Additive effects of lengthening on the utterance-final word in child-directed speech

Eon-Suk Ko

University at Buffalo, State University of New York

Melanie Soderstrom

University of Manitoba

Acknowledgements:
We thank the theatre students and mothers who participated in our study including the pilot stage. In addition, we are grateful to Lindsay Bacala and Dana Bernier for their assistance with data collection, and Alex Walters, Mengyang Qiu, and Sandra Hunt for their assistance with data analysis. The project was partially funded through an NSERC Discovery grant to the second author.

Corresponding Author's contact information:
Eon-Suk Ko
122 Cary Hall
Communicative Disorders and Sciences
University at Buffalo, State University of New York, Buffalo, New York 14214

Email: eonsukko@buffalo.edu
Phone: 716-829-5572
Fax: 716-829-3979

Abstract

Purpose: We investigate lengthening effects in child-directed speech (CDS) across the sentence, testing the additive effects on duration of Word Position, Register, Focus, and Sentence Mode (statement/question).

Method: Five theatre students produced 6 sentences containing 5 monosyllabic words in a simulated dialogue, varying in Register, Focus, and Sentence Mode. A total of 1,800 sentences were segmented using forced-alignment tools, and the duration of each word was analyzed.

Results: The results show significant effects of Register, Word Position and their interactions. The simple effect of Register was significant in all 5 word positions, indicating a global elongation effect in CDS. Interestingly, there was no proportional increase of the final word in CDS. Additionally, the three-way interactions Register x Word Position x Focus and Register x Word Position x Sentence Mode were significant, which converge to the conclusion that the utterance-final word in CDS is additively elongated when it is focused, and in a statement.

Conclusion: Elongation in CDS is a global effect, but the additive effects of duration demonstrated in our data suggest that the effect of enhanced utterance-final lengthening in CDS in naturalistic samples may be a by-product of discourse characteristics of CDS.

Key words: child-directed speech, duration, focus, utterance-final lengthening, speaking rate, prosodic bootstrapping

Additive effects of lengthening on the utterance-final word[1] in child-directed speech

It is well-established that child-directed speech (CDS) has a slowed tempo, or elongation of utterance duration, compared to the speech addressed to adults (ADS; adult-directed speech) (e.g. Garnica 1977; Lam & Kitamura, 2010; Morgan 1986). Previous research has suggested that elongation effects in CDS are more pronounced at, or limited to the final syllable of the sentence. Bernstein Ratner (1986), for example, found that utterance-final lengthening is more exaggerated in CDS compared to ADS when the differences between the duration in utterance-medial and utterance-final syllables in each register are considered. More recently, Church, Bernhardt, Pichora-Fuller, and Shi (2005) suggested that if final syllables are removed from the analysis, durational differences between ADS and CDS disappear altogether. However, a number of underlying subtleties in the temporal characteristics of CDS could have been masked in previous research due to the difficulty of controlling for various factors influencing duration in naturalistic speech. These subtleties may have important implications for understanding the nature of temporal patterns in CDS. In the current study, we investigate how the duration of words across sentences are affected by register, focus (Cooper, Eady, & Mueller, 1985) and sentence mode (Eady & Cooper, 1986), using a highly controlled set of elicited speech stimuli.

Understanding how durational effects are exhibited in CDS is theoretically important

[1] In our study, all words are monosyllabic. Thus, the term 'word' in this article relates specifically to a monosyllabic word. We refrain from using the term 'syllable' because our data does not directly test if the elongation effects we investigate in this article more generally apply to a syllable in a different environment, e.g. unstressed final syllable in a multisyllabic utterance-final word. Related findings in previous research (e.g. Oller, 1973), however, point to the possibility that the effects may indeed apply to the final-syllable regardless of the stress or the number of syllables in the final word.

because the utterance-final lengthening in CDS has been hypothesized to provide young children with important information about the boundaries between syntactically relevant units of speech (e.g. Hirsh-Pasek et al., 1987; Morgan, Meier, & Newport, 1987; Seidl, 2007; Soderstrom, Kemler Nelson, & Jusczyk, 2005). Implicit in this argument is the view that the ***purpose*** of exaggerated utterance-final lengthening in CDS may be to provide enhanced prosodic cues to language structure. Recently, however, it has been convincingly argued that at least some of the prosodic characteristics of CDS that are thought to be helpful for infants' language learning are side effects of the affective function in CDS (e.g. Singh, Morgan, & Best, 2002; Trainor, Austin & Desjardins, 2000), which may serve to hold infants' attention to language and its structure. These findings raise the possibility that exaggerated utterance-final lengthening in CDS could also be a by-product of other linguistic patterns characteristic of CDS.

