of acoustic correlates of segmental contrast will be correlated with the generally reported exaggeration in nonsegmental cues such as intonation, accent, and pause. It is very presumable that other cues for prosodic information involve strengthening, which should be further analyzed in the future study.

The revealed acoustic variation indicates a dissimilar strategy to signal the prosodic cues in addressing single-word-producing infants. At least, the segmental cues are not employed for the prosodic information as much as they are in ADS. One possible explanation on the given results is the presumable caregivers' sensitivity to the child's expressive linguistic ability. Based on the correlation reported by Roy et al. (2009), the reduced prosodic effects in the current study can be understood as a result of the presumable attention on the words and segmental realization in the speech to word-producing infants. Greater attention on micro segmental representation may override the concern on rather macro-structures of prosody and syntax at least at the segmental representation. An interaction between language acquiring children and the external structure of CDS is suggested from the current research but not directly measured. Therefore, subsequent research is introduced finally to confirm the interaction.

If caregivers pay attention to children's linguistic stages in addressing their children, the related prediction is, then, the speech directed to further syntactically developed children will convey more information of phrasal structures possibly even in segmental representations. If so, prosodic factors will be found with greater significance in the speech directed to children with greater MLU. This hypothesis will be tested with a planned longitudinal study in the near future. Acoustic enhancement in prosodically prominent contexts in CDS is expected to vary depending on the child's developmental stages to facilitate the optimal structure to develop their language skills. This functional strategy in CDS is understood to provide immature language acquirers with maximum cues to develop their language skills. Children at different linguistic stages will be provided with linguistic inputs tailored to their emergent needs. Another related question is whether and how much the child get influenced by the provided speech. The bidirectional interaction between the child's linguistic development and CDS should be explored further, and the interactive patterns need to be compared cross-linguistically as well.

REFERENCES

Beckman, M. & Edward, J. (1990). "Lengthening and shortening and the nature of prosodic consistency", *Papers in Laboratory Phonology I: Between Grammar and the Physic of Speech*, pp. 152-178.

Boersma, P. & Weenink, D. (2000). *The Praat program*, Version 4.0.13, University of Amsterdam.

Byrd, D., & Tan, C. Co (1996). "Saying consonant clusters quickly", *Journal of Phonetics*, Vol. 24, pp. 263-282.

Cho, T. (2001). "Effects of Prosody on Articulation in English", PhD dissertation, UCLA.

Cho, T. & Jun, S. (2000). "Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean", *Chicago Linguistics Society*, Vol. 36, No. 1, pp. 31-44.

Cho, T. & Keating, P. (2001). "Articulatory and acoustic studies of domain-initial strengthening in Korean", *Journal of Phonetics*, Vol. 29, pp. 155-190.

Cho, T. & Keating, P. (2009). "Effects of initial position versus prominence in English", *Journal of Phonetics,* Vol. 37, pp. 466-485

Choi, H. (2009). "The role of contrast in prosodically induced acoustic variation", *Phonetics and Speech Sciences*, Vol. 1, No. 3, pp. 29-37.

Choi, H. (2006). "Prosodically induced acoustic variation in English, Korean, and Hindi", ProQuest / UMI, ISBN: 0542228998.

Cole, J., H. Kim, H. Choi, & M. Hasegawa-Johnson. (2007). "Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech", *Journal of*

Phonetics, Vol. 35, pp. 180-209.

Cutler, A. (1994). "Segmentation problems, rhythmic solutions", *Lingua,* Vol. 92, pp. 81-104.

de Jong, K., Beckman, M. & Edwards, J. (1993). "The interplay between prosodic structure and coarticulation", *Language and Speech*, Vol. 36, pp. 197-212.

Fernald, A. & Mazzie, C. (1991). "Prosody and focus in speech to infants and adults", *Developmental Psychology*, Vol. 27, No. 2, pp. 209-221.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., Boysson-Bardies, B., & Fukui, I. (1989). "A cross-language study of prosodic modifications in mother's and father's speech to preverbal infants", *Journal of Child Language*, Vol. 16, pp. 477-501.

Frascarelli, M. (2000). "*The syntax-phonology interface in focus and topic construction in Italian*", Dordrecht: Kluwer.

Gay, T. (1978). "Effect of speaking rate on vowel formant movements", *Journal of the Acoustical Society of America*, Vol. 63, pp. 223-230.

