Prosodic Marking of Continuation versus Completion in Children's Narratives

Melissa A. Redford[1], Laura C. Dilley[2,3], Jessica L. Gamache[3], and Elizabeth A. Wieland[2]

[1]Department of Linguistics, University of Oregon, Eugene, Oregon, USA
[2]Department of Comm. Sci. & Disorders, Michigan State University, East Lansing, Michigan, USA
[3]Department of Linguistics & Germanic, Slavic, Asian, and African Languages, Michigan State University, East Lansing, Michigan, USA

redford@uoregon.edu, ldilley@msu.edu, gamache1@msu.edu, wielande@msu.edu

Abstract

Discourse prosody in school-aged children's narratives was investigated to test for developmental changes in transitional prosody and to characterize the acquisition of key contrastive features for marking continuation versus completion. Spontaneous narratives were obtained from 42 children (5 to 7 years old) and 14 adult caregivers. The narratives were prosodically transcribed using a formal annotation system that relies on perceptual and acoustic analyses. Metalinguistic judgments of prosodic function and appropriateness were also obtained. Analyses examined effects of age on the use of transitional prosody as well as on prosodic cues to continuation and completion. Results were that children marked phrases for completion less often and less appropriately than adults. Children and adults both used phrase-final tones and post-boundary pauses to mark continuation versus completion; however, children's overall higher rate of pausing lessened the extent to which continuation and completion were differentiated through pausing.

Index Terms: speech acquisition, speech prosody, discourse prosody, prosodic boundaries

1. Introduction

Discourse is hierarchically structured in that larger ideas are constructed out of smaller ones. The hierarchical structure of discourse is expressed using morphosyntactic devices such as anaphor [1] and prosodic devices that cue relationships between ideas and events [2]. An example of a discourse prosodic device is something we will call *transitional prosody*: the prosody that glues spoken "sentences" into spoken "paragraphs." Specifically, transitional prosody indicates whether a speaker intends to continue with a thought or has completed it [3].

Transitional prosody creates a global prosodic structure through the alternation of phrases prosodically marked for continuation versus completion. This structure parallels the information structure referenced by the morphosyntactic system. According to some functional theories of discourse, there is a direct, quasi-reflexive relationship between complex thought and prosody [3], [4]. Whereas such theories are consistent with the idea that prosody is independent of other linguistic systems such as syntax (e.g., [5]), the hypothesis of prosody-as-thought in some senses belies, or at least diminishes, the thesis that prosody is a sophisticated linguistic marking system in its own right. By contrast, the current study departs from this thesis. Accordingly, we investigated the acquisition of transitional prosody on its own terms.

Transitional prosody represents a challenge for acquisition in that continuation and completion are cued using a combination of tones and timing (e.g., [6]). Moreover, the cues are used in a probabilistic fashion and the relative salience of each is poorly understood. For example, completion judgments are more often associated with final low tones than with final high tones, but can be made in the presence of high tones [7], [8]. Similarly, completion judgments are more likely in the presence of a pause, but can be made in the absence of a pause [ibid.]. Insofar as developmental patterns are driven by children's efforts to optimally match salient adult targets, studying development may help us identify the key features and combinations of features used to mark continuation versus completion.

There is little prior research on the acquisition of functional prosody. The few studies that exist indicate that development continues through middle childhood. For example, some functional prosodic contrasts are not mastered until 8 years of age [9]. Also, the on-line use of some prosodic contrasts differs in younger and older children. For example, the relative frequency of continuation versus completeness marking is lower in 7-year-olds than in 11-year-olds [10]. The current study focused on the use and marking of transitional prosody in school-age children between the ages of 5 and 7 years old, a period that represents a reasonably early stage in acquisition of discourse-related prosody. Analyses examined the frequency and appropriateness of transitional prosody in spontaneous narratives, as well as the key characteristics of the prosodic marking based on a systematic, perceptual- and acoustically-based annotation of boundary tones, pitch accents, prominences, and pauses.

2. Methods

2.1. Participants

Forty-two native English-speaking children and their mothers participated in the study. All participants resided in the Eugene-Springfield area in Oregon and spoke a west coast dialect of American English. Fourteen of the children were 5 years old (*M* = 65 mos., *SD* = 1.63 mos.), 14 were 6 years old (*M* = 77 mos., *SD* = 2.17 mos.), and 14 were 7 years old (*M* = 89 mos., *SD* = 1.61 mos.). All children were typically developing as determined by parental report. All passed a pure tone hearing screening.