One possible explanation for the exaggerated phrase-final lengthening in CDS could be found in the discourse characteristics of CDS. Previous research found that mothers tend to place focused words in utterance-final position more frequently when addressing infants than adults (Fernald & Mazzie, 1991; Woodward & Aslin, 1990). There are therefore multiple factors driving lengthening effects on these words (CDS, utterance-final lengthening, focus effects). Analyses of the effects of multiple factors on duration in the same syllable have found that the combined lengthening effects approximate the addition of each individual factor independently (Klatt, 1976), subject to some constraints on expandability (Cooper, Eady, & Mueller, 1985). Thus, an utterance-final word under focus is predicted to undergo greater elongation than an utterance-final word without a focus due to the combination of the two factors affecting duration. A conflation of these two factors with the elongation effect of register is expected to yield a much greater elongation of the utterance-final word in CDS compared to ADS. Such additive

effects could be further shaped by the particular sentence mode (e.g. statement/question) of a given utterance.

A crucial factor to be considered in a study of durational aspects of speech is the trade-off between the length of utterances and speaking rate. That is, longer sentences tend to be spoken at a faster rate due to the compression of syllable duration known as "anticipatory shortening" (Nooteboom, 1972; Lindblom & Rapp, 1973). This effect is particularly clear in shorter utterances containing 7 words or less (Yuan, Liberman, & Cieri, 2006). Although these analyses were done with ADS, the findings are likely applicable across registers. At the very least, they suggest that models of durational effects in CDS should take sentence length under consideration, especially given its characteristic of short utterance length. Results of previous studies investigating durational effects in CSD may not have accounted adequately for this source of variance (e.g. Church et al., 2005). In addition, the extent of vowel lengthening differs in the final, the penultimate, and the anti-penultimate syllable (Yuan, Liberman, & Cieri, 2006), thus a gross categorization of all non-peripheral vowels as "phrase-medial" as in earlier research (e.g. Bernstein Ratner, 1986) could mask certain durational effects in CDS.

In the current study, we investigate the patterns of elongation in CDS using a highly controlled set of stimuli. Given that there are a variety of factors that influence duration of a given word such as affect (Trainor et al., 2000) and lexical probability (Jurafsky, Bell, Gregory, Raymond, 2001), an ideal investigation of temporal adjustments in CDS requires comparing the temporal pattern of CDS with ADS utterances containing identical strings of words produced by the same speakers in the same discourse contexts. This, however, is impractical with naturalistic samples. While effects of sentence length and sentence mode might be adequately controlled for under more naturalistic conditions using a careful selection process, there are practical

difficulties with determining the focus context in any given utterance, since our perceptual judgments of focus are conflated with the acoustic features in question. The current study thus investigated the elongation effects of CDS on duration based on a set of simulated data using trained speakers. Although there are limitations to this approach, previous research suggests that simulated CDS reflects the properties observed in spontaneous CDS at least to a certain extent (Fernald & Simon, 1984; Schaeffler, Kempe, & Biersack., 2006). In addition, Knoll, Scharrer, and Costall (2009) suggest that simulated speech generated by trained speakers is much more likely to be generalizable to real world than data generated by untrained speakers.

Our specific aims in this study are as follows. First, we investigate the elongation effect of CDS in each word position of a sentence to examine the claim that the final position has a special status regarding the CDS elongation effect. In particular, we re-evaluate the claim that the speaking rates in ADS and CDS are approximately equal after the elimination of the phrase-final syllable (Church et al., 2005). Second, we investigate the effects of register, word position, focus, and sentence mode on duration and their interactions to investigate the additive durational effects of these factors.

Methods

We recruited five trained speakers (theatre students) in the Winnipeg area.[2] The participants were presented with a photo of a young child about 14 months old projected onto the

[2] During piloting, we initially attempted recording the data with mothers in the presence of their child, but it was clear that the mothers had considerable difficulty attending to the register manipulations due to our complex design. The child's presence did not seem to help in eliciting CDS, presumably because the mothers were not actually interacting with their child. One of the mothers' data were analyzed along with the theatre students. In addition, data from four mothers were included in the perceptual rating task reported in this section. However, since the extent to

computer monitor along with the stimulus. The stimuli consisted of 6 sentences composed of 5 words. The lexical content of the sentences were devised to be appropriate in both the CDS and ADS registers, albeit not the most typical CDS or ADS sentences. A list of the sentences used in this experiment is presented in (1).