Gleitman, L.R. & Wanner, E. (1982). "Language acquisition: The state of the art", In Wanner & Gleitman (Eds.), *Language acquisition: The state of the art*, Cambridge, pp. 3-48.

Hirsh-Pasek K. & Golinkoff, R. M. (1996). *The origins of grammar: Evidence from early language comprehension*, MIT Press.

Hoffman, L. & Rovine, J. (2007). "Multilevel models for the experimental psychologist: Foundation and illustrative examples", *Behavior Research Method*, Vol. 39, pp. 101-117.

Hsu, C. & Jun, S. (1998). "Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic?", *UCLA Working Papers in Phonetics,* Vol. 96, pp. 69-89.

Jun, S. (1993). "The phonetics and phonology of Korean prosody", PhD dissertation. Ohio State University.

Jusczyk, Hirsh-Pasek, Nelson, Kennedy, Woodward, & Piwoz (1992). "Perception of acoustic correlates of major phrasal units by young infants", *Cognitive Psychology*, Vol. 24, pp. 252-293.

Keating, P. (1984). "Phonetic and phonological representation of stop consonant voicing", *Language*, Vol. 60, pp. 286-319.

Kuhl, P., Andruski, J., Chistovich, I., Chistovish L., Kozhevnikova, E., Ryskina, V., Stolyarovw E., Sundberg, U. & Lacerda, F. (1997). "Cross-language analysis of phonetic units in language addressed to infants", *Science,* Vol. 277, pp. 684‑686.

Ladd, R. (1996). "Intonational phonology", *Cambridge Studies in Linguistics,* Vol. 79, Cambridge: Cambridge University Press.

Lisker, L. & Abramson, A. (1964). "A cross-linguistic study of voicing in initial stops: acoustic measurements", *Word*, Vol. 20, pp. 384-422.

Malsheen, B. (1980). "Two hypotheses for phonetic clarification in the speech of mothers to children", In G. Yeni-Komshian, G., Kavanagh J., & Ferguson, C. (Eds.), *Child Phonology: Perception*, New York: Academic Press, pp. 173-184.

Manuel, S. (1990). "The role of contrast in limiting vowel-to-vowel coarticulation in different languages", *Journal of the Acoustical Society of America*, Vol. 88, pp. 1286-1298.

Manuel, S. (1999). "Cross-language studies: relating language-particular coarticulation patterns to other language-particular facts", *Coarticulation: Theory, data and techniques*, pp. 179-198.

Moon, S. & Lindblom, B. (1994). "Interaction between duration, context, and speaking style in English stressed vowels", *Journal of the Acoustical Society of America,* Vol. 96, pp. 40-55.

Munhall, K., & Lofqvist, A. (1992). "Gestural aggregation in speech: Laryngeal gestures", *Journal of Phonetics*, Vol. 20, pp. 111-126.

Oberecker, R. & Friederici, A. (2006). "Syntactic event-related potential components in 24-month-olds' sentence comprehension", *NeuroReport,* Vol. 17, pp. 1017-1021.

Pierrehumbert, J. & Talkin, D. (1992). "Lenition of /h/ and glottal stop", *Papers in Laboratory phonology II: Gesture, Segment, Prosody*, pp. 90-117.

Ratner, B. (1984). "Patterns of vowel modification in mother-child speech", *Journal of Child Language*, Vol. 11, pp. 557-578.

Roy, B., Frank, M., & Roy, D. (2009). "Exploring word learning in a high-density longitudinal corpus", *http://www.media.mit.edu/cogmac/publications/roy-frank-roy-cogsci2009.pdf*

Whalen, D., Abramson, A., Lisker, L. & Mody, M. (1993). "F0 gives voicing information even with unambiguous voice onset time", *Journal of the Acoustical Society of America*, Vol. 93, pp. 2152-2159.

White, L. (2002). "English Speech Timing: A Domain and Locus Approach", PhD dissertation, The University of Edinburgh.

Zsiga, E. C. (1994). "Acoustic evidence for gestural overlap in consonant sequences", *Journal of Phonetics*, Vol. 22, pp. 121-140.

• **Choi, Hansook**
Research Affiliate (2009~present)
Haskins Laboratories
New Haven, CT USA
Email: choi@haskins.yale.edu