Speech Prosody 2016
31 May - 3 Jun 2106, Boston, USA

Acquisition of prosody: The role of variability*

Angeliki Athanasopoulou[1], *Irene Vogel*[1]

[1] University of Delaware

angeliki@udel.edu, ivogel@udel.edu

Abstract

Although some phonetic variability is inevitable in speech production, adult speech is fairly consistent. Thus, part of becoming a competent adult speaker is learning to appropriately limit the variability in one's speech. It is generally believed that phonology is mastered relatively early; however, this does not take into account the refinement of articulation required to reign in the variability in production. In this paper, we investigate the development of the acoustic properties (i.e., F0, Duration, and Intensity) of two prosodic patterns of English, compound and phrasal prominence, as well as their patterns of variability, in the speech of 6-, 8-, and 11-year-olds. While the 11-year-olds show adult-like acoustic patterns in their compound prosody, no children are adult-like in phrasal prominence. Examination of the variability at the different ages shows, moreover, that even the 11-year-olds have not yet refined their speech to the same extent as adults. Thus, examination of these more subtle aspects of phonological acquisition demonstrates that the process continues for much longer than is typically assumed, and usually studied.

Index Terms: prosody, language acquisition, variability

1. Introduction

Prosody provides initial cues to infants acquiring language [1, 2, 3]. Indeed, very young infants are sensitive to certain prosodic patterns (e.g., the difference between the rhythmic patterns of their language vs. others) [4, 5, 6, 7], suggesting that the prosodic system of a language would be acquired quite early. Nevertheless, at least some aspects of prosodic acquisition seem to appear relatively late (e.g., while children produce word stress at 2 years [8], they only distinguish compounds vs. phrases in English around 11 years [9, 10]. Different methodologies across studies, however, make comparisons difficult, and overall there are generally fewer production than perception studies. Moreover, most production studies rely on adult listener judgments to determine prosodic acquisition, but such impressionistic data cannot provide information about the subtler acoustic aspects of prosody production, including the extent of variability present in speech at different ages. Although reduction of variability is recognized as an issue in language acquisition [11, 12], little is known about it, and if speech is like other behaviors, it could take until about age 16 for adult variability patterns to be achieved, well beyond the age usually studied in phonology acquisition research. In this paper, we examine the acquisition and variability of the acoustic properties of phrasal and compound prosody in English in the speech of 6-, 8-, and 11-year-olds in comparison with adults. We interpret the acoustic patterns and our finding that children gradually exhibit less variability with age, but at 11 years still have more variability than adults, as evidence that phonological acquisition, or at least the acquisition of prosody, extends to the early teen years. It is thus not fully accurate to claim that phonology is acquired by about the time children start school.

* The study was partially supported by NSF DDIG 1349032 to the authors.

2. Background

2.1. Phonological Variability and Acquisition

In studies of phonological acquisition that focus on contrasts between pairs of segments, it is considered that the sounds have been learned if a child can distinguish between them in perception (e.g., IPLP [13]), and possibly also production. Another indication that a child has learned some aspect of phonology, or other linguistic phenomenon, is that s/he appears to behave in the same way as adults. Both of these considerations, often based on auditory assessments of child speech, tend to lead to the belief that phonological acquisition is complete at a relatively young age.

Acoustic measurements provide more subtle information about the similarities between child and adult productions, but means calculated for groups of speakers tell only part of the story. That is, while they show overall patterns, they also give the impression of uniform behavior within the groups. There is, however, always some variability in speech, although as competent speakers, adults tend to be fairly consistent in their productions (aside from sociolinguistic variability which is not dependent on articulatory mastery). By contrast, children have been found to exhibit more variability, controlling their consonantal articulatory gestures only around age 14-16 [11], and the acoustic properties of vowels (i.e., duration, F0 and formants) only around age 14-15 [12]. Thus, if language acquisition means becoming fully adult-like in linguistic behavior, the time-frame studied must extend to the teen years.

2.2. Acquisition of English Compounds and Phrases

It is well-known that most English compounds differ from phrases in their stress patterns, with the former stressed on the first element and the latter on the final element (e.g., *hot-dog* = food vs. *hot dog* = hot canine). This refers to additional prominence since in both cases, each word has a primary stress manifested mainly by f0, duration and intensity [14, 15, 16].

There is some contradiction in the literature about the age of acquisition of the compound/phrase stress distinction. Some production studies [17] have reported that children as young as 2 years are able to produce compound prosody, however, more

 doi: 10.21437/SpeechProsody.2016-147

recent studies have shown that 6-year-olds do not produce compound prosody correctly, and that they are able to differentiate between compound and phrase minimal pairs only around age 11 [10, 9]. Based on these studies, it is difficult to determine at what age the relevant prosody is actually acquired since the stimuli often have a confound with focus on one of the elements [17], and more problematically, the findings typically rely on adult listener judgments which provide global assessments, but not precise information about the acoustic properties of the different constructions. Thus, while the exact age of acquisition is unclear, what does seem clear is that at least some aspects of prosodic acquisition take longer to master than might be expected.

3. Present Study

3.1. Hypotheses

The goals of this study of the acquisition of compound and phrasal prosody are two-fold: to identify developmental patterns in their acoustic properties (f0, duration, intensity), and to determine whether variability in these properties shows gradual reduction with age, extending at least until 11 years. We thus formulate the following hypotheses:

Hypothesis 1: Children exhibit adult-like prosody in compounds and phrases prosody at 11 years, but not earlier.

Hypothesis 2: The variability of compound and phrasal prosodic properties is reduced as children get older, but even 11-year-olds may have greater variability than adults.

3.2. Participants

The participants included three groups of children