(1) a. Lee held the long worm.

b. Ray made a blue bow.

c. May gave Kim the doll.

d. Moe ran by the lawn.

e. Ann drew a blue moon.

f. Neil threw the new ball.

Each sentence was read in intonation indicating a Sentence Mode of statement or question, with systematically varied location of Focus, i.e. initial-focus, neutral, and final-focus, in CDS or ADS Register. This yielded 6 utterance types (2 Sentence Modes x 3 Focus Conditions) for each of the sentence items in (1a)-(1f). To elicit each target sentence with the appropriate Sentence Mode and Focus conditions, 72 pairs (6 sentence items x 6 utterance types x 2 Registers) of prompt-response dialog were composed. Examples are illustrated in Table 1.

((Table 1 about here))

---

which their register manipulations were successful is questionable, we do not include their data in this report. Overall their data patterned similarly to that of our theatre student participants.

Participants were seated in front of a computer screen with a slide display showing the prompt and response, with the focused item (if any) highlighted. The procedure began with a demonstration of 3-6 pre-recorded prompt-response examples, followed by a practice session of 5-7 practice sentences for each of CDS and ADS. During the recording session, the experimenter read one of the prompt sentences (either the CDS or ADS version), and the participant responded with the appropriate test sentence response. Sentences were divided into 4 blocks of 9 sentences in each speech register and presented in pseudo-randomized orders that differed for each participant. Participants alternated whether CDS or ADS blocks were presented first. The whole set of sentences was repeated 5 times. The recording was done using an Audio-technica Cardiod condenser microphone, and was digitized at the sampling rate of 22,050 Hz with a 16 bit resolution. A total of 360 sentences were elicited from each participant, which resulted in a final set of 1,800 sentences containing 9,000 word tokens.

The performance of the five participants in producing the simulated ADS and CDS was evaluated by the second author and four psychology students through a rating task. Using Praat (Boersma & Weenink 2010), a selection of 12 sentences (1 sentence item x 6 utterance types x 2 Registers) from each of the 5 participants' speech was presented to the 5 raters 5 times, with a self-controlled duration of break after every 50 stimuli. The selected sentence item differed for each speaker. The order of the stimuli was randomized individually for each rater. The raters were given the following instruction: "On a scale of 1-5, please indicate how likely the sentence was spoken to a young child." The mean rating score in ADS target tokens (2.2, SD=1.4) was different from that for the CDS target tokens (4.4, SD=1.1), which was significant by a two-tailed paired t-test, $t(749) = -34.5$, $p < 0.001$. This result indicates that the CDS target tokens

were rated as significantly more likely to have been said to a child than the tokens produced as ADS.

The size of our data set is substantially larger than the ones in similar previous studies, for example, 90 sentences in Swanson, Leonard, & Gandour (1992), and 413 sentences in Church et al. (2005). To facilitate the segmentation procedure, and also to apply the criteria in exactly the same manner for all the tokens in the data set without involving the risk of annotator bias, we adopted the Penn Phonetics Lab Forced Aligner tool kit (Yuan & Liberman, 2008) to time-align the target sentences with the transcripts at the word level. The duration of each word was extracted based on the time-aligned word boundaries using a Praat script. Occasional small pauses between words as detected by the forced aligner were not included in the measurement. To check the accuracy of the word boundaries generated by the forced alignment, a trained research assistant manually coded the word boundaries in a total of 228 sentences containing 1,140 words, which cover about 13% of the entire data. There was a significant correlation held between the word durations in the raw and adjusted data sets in both speech registers ($r=0.97$, $n=1140$, $p < 0.001$). The median and the interquartile range for the absolute differences in word duration between the raw and corrected annotations was -1.17 ms (Q1: -15.04 ms, Q3: 13.58 ms). Considering that the average duration of each word was 292 ms, this indicates a high level of agreement. The results thus assure that we can obtain valid statistical analysis of the duration data based on the boundaries marked by the forced aligner.

Results

We analyzed the data using a series of repeated measures ANOVAs with the word duration and ratio as the dependent measures, and the participants (F_1) and the items (F_2) as the random variable. Fixed factors were a combination of Register (CDS and ADS – 2 levels), Word

Position (5 levels), Focus (focused and unfocused – 2 levels), and Sentence Mode (statements and questions – 2 levels), depending on the specific question being addressed. Below, we will first report the statistical analyses of a model containing two fixed factors of Register and Word Position, followed by a more complex model containing an additional factor of either Focus or Sentence Mode.