Although an accompanying caregiver always participated in the task, we randomly identified a subset of 14 mothers to serve as our adult comparator group. In this way, the study included 14 individuals in each of four age groups: 5-year-olds, 6-year-olds, 7-year-olds, and adult females.

2.2. Task

A story telling task was used to elicit spontaneous narratives. The child and his/her caregiver each chose one of 4 wordless picture books to tell a story to the other. While the caregiver familiarized herself with her choice, an experimenter helped the child look through his/her book. Once both participants were ready, each told a story to the other. The stories were digitally recorded for later analysis. The participant not telling the story was instructed to not interrupt.

Once both participants had finished with one telling of the story, each retold the same story to the other one more time. The second story telling was used to elicit the most fluent spontaneous speech possible and the best possible story [11].

2.3. Prosodic Labeling

Prosodic labeling is a time-intensive endeavor. For this reason, we restricted labeling to a 30-second sample extracted from each participant's second story telling. After excluding the first 30 seconds to avoid the stereotyped prosody associated with the beginning of stories, a speech sample was chosen by taking the earliest 30-second fragment that maintained complete sentences from each narrative. Any interruptions that may have occurred during the telling (e.g., a cough or interjection from a listener) were excluded.

We used the Rhythm and Pitch (RaP) system of prosodic annotation to identify rhythmically prominent syllables, prosodic boundaries, tones, and pitch accents [12]. RaP is similar in many respects to the well-known ToBI system [13]. RaP judgments are based on auditory-perceptual evaluations of speech. The mapping from phonetic correlates to labels is simpler and more consistent in RaP than in ToBI [12]. Also, unlike in ToBI, pitch accents are reserved for prominences with observable pitch excursions in RaP RaP has been validated on both adult and child speech [12], [14].

A team of trained analysts used the RaP system to label the location and types of perceptual prominences, prosodic boundaries, pitch accents, and tones (low, high, equal) in the speech samples. Analysts also orthographically transcribed the speech, and noted any pauses and disfluencies. Reliability between analysts, assessed using the kappa (κ) metric, was generally good. There was a fair level of agreement for tone type ($\kappa = .40$), a moderate level of agreement for prominence type ($\kappa = .60$), and a good level of agreement for boundaries ($\kappa = .62$) and pitch accents ($\kappa = .70$).

Once the systematic prosodic labeling was complete, analysts made two critical metalinguistic judgments. The first was whether the prosody of phrases delimited by strong boundaries indicated that the speaker intended to continue with a thought (A), or whether the speaker had completed a thought (B). Phrases with judgment A will be referred to as marked for continuation. Those with judgment B as marked for completion. The second metalinguistic judgment was whether the marking was appropriate given the local semantic and discourse context. Analysts had good agreement in their metalinguistic judgments (prosodic marking, $\kappa = .86$; appropriateness, $\kappa = .88$).

All strong boundaries were counted for every sample from every speaker. Boundary-delimited phrases were coded as continuing or completed, and the sequencing pattern of continuation and completion marking was quantified for each speaker. Appropriateness judgments and number of words were also tallied for every phrase.

Given that multiple prosodic cues to continuation and completion likely exist, we chose to investigate a wide range of tone and timing features. The dependent variables were as follows: a count of initial high (versus low) tones; occurrences of melodic continuity between phrases (i.e., initial equal tone versus other initial tones); a count of final pitch accents and final prominences, whether weak or strong; occurrences of post-boundary pausing; and pause duration in millisecond. All count measures were normalized against the total number for each functional marking type and for each speaker. Several speakers did not produce any post-boundary pauses after phrases with a particular functional marking, which is why the degrees of freedom in analyses of pause duration differ from those of the other analyses.

3. Results and Discussion

Three sets of analyses were conducted. The first was used to characterize the speech samples obtained from children and adults, the second to test for age-related differences in the use of transitional prosody, and the third to describe the prosodic marking of continuation and completion in child and adult speech.

