Prosodic structure in child speech planning and production

Ivan Yuen[1], Katherine Demuth[2] and Mark Johnson[3]

[1,2]Department of Linguistics, [3]Department of Computing, Macquarie University
[1]ivan.yuen@mq.edu.au, [2]katherine.demuth@mq.edu.au, [3]mark.johnson@mq.edu.au

ABSTRACT

English-speaking children have acquired phrase-final lengthening by the age of 2, but other aspects of prosodic organization appear to be later acquired. This study investigated 2-year-olds' prosodic organization of function words in an elicited imitation task. In particular, we wanted to know if children would prosodify pronouns 1) as part of a trochaic foot with the preceding word, or 2) as a separate prosodic word. The results showed that the function word was produced as an independent prosodic unit, in contrast to the adult model being imitated. Implications for a developmental model of speech planning and production are discussed.

Keywords: prosodic structure, speech planning, language acquisition.

1. INTRODUCTION

It has long been known that prosodic structure plays an important role in the organization of speech. For instance, early research on utterance-final lengthening [3] led to further studies showing that the degree of constituent-final lengthening is proportional to the level of constituency in the prosodic hierarchy ([5], [7] and [12]). In a review on prosodic organization, Shattuck-Hufnagel & Turk [9] raised many questions about how the instantiation of phonetic parameters such as duration, F0 and amplitude are influenced by the prosodic structure of a particular spoken utterance.

At the same time, results from studies of early language acquisition have shown that 2-year-olds' omission of certain grammatical morphemes is influenced by the prosodic structure in which it appears. Using an elicited imitation task, Gerken [6] showed that target function words that could be prosodified as part of a disyllabic trochaic foot with the preceding word (1a) were more likely to be produced than those that occurred in an unfooted context (1b).

1a. Tom [likes <u>the</u>]$_{FT}$ piggy.
1b. Tom [catches]$_{FT}$ <u>the</u> piggy.

Demuth & McCullough [4] replicated this finding in a longitudinal study of children's spontaneous speech, suggesting that this is a robust effect. However, one child produced articles as an independent stressed unit (as a separate prosodic word (PW)) rather than prosodifying it as part of a disyllabic foot. This raises the question of how and when children begin to show adult-like prosodic organization.

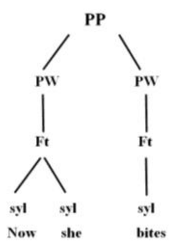
The goal of the current study was to further explore this issue of prosodic organization in early speech. To do this, we examined children's productions of minimal pair sentences containing a target verb and a personal pronoun, and compared their productions with the pre-recorded adult model they heard during an elicited imitation task. In one set of stimulus sentences the target verb was in medial position, preceded by an utterance-initial target pronoun (2a). In the paired sentence the target verb occurred in utterance-final position, preceded by a now utterance-medial target pronoun (2b).

2a. <u>She</u> <u>bites</u> pears.
2b. [Now <u>she</u>]$_{FT}$ <u>bites</u>.

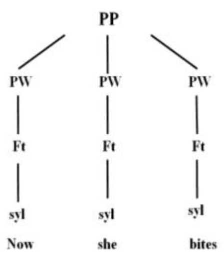
Given that previous studies had shown that 2-year-old children speaking American English control phrase-final lengthening [10], we predicted that the vowel of the lexical item in final position would be longer than the vowel of the same word in utterance-medial position. This served as a control to ensure that the task was tapping into children's abilities to signal prosodic structure in production. The primary interest, however, was in the pronoun. If children exhibited adult-like prosodic organization, we predicted they would produce the function word as an independent, stressed prosodic word when it occurred utterance initially (2a), but as a weak (unstressed) syllable of a trochaic foot when it occurred utterance medially (2b). This would be realized by a shorter vowel in the footed condition (2b). On the other hand, if children produced the function word in (2b) as an independent prosodic word, there should be no vowel duration differences for the function word across the two sentence conditions. This would indicate that these children are not yet adult-like in the prosodic organization of their utterances.

These two possible prosodic organizations for (2b) are illustrated below in terms of the prosodic hierarchy consistent with [6] and [8]. The form in (3a) represents the adult-like form with the pronoun which is prosodified as part of a foot. In contrast, the form in (3b) represents the possible child form, where the pronoun is realized as a separate foot (Ft) and prosodic word (PW) within the phonological phrase (PP):

(3a)



(3b)



2. METHODOLOGY

2.1 Subjects

Data were analyzed from eight monolingual children (5 girls and 3 boys) aged 2;0-2;6 (mean 2;3 years) from southern New England, USA. MacArthur CDI percentile scores on vocabulary size ranged from 25-99%, with a mean of 69%. Data from an additional 6 children were not used in this study due to poor acoustic quality, difficulty identifying segmental boundaries (in the case of /h/ onsets), exaggerated speech behaviour, or use of non-target word items, thereby reducing the number of useable data points.

2.2 Stimuli and Procedure

The 16 3-syllable, 3-word stimulus sentences were constructed containing 8 different target verbs placed in either sentence-medial or sentence-final position. These were all preceded by the pronoun *he* or *she*, which occurred either initially or medially within the sentence, as shown in Table 1.