Effects of Register and Word Position

We first sought to replicate the observation of exaggerated utterance-final lengthening in CDS in previous research by constructing a repeated measures ANOVA model with Register and Word Position as fixed factors. The results show significant effects of Word Position, $F_1(4,16) = 303.6$, $p < 0.001$; $F_2(4,20) = 15.2$, $p < 0.001$, Register, $F_1(1,4) = 16.6$, $p < 0.05$; $F_2(1,5) = 327$, $p < 0.001$, and the interaction between the two, $F_1(4,16) = 21.4$, $p < 0.001$; $F_2(4,20) = 8.2$, $p < 0.001$. The main effect of Word Position indicates that the duration of words differ as a function of their position in the sentence, and the main effect of Register indicates that the duration of words differ in the two registers. Their interaction suggests that the degree of register-induced lengthening differs depending on the position of the word. An inspection of this interaction revealed that the final word, which had a greater duration than any other word in the sentence, underwent the greatest amount of register-induced lengthening in CDS (see Figure 1; row 1, left, and Table 2). Our result here is therefore consistent with previous research suggesting greater utterance-final lengthening in CDS. By comparison, a parallel analysis with ratio as the dependent variable found only a main effect of Word Position, $F_1(4,16) = 437.9$, $p < 0.001$; $F_2(4,20) = 231.9$, $p < 0.001$, indicating that words differ systematically in ratio depending on position in the sentence. The lack of Register main effect or its interaction with Word Position

suggests that a highly consistent ratio is maintained in each word position across the two registers (Figure 1; row 1, right), which implies that CDS lengthening occurs *across* the sentence.

((Figure 1 about here))

We next examined directly whether CDS elongation occurs only in final position or more generally throughout the sentence. In separate by-participant and by-item repeated measures ANOVAs, we found significant effects of the CDS register for each word position (all p's < 0.05 or smaller). The one exception was the by-participant analysis in Word 3, $F_1(1,4) = 5.2$, $p = 0.08$. The significant simple effect of Register in all 5 words means that the elongation effect of CDS register is distributed over the entire sentence rather than on the final syllable only. Consequently, the slower speaking rate in CDS (3.03 syllables/s) compared with ADS (3.89 syllables/s) persisted after the elimination of the final syllable in calculating the speaking rate (CDS: 3.61 syllables/s; ADS: 4.58 syllables/s). Our finding thus contradicts the claim made by Church et al. (2005) that the slow speaking rate in CDS is mainly due to the exaggerated lengthening of the utterance-final syllables.

((Table 2 about here))

To summarize these findings, CDS elongation was found globally when raw durations are considered, but there was a greater durational increase in CDS on the final syllable. In contrast, there was no comparative increase in utterance-final lengthening in CDS when ratio is considered. Therefore, enhanced utterance-final lengthening in CDS in this analysis is

attributable to the additive combination of final-syllable lengthening (Klatt, 1973) and overall lengthening across the sentence in CDS.

Additive effect of Focus

We next examined how Focus might influence the utterance-final lengthening in CDS. If Focus, which elongates the duration of words, has an additive effect on the interaction Word Position x Register we found earlier, the pattern of interaction Word Position x Register should differ depending on the focus condition of the sentence. In other words, we predict the elongation effect of CDS would be greater in Word 1 under the Initial-focus condition whereas the same effect would be greater in Word 5 under the Final-focus condition. To test this hypothesis, we constructed a model which included Focus as a fixed factor, in addition to Word Position and Register. In our results, all main effects and interactions were significant (p's < 0.05 or smaller), including the three way interaction Focus x Word Position x Register, $F_1(8,32) = 6.1$, $p < 0.001$; $F_2(8,40) = 9.5$, $p < 0.001$. The significant three-way interaction indicates that the additive effect of Word Position x Register differs depending on the focus condition of the sentence. As predicted, the magnitude of CDS elongation in Word 1 is greatest in the Initial-Focus condition whereas the same effect in Word 5 is greatest in the Final-Focus condition (Figure 1, row 2).

Overall, these results indicate that utterance-final lengthening, which is augmented in CDS due to the register elongation effect, is further enhanced when the utterance-final word has a focal pitch accent.