3.1. Speech Sample Characteristics

Nonparametric (Kruskal-Wallis) one-way ANOVA tests indicated a significant effect of age group on the number of words produced per sample [$H(3) = 23.73, p < .001$], the number of strong boundaries per sample [$H(3) = 8.81, p = .032$], and the number of disfluencies per sample [$H(3) = 20.88, p < .001$], but not on the total number of pauses per sample [$H(3) = 5.93, p = .12$]. As for the direction of these effects, the mean number of total words was higher in the adult samples than in the child samples (child $M = 51.62$; adult $M = 77.35$). The mean number of disfluencies was lower in the adult samples than in the child samples (child $M = 4.64$; adult $M = 1.21$). The effect of age group on the number of boundaries was driven by 7-year-olds, who produced fewer identifiable strong boundaries on average than participants in any other age group (5-year-old $M = 17.57$; 6-year-old $M = 20.00$; 7-year-old $M = 15.93$; adult $M = 19.07$). The mean number of pauses produced per sample was 11.57.

Apart from the effect of age on the number of strong boundaries, all results were consistent with expectations based on previous comparisons of child and adult language (e.g., [11]). More importantly, the analyses suggested that the 30-second speech samples provided sufficient information to test for group differences in the use of transitional prosody.

3.2. Use of Transitional Prosody

One-way ANOVA tests indicated significant effects of age group on the proportion of phrases marked for completion [$F(3,55) = 3.69, p = .017$], and on the proportion of phrases judged to have appropriate transitional prosody given the local semantic and discourse context [$F(3,55) = 6.41, p = .001$]. There was no effect age group on the sequencing of continuation and completion marking. The results for the metalinguistic judgments are shown in Figure 1.
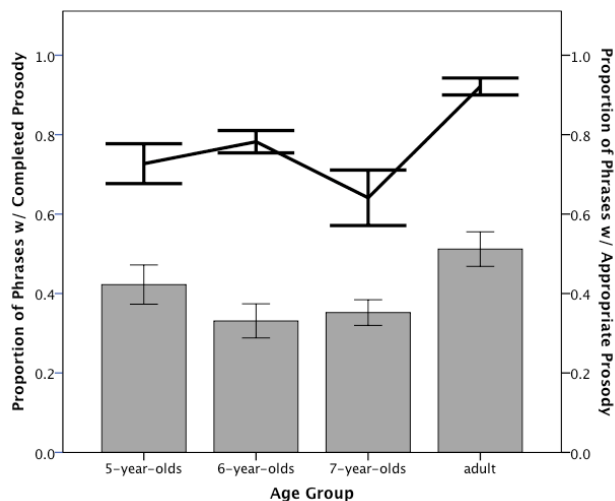
Figure 1: Proportion of phrases marked for completion (bar chart) and the appropriateness with which phrases were marked (line graph) shown for child and adult narratives.

It is evident in Figure 1 that adults produced a higher proportion of phrases marked for completion than children did (child $M = .37$; adult $M = .51$) in keeping with [10], though this result did not hold for all age groups. Post-hoc comparisons of 5-year-old and adult speech show no significant differences. 5-year-old speech was also not different from 6- and 7-year-old speech. Figure 1 also shows that adult transitional prosody was deemed more appropriate given the semantic/discourse context than child transitional prosody (child $M = .72$; adult $M = .92$).

In sum, the analyses indicate two age-related differences in how transitional prosody is used. The finding that adults produce higher proportion of phrases marked for completion than children may reflect age-related differences in narrative structure. In contrast, the finding that children use transitional prosody less appropriately than adults suggests that it is acquired independently from the morphosyntactic devices used to mark discourse structure. We will investigate the types of morphosyntactic devices that the children used in follow up work in order to more directly assess the notion of independence.

3.3. Marking Continuation versus Completion

Repeated measures ANOVA tests, with age group as a between-subjects factor and functional marking as the within-subjects factor, were used to test for differences in the relative frequency of phrase-initial high versus low tones and phrase-final low versus high tones, the melodic continuity between phrases, phrase-final pitch accents, phrase-final prominences, post-boundary pausing, and mean pause duration. The analyses revealed several differences between the age groups and several robust prosodic cues to continuation and completion.

The effect of age group was significant for the proportional number of phrases produced with final prominences [$F(3,52) = 8.25$, $p < .001$] and post-phrasal pauses [$F(3,52) = 8.61$, $p < .001$], and for mean pause duration [$F(3,49) = 3.33$, $p = .027$]. As before, children patterned together and adults differed. Children produced a higher proportion of their phrases with final prominences compared to adults (child $M = .83$; adult $M = .72$). They also delimited a higher proportion of their phrases with subsequent pauses (child $M = .72$; adult $M = .54$), and produced

longer pause durations than adults (child $M = 523$ msec.; adult $M = 303$ msec.). The age-related differences in the frequency and duration of pausing are consistent with results from prior work [11]. To the best of our knowledge, the effect of age on phrase-final prominences is new and deserves further exploration.