Table 1. Stimulus set

She *hits* balls	Now **she** *hits*
He *fights* bears	Now **he** *fights*
He *cooks* peas	There **he** *cooks*
She *bakes* pies	There **she** *bakes*
He *sits* back	There **he** *sits*
She *bites* pears	There **she** *bites*
She *talks* back	Now **she** *talks*
He *takes* books	Now **he** *takes*

All target verbs were CVCC present tense inflected forms, both picturable and familiar to 2-year-olds. The stimulus sentences were then paired with 8 pictures and accompanying auditory prompts which had been prerecorded by a female American-English speaker in child-directed speech. These were then embedded in a power point presentation and played for the participants on a computer. This also provided the adult model that was analyzed for this study.

The children were invited into a sound-attenuated test room to play a game with the experimenter. Children looked at the computer monitor and repeated what the 'puppet' said. The room was equipped with a lavalier microphone (Audio-technica 700 Series) connected to a computer in an adjoining room via an MBox 2 Audio Interface (Digidesign) for recording. The data were later downloaded for analysis. (See [11] for full details of the study from which our study was drawn.)

2.3 Coding

A total of 8 sentence pairs from the adult speaker and 42 sentence pairs from the 8 children were analyzed using Praat [2]. As we were interested in the relative vowel durations as a function of prosodic position within the utterance, two vowel durations (function word, lexical item) were measured for each sentence. The pronouns were coded as FUNCi (function word – initial) and FUNCm (function word – medial). The target verbs were coded as LEXm (lexical item – medial) and LEXf (lexical item – final).

The following criteria were used to identify the two vowels of interest. F2 onsets and offsets were used to determine the vowel onsets and offsets. As the vowels in FUNCi and FUNCm were preceded by either /ʃ/ or /h/, offset of frication was used to aid the identification of vowel onset. At the vowel offset, two additional cues were used to aid segmentation, namely, the onset of frication or the onset of closure, given the types of onset consonants in the lexical verbs. However, if either a lag or an overlap of the acoustic cues was observed, the F2 cue took precedence.

The same criteria were applied to annotating the vowel onsets and offsets of LEXm and LEXf. Identification of the vowel offset was also aided by the onset of closure in the following plosive coda.

Vowel durations for FUNCi, FUNCm, LEXm, and LEXf were then automatically extracted using Praat, and consolidated for statistical analysis.

3. RESULTS

Mean vowel duration for FUNCi, FUNCm, LEXm, and LEXf is shown in Table 2. Mean vowel duration reflects performance across stimuli and speakers for the children, and across stimuli for the adult. These data are shown in Figure 1.

Table 2: Average vowel durations of function and lexical items in two utterance positions.

	FUNCi	FUNCm	LEXm	LEXf
Child	183 ms	170 ms	164 ms	187 ms
Adult	187 ms	97 ms	155 ms	168 ms

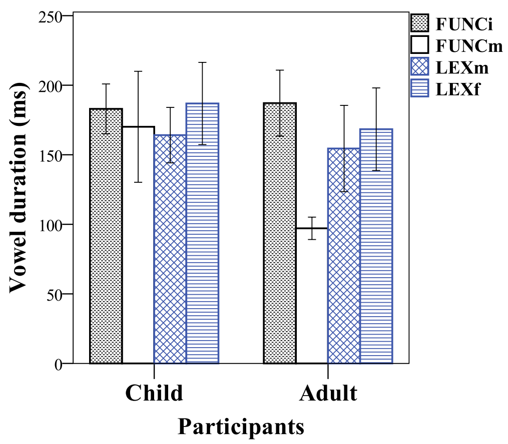
As described earlier, we predicted that the adult would shorten vowel in FUNCm due to prosodification of FUNCm with the previous word. For the children, we made two predictions. First, the vowel durations in LEXf would be longer than LEXm as a result of phrase-final lengthening. This served as a control to ensure that the elicitation task tapped into children's abilities to signal prosodic organization. The second prediction concerned prosodification of the function word. If children have acquired both levels of prosodic organization as in the adult, we predicted that the child data would pattern in the same way as the adult data.

Two paired t-tests were conducted for the adult and child data each. In each data set the first t-test compared mean vowel duration across positions (initial vs. medial) for the function word, and the second across positions (medial vs. final) for the lexical item.

First, we reported the results from the adult. An item analysis of the 16 prompts presented to the children showed that vowel duration of FUNCm was significantly shorter than vowel duration of FUNCi as predicted [$t(7) = -8.415$, $p < 0.0001$], indicating that pronouns in medial position were in fact prosodified as a trochaic foot with the pervious word. In terms of vowel durations for the lexical items, results showed no evidence that LEXm was shorter than LEXf [$t(7) = 1.52$, $p = 0.172$], indicating that phrase-final lengthening as measured in vowel duration was not observed, though there was a numerical trend in the appropriate direction. This suggests that the adult might lengthen the rime or word as a whole, rather than the vowel nucleus to signal phrase-final lengthening. In fact, a follow-up analysis in [11] confirmed that phrase-final lengthening was observed using both word duration and morpheme duration as metrics.