Additive effect of Sentence Mode

Lastly, we turn to the effects of Sentence Mode, i.e. whether the sentence is a statement or question, to investigate if there is any difference in the extent of utterance-final lengthening in

CDS conditioned by the different modes of sentences. To investigate this question, we constructed a model with the inclusion of the fixed factor Sentence Mode, in addition to Word Position and Register. Repeated measures ANOVAs show all main effects and interactions to be significant (all p's < 0.05 or smaller), including the three-way interaction Word Position x Register x Sentence Mode ($F_1(4,16) = 4.8$, $p < 0.01$; $F_2(4,20) = 9.5$, $p < 0.001$). These effects reflect the tendency for a greater extent of utterance-final lengthening in CDS statements than in CDS questions (Figure 1, row 3).

To sum up, elongation effects of CDS register were found *across* the sentence, but the magnitude of the elongation was greatest on the utterance-final word. This enhanced utterance-final lengthening in CDS was further augmented by the additional factor of final-focus and the Sentence Mode statement. In other words, the Focus and Sentence Mode factors contribute additively with CDS and utterance-final lengthening.

Discussion

Our study had two main goals: to re-evaluate the previous claim that CDS lengthening is exclusively, or primarily located in the final position of a sentence, and to test the hypothesis that the often-cited exaggerated utterance-final lengthening in CDS is due to the additive effects of multiple factors that are present in a high proportion in CDS.

With regard to the first question, our results show that the elongation effect in CDS is not limited to the final syllable, but rather is a more general effect across the sentence. Thus our results failed to support the idea that the slow tempo in CDS is solely or primarily due to final-lengthening effects as previously claimed. However, we did find a greater extent of lengthening in the final syllable compared to the lengthening in other word positions in CDS. Our analyses suggest that this effect is due to the conflated effects of utterance-final position and the register.

Interestingly, we found that the ratio of each word, including the final syllable, in the sentence remained nearly constant across the two speech registers. In other words, although the final syllable underwent the greatest elongation in CDS, its overall proportion in the sentence remained the same. This finding raises an interesting question about how infants might process the durational cues at the sentence level. As previously discussed, many researchers (including the second author) have at least implicitly argued that utterance-final lengthening is exaggerated in CDS to provide enhanced utterance-boundary cues in duration. However, little is known about the exact mechanism of how infants process temporal cues. In studies of speech perception, it has been suggested that adult listeners might be using relative rather than absolute duration in mapping the signal to phonological categories (Port & Dalby, 1982; Volaitis & Miller, 1992). If these findings hold at the utterance level, an enhanced duration cue for the boundary would mean an increase in the proportion of the final syllable. Our results, however, failed to provide evidence for such hypothesis. Alternatively, some studies suggest that timing is intrinsic to the articulatory specifications of a phonetic category that listeners experience with speech, which they preserve in long term memory and directly match with the temporal dynamics in the perceptual structure (Pisoni, 1993; Summerfield, 1981). Under this hypothesis, children would store a host of cues signaling the boundary including prosodic (e.g. duration, F0, and intensity) and spectral characteristics (e.g. Tabain, 2003) in memory, and the longer duration could provide extra temporal capacity for processing the various other information relevant to identifying the boundary. Thus, while our data does not provide direct evidence regarding the processing mechanism of temporal cues, there seems to be a good reason to think that the elongation in duration of the final syllable will be beneficial for children to process information on phrasal boundaries and thus may increase the salience of the boundary cues for the child listener.

Importantly, the increase in utterance-final lengthening in CDS is further enhanced in statements and by the additive effect of final-focus. Given the tendency for CDS to place focus on the final word of an utterance (Fernald & Mazzie, 1991; Woodward & Aslin, 1990), the additive effect of Focus on utterance-final lengthening is expected to be influential in natural speech. The additive effect of Sentence Mode on utterance-final lengthening is a somewhat less discussed issue, but the tendency of the extra utterance-final lengthening in final-focused statements compared to final-focused questions has been reported in Eady & Cooper (1986). In our data, such effect was further augmented by the addition of the Register factor. Although CDS is characterized by a relatively higher proportion of questions compared to ADS, statements still account for the greatest proportion of the sentence modes in CDS (Fisher & Tokura, 1996; Fernald & Mazzie, 1991; Soderstrom, Blossom, Foygel, & Morgan, 2008). Therefore, while the relatively higher proportion of questions in CDS may not contribute to the reason that previous studies found strong effects of utterance-final lengthening in CDS compared with ADS, CDS statements with final focus are nonetheless likely to be frequent in a child's input. Our data suggest that these sentences types will converge to the strongest utterance-boundary cues. Our findings may therefore have implications for clinical interventions in children with language disorders and delays. While utterance-boundary lengthening itself is not something under conscious control, a speaker may well choose to use a high proportion of sentences with focused words in final position. Our findings suggest that in addition to increasing the salience of the focused item for lexical acquisition, the focal pitch accent on the final syllable may also have important influences on the perception of sentence boundaries due to the increased temporal capacity to process relevant cues (Marinis, 2011). This hypothesis is supported by the finding