The effect of functional marking was significant for the proportional number of phrases with final low tones [$F(3,52) = 8.25$, $p < .001$], final pitch accents [$F(3,52) = 8.25$, $p < .001$], post-phrasal pauses [$F(1,52) = 46.15$, $p < .001$], and mean pause duration [$F(1,49) = 13.39$, $p = .001$]. As expected, final low tones and pausing were more often associated with completion marking than with continuation marking, and pauses were shorter when they occurred after a phrase marked for continuation than when they occurred after a phrase marked for completion. In contrast, final pitch accents were more commonly associated with continuation marking than with completion marking.

The overall analyses revealed no significant interactions between age group and continuation or completion marking. Nonetheless, an age-related difference in the distribution of pauses across the prosodic categories was suggested in the data (see Figure 2). In particular, children were more likely than adults to pause after a phrase marked for continuation. When the data were split by category this difference was significant [$F(3,52) = 6.86$, $p = .001$]. No similar difference was found for phrases marked as completed.
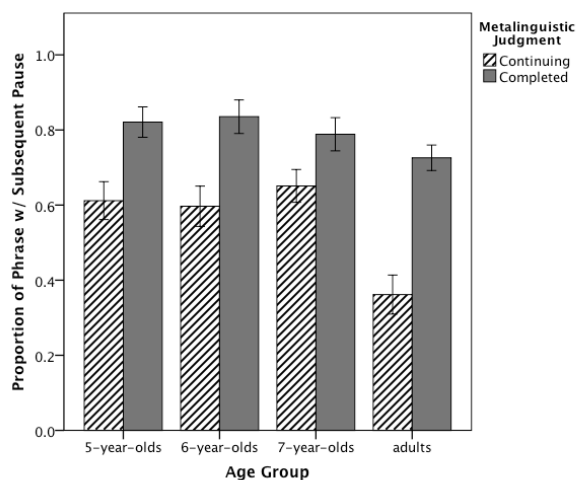

Figure 2: The relative frequency of pauses after child and adult phrases marked for continuation versus completion.

When the effect of age on the other dependent measures was analyzed separately for each prosodic category, we found a trend for pause durations to vary with age group after phrases marked for completion [$F(3,51) = 2.59$, $p = .063$], but not after those marked for continuation. Specifically, 7-year-old children produced especially long pauses after prosodically completed phrases. This result, shown in Figure 3, is noted because it nicely complements the results on pause frequency.

In contrast to pausing, the results from the split data analyses were as before for final low tones and final pitch accents—these varied systematically with prosodic marking, but in the same way for child and adult speech. Figure 4 exemplifies these results.
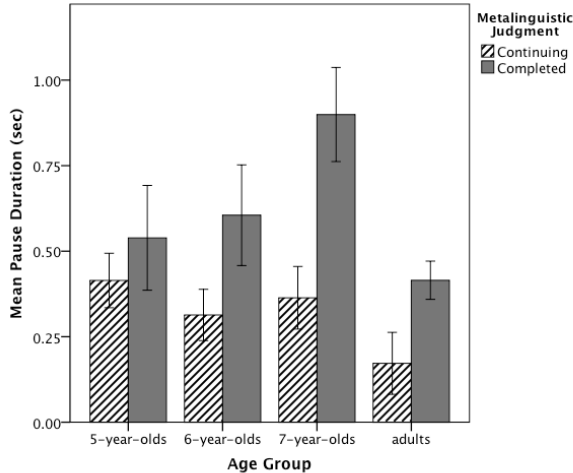
Figure 3: The mean duration of pauses after child and adult phrases marked for continuation versus completion.
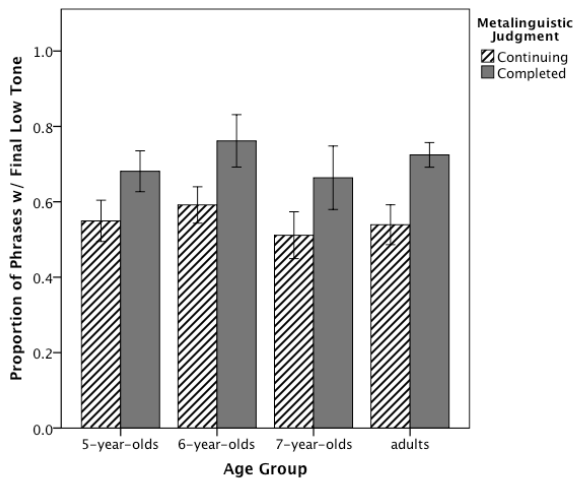

Figure 4. The relative frequency of phrase-final low tones in child and adult phrases marked for continuation versus completion.