In terms of the children, results showed that, unlike the adult, there was no reliable difference in vowel duration for FUNCi and FUNCm [$t(7) = -0.813$, $p = 0.443$]. This indicates that children were producing the function word in medial position as a unique prosodic word, and not part of a larger prosodic unit as was observed for the adult. In addition, the results showed a difference in vowel duration for LEXm and LEXf [$t(7) = 2.693$, $p = .031$]. This suggests that the children lengthened vowel of the lexical verb phrase-finally to signal prosodic structure.

Figure 1: Vowel durations of FUNCi, FUNCm, LEXm and LEXf for children and adults. (Error bars: +/- 2 SE)



4. DISCUSSION

Consistent with [10], our findings confirmed the phrase-final lengthening effect for the children. Using vowel duration as the metric, the adult also showed a trend in the same direction.

As regards the prosodic organization of the function item (pronoun) we also observed a difference between children and the adult. The adult showed vowel shortening of the function word utterance-medially as a result of prosodifying the pronoun as part of a foot with the previous word, whereas the children did not. This suggests that 2-year-olds have yet to fully acquire aspects of prosodic organization which will lead to reduction of some words within the phonological phrase (cf. [4]).

There was also a large amount of inter-speaker variability in the duration of the FUNCm vowel, with some children approximating adult-like durational values and others not. Although there were no correlations with either vocabulary size or chronological age, it appears that some 2-year-olds may exhibit more advanced prosodic and sentence planning organization than others. Further research would want to probe these issues further, exploring the various factors that may influence utterance-level planning [1].

Our findings then suggest that children as young as 2;3 are aware of phrase-level lengthening. However, within the phrase, their prosodic organization is fairly flat (3b), with the hierarchical adult structures still in the process of being acquired. This is consistent with previous results in the field using slightly different methods [4], [6]. These studies also suggest that access to higher-level prosodic representations may be developing around 2;6. We might therefore expect more adult-like prosodic organization in the speech of children by the age of 3.

5. CONCLUSION

This study aimed to examine children's prosodic organization of function words and lexical items in an elicited imitation task. As predicted, the results showed that the child participants lengthened the vowel of the lexical item in utterance-final position, exhibiting domain-edge effect. This indicates that the task successfully tapped into children's ability to signal phrase-final prosodic organization. However, children did not organize the function word in the same way as the adult. Unlike the adult, children did not prosodify the pronoun as part of a foot with the previous word, but produced it as an independent prosodic word. This suggests that there is still much to be learned about the prosodic organization of children's utterances, and how this interacts with the acquisition of speech planning more generally.

6. REFERENCES

[1] Aylett, M., Turk, A. 2004. The Smooth Signal Redundancy Hypothesis: A Functional Explanation for Relationships between Redundancy, Prosodic Prominence and Duration in Spontaneous Speech. *Language and Speech*, Volume 47(1), 31-56.

[2] Boersma, P., Weenink, D. 2005. Praat:doing phonetics by computer. http://www.praat.org/.

[3] Cooper, W.E., Paccia-Cooper, J. 1980 *Syntax and speech*. Cambridge, MA: Havard University Press.

[4] Demuth, K., McCullough, E. 2009. The prosodic (re)organization of children's early English articles. *Journal of Child Language*, Volume 36, 173.

[5] Ferreira, F. 2007. Prosody and performance in language production. *Language and Cognitive Processes*, Volume 22 (8), 1151-1177.

[6] Gerken, L.A. 1996. Prosodic structure in young children's language production. *Language*, Volume 72, 683–712.

[7] Ladd, D.R., Campbell, N. 1991. Theories of prosodic structure: evidence from syllable duration. In *Proc. 12th ICPhS*, Aix-en-Provence, 290-293.

[8] Selkirk, E.O. 1996. The prosodic structure of function words. In *Signal to syntax: bootstrapping from speech to grammar in early acquisition*, 187-213. Mahwah, NJ: Lawrence Erlbaum.

[9] Shattuck-Hufnagel, S., Turk, A.E. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, Volume 25(2), 193-247.

[10] Snow, D. 1998. Phrase-final syllable lengthening and intonation in early child speech. *Journal of Speech and Hearing Research*, Volume 37, 831-840.

[11] Theodore, R., Demuth, K., Shattuck-Hufnagel, S. In submission. An acoustic examination of the effects of word frequency and utterance position on 2-year-olds' production of 3rd person singular –s.

[12] Wightman, C., Shattuck-Hufnagel, S.,Ostendorf, M. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of Acoustical Society of America*, Volume 91, 3, 1707-1717.

7. ACKNOWLEDGEMENTS

We thank Stefanie Shattuck-Hufnagel and Rachel Theodore for helpful comments and suggestions. This research was funded in part by Macquarie University, and the following grants: NICHD R01 HD057606 (Demuth & Shattuck-Hufnagel) and ACR DP110102506 (Johnson & Demuth).