that children with SLI perform better in word learning tasks if the target words are presented in slower speaking rate (e.g. Ellis Weismer & Hesketh, 1996).

Before closing the discussion, one important limitation of our design is worthy of consideration. Our study is based on data from simulated dialogues rather than naturally occurring speech, thus we cannot be certain that our simulated CDS exactly mimics the characteristics of true CDS under natural conditions. Nevertheless, there is good reason to believe that our samples tap into real speech phenomena. While our participants were not given any direction regarding the characteristics we were expecting to see in their speech, all of them elongated the sentences in CDS, and showed systematic differences across the focus conditions and sentence modes. As mentioned earlier, previous research indicates that characteristics of CDS observed in spontaneous speech is replicated in simulated CDS at least to a certain extent, and with greater efficiency when elicited from trained speakers (Knoll et al., 2009). As methods for analyzing large quantities of spontaneous speech (and in particular, for controlling for focus-type effects in such samples) improve, it will be important to replicate our findings under more naturalistic conditions.

Conclusion

We investigated the durational patterns of CDS across the sentence, and the additive effects of speech register, word location, focus and sentence mode on the utterance-final monosyllabic word under systematically controlled conditions using simulated speech data. Our data suggest that the elongation in CDS is not restricted to the final syllable. In addition, the discourse-related factors of focus and sentence mode showed additive effects with register and word position. We propose that the exaggerated utterance-final lengthening found in CDS in

natural speech is due to the discourse characteristic of CDS which contains a high proportion of final-focus sentences.

References

Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother-child speech. *Journal of Phonetics, 14* (2), 303-309.

Boersma, P. & Weenink, D. (2010). Praat: doing phonetics by computer [Computer program]. Version 5.1.42 retrieved from http://www.praat.org/.

Church, R., Bernhardt, B., Pichora-Fuller, K., & Shi, R. (2005). Infant-directed speech: Final syllable lengthening and rate of speech. *Canadian Acoustics*, *33* (4), 13-20.

Cooper, W.E., Eady, S.J., Mueller, P.R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America*, 77, 2142-2156.

Eady, S. J & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America, 80*, 402-416.

Ellis Weismer, S., & Hesketh, L. (1996). Lexical learning by children with specific language impairment: Effects of linguistic input presented at varying speaking rates. *Journal of Speech and Hearing Research, 39*, 177-190.

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, 104-113.

Garnica, O. (1977). Some prosodic and paralinguistic features of speech to young children. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition*. Cambridge, MA: Cambridge University Press, 63-88.

Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Wright Cassidy, K., Druss B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition* 26, 269-286.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In Bybee, J., & Hopper, P. (Eds.), *Frequency and the Emergence of Linguistic Structure*. Amsterdam: Benjamins, 229–254.

Klatt, D. H. (1973). Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America, 54(4),* 1102-1104.

Knoll, M., Scharrer, L., & Costall, A. (2009) Are actresses better simulators than female students? The effects of simulation on prosodic modifications of infant- and foreigner-directed speech. *Speech communication, 51(3),* 296-305.

Lam, C., & Kitamura C. (2010). Maternal Interactions with a Hearing and Hearing-impaired Twin: Similarities and Differences in Speech Input, Interaction Quality, and Word Production, *Journal of Speech Language and Hearing Research*, 53, 543-555.

Lindblom, B., and Rapp, K. (1973). Some temporal regularities of spoken Swedish. *Papers in Linguistics from the University of Stockholm, 21*, 1–59.

Marinis, T. (2011). On the nature and cause of Specific Language Impairment: A view from sentence processing and infant research. *Lingua*, 121, 463-475.

Morgan, J.L. (1986). From simple input to complex grammar. Cambridge, MA: MIT Press.

Morgan, J. L., Meier, R. P., and Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology, 19*, 498-550.

Nooteboom, S. (1972). Production and perception of vowel duration: A study of durational properties of vowels in Dutch. Ph.D. thesis, Utrecht University.