On the whole, children and adults appeared to use the same prosodic cues to mark continuation versus completion. These cues were confined to phrase-final prosody, contra some indications in the literature regarding the functional importance of initial contours (see [2]). The absence of a significant interaction between age and functional category for any of the dependent measures in the overall analyses suggests that the key features of transitional prosody are acquired fairly early. This broad conclusion is somewhat tempered by results from additional analyses, which suggest that the pause-related cues to continuation and completion vary with age. Although this variability is likely due to the overall higher rate of pauses to speech in children's narratives compared to adults', the use of contrastive durations, especially in 7-year-old speech, could underscore the importance of pausing as a cue to completion.

4. Conclusion

The current study provides some evidence that transitional prosody is acquired separately from the morphosyntactic and/or lexicogrammatical aspects of discourse. The study also shows that continuation and completion is marked in the same way in child and adult speech with a combination of phrase-final accents, tones, and pauses. These results are in agreement with the view that phrase-final prosody defines the relationship between intonational phrases [3]. Finally, the results underscore the importance of pausing as a cue to completion, but also suggest that the strength of this cue interacts with the higher rate of pausing in child compared to adult speech.

5. Acknowledgements

6. References

[1] T. Givón, "Topic continuity in spoken English," in *Topic Continuity in Discourse*, T. Givón, Ed. Amsterdam: Benjamins, 1983, pp. 343-364.

[2] E. Couper-Kuhlen, "Intonation and discourse: current views from within," in *The Handbook of Discourse Analysis*, D. Schiffrin et al., Eds. Malden, MA: Blackwell, 2001, pp. 13-34.

[3] J.W. Du Bois et al., "Outline of discourse transcription," in *Talking Data: Transcription and Coding in Discourse Research*, J. A. Edwards and M. D. Lampert, Eds. Hillsdale, NJ: Erlbaum, 1993, pp. 45-89.

[4] W. Chafe, "Linking intonation units in spoken English," in *Clause Combining in Grammar and Discourse*, J. Haiman and S. A. Thompson, Eds. Amsterdam: Benjamins, 1988, pp. 1-27.

[5] S. Shattuck-Hufnagel and A. Turk, "A prosody tutorial for investigators of auditory sentence processing," *J. of Psycholinguistic Research*, 25(2), 193-247.

[6] M. Swerts, "Prosodic features at discourse boundaries of different strength," *J. of the Acoust. Soc. of Amer.,* 101(1), 514-521, 1997.

[7] S. Schuetze-Coburn et al., "Units of intonation in discourse: A comparison of acoustic and auditory analyses," *Language and Speech*, 34(3), 207–234, 1991.

[8] M. Swerts and R. Geluyken, "Prosody as a marker of information flow in spoken discourse," *Language and Speech*, 37, 21–43, 1994.

[9] B. Wells et al., "Intonation development from five to thirteen," *J. of Child Lang.*, 31, 749-78, 2004.

[10] M. Vion and A. Colas, "Intentional control and operational constraints in prosodic phrasing: a study of picture-elicited narrations by French children," *Language and Speech*, 52, 481-513, 2001.

[11] M.A. Redford, "A comparative analysis of pausing in child and adult storytelling," *Applied Psycholinguistics*, 1(1), 1-21, 2012.

[12] M. Breen et al., "Inter-transcriber agreement for two systems of prosodic annotation: ToBI (Tones and Break Indices) and RaP (Rhythm and Pitch)," *Corpus Linguistics and Linguistic Theory*, in press.

[13] K. Silverman et al., "*TOBI: A standard for labeling English prosody*," 2nd *International Conference on Spoken Language Processing*, Banff, Alberta, CA, 1992, pp. 867-870.

[14] L.C. Dilley et al., "Age-related changes to spectral voice characteristics affect judgments of prosodic, segmental, and talker attributes for child and adult speech," *J. of Speech, Language, and Hearing Research*, in press.