Oller, K. (1973). The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America, 54(5),* 1235-47.

Pisoni, D. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning. *Speech communication*, *13*, 109-125.

Port, R. F. and Dalby, J. (1982). Consonant/vowel ratio as a cue for voicing in English. *Perception and Psychophysics, 32(2)*, 141–152.

Schaeffler, F., Kempe, V., Biersack, S. (2006). Comparing vocal parameters in spontaneous and posed child-directed speech. Paper presented at the 3rd Speech Prosody Congress, Dresden, Germany. Retrieved from http://www.isca-speech.org/archive/sp2006/papers/sp06_227.pdf

Seidl, A. (2007) Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language, 57,* 24-48.

Singh, L., Morgan, J. L., & Best, C. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy, 3,* 365-394.

Soderstrom, M., Kemler Nelson, D. G., & Jusczyk, P. W. (2005). Six-month-olds recognize clauses embedded in different passages of fluent speech. *Infant Behavior and Development, 28,* 87-94.

Soderstrom, M., Blossom, M., Foygel, R., & Morgan, J.L. (2008). Acoustic cues and grammatical units in speech to two preverbal infants. *Journal of Child Language, 35*, 869-902.

Summerfield, A . Q. (1981). Articulatory rate and perceptual constancy in phonetic perception. *Journal of Experimental Psychology: Human Perception Performance*, *7*, 1074-1095.

Swanson, L. & Leonard, L. (1994). Duration of function-word vowels in mothers' speech to young children. *Journal of Speech and Hearing Research*, 37, 1394-1405.

Trainor, L., Austin, C. & Desjardins, R. (2000) Is infant-directed speech prosody a result of the vocal expression of emotion?. *Psychological Science,* 11(3), 188-195.

Volaitis, L. E. & Miller, J. L. (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *J. Acoust. Soc. Am.* 92, 723-735.

Woodward, J. Z., & Aslin, R. N. (1990, April). Segmentation cues in maternal speech to infants. Poster presented at the International Conference on Infant Studies, Montreal, Canada, cited in Aslin,R., Woodward, J., Lamendola, N. & Bever, T. (1996). Models of word segmentation in fluent maternal speech to infants. In Morgan, J. & Demuth, K. (eds). *Signal to Syntax*, Mahwah NJ: Lawrence Erlbaum Associates.

Yuan, J. & Liberman, M. (2008). Speaker identification on the SCOTUS corpus, *Proceedings of Acoustics 2008*, 5687-5690.

Yuan, J., Liberman, M. & Cieri, C. (2006). Towards an Integrated Understanding of Speaking Rate in Conversation. *Proceedings of Interspeech 2006*, 541-544. Pittsburgh, PA.

Table 1.

Sample dialogues for eliciting a sentence in two registers (ADS/CDS), two sentence modes (statement/question) and three focus conditions (neutral/initial-focus/final-focus).

ADS	CDS	Sentence Mode	Focus
A: What happened? B: Lee held a long worm.	A: Tell Johnny what happened. B: Lee held a long worm.	statement	neutral
A: Who held a long worm? B: Lee held a long worm.	A: Tell Johnny who held a long worm. B: Lee held a long worm.	statement	initial
A: What did Lee hold? B: Lee held a long worm.	A: Tell Johnny what Lee held. B: Lee held a long worm.	statement	final
A: Ask what happened. B: Lee held a long worm?	A: Ask Johnny what happened. B: Lee held a long worm?	question	neutral
A: Ask who held a long worm. B: Lee held a long worm?	A: Ask Johnny who held a long worm. B: Lee held a long worm?	question	initial
A: Ask what Lee held. B: Lee held a long worm?	A: Ask Johnny what Lee held. B: Lee held a long worm?	question	final

Table 2.

Mean duration (ms) and ratio of words in ADS and CDS by Sentence Mode. The duration in each word position is increased across the sentence in CDS, but the ratio remains almost constant across the two registers.

Word Position	ADS statements (SD)	CDS statements (SD)	ADS questions (SD)	CDS questions (SD)	mean ratio statement (CDS/ ADS)	mean ratio question (CDS/ ADS)
W1	315.2 (71)	417.3 (118)	324.7 (77)	417.5 (124)	1.32	1.29
W2	242.2 (40)	303.8 (104)	233.3 (40)	296.5 (86)	1.25	1.27
W3	121.0 (90)	149.3 (121)	121.5 (85)	149.1 (112)	1.23	1.23
W4	211.3 (109)	284.0 (168)	203.1 (103)	264.3 (153)	1.34	1.30
W5	417.9 (83)	560.8 (139)	406.7 (71)	497.1 (124)	1.34	1.22
Total	1307.6 (149)	1715.2 (249)	1289.2 (121)	1624.5 (256)	1.31	1.26

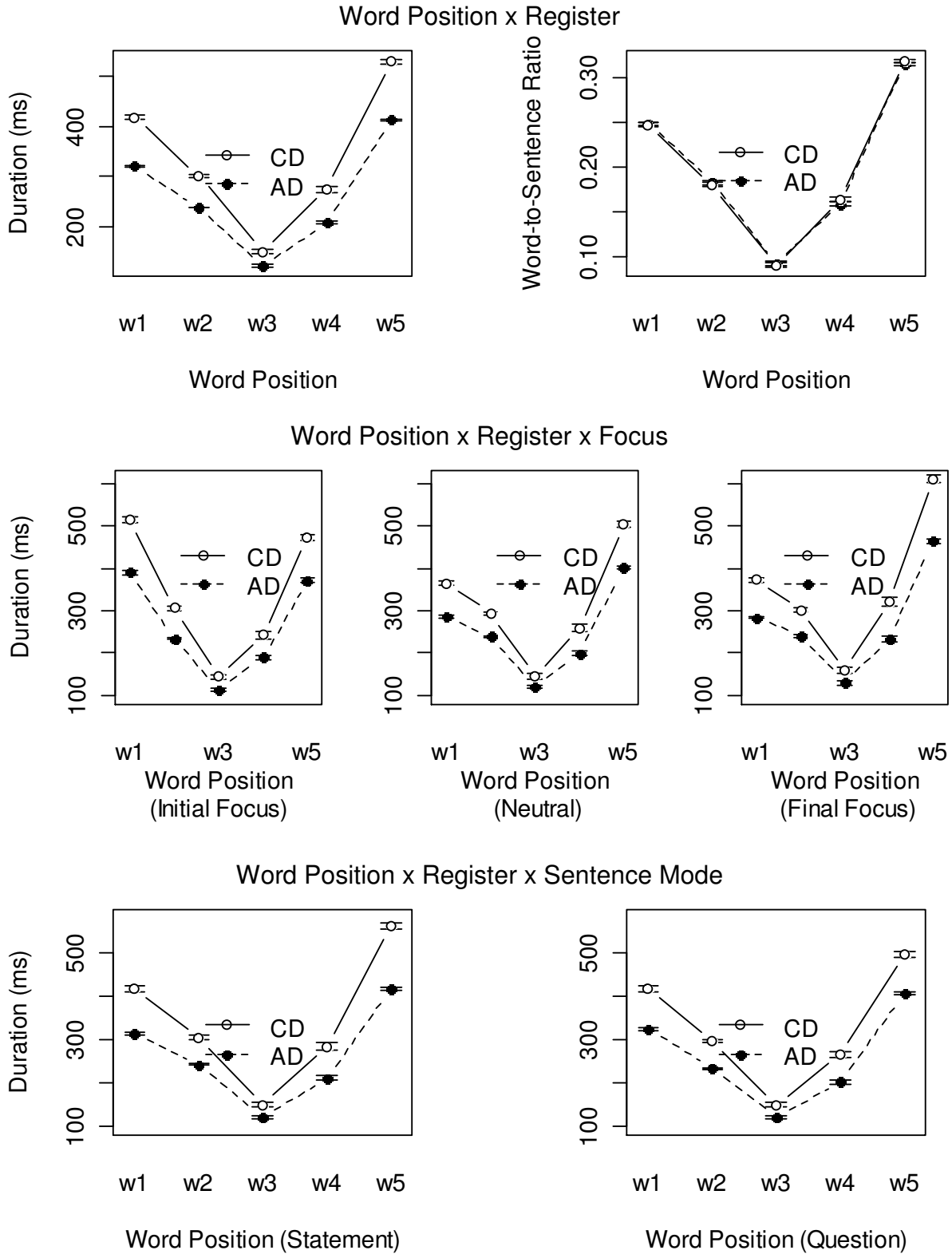

Figure 1. Two- and three-way interactions in Word Position, Register, Focus, and Sentence Mode. All illustrated interactions were significant, converging to the greatest extent of elongation on the pre-boundary syllable (W5) in CDS in the final-focus condition and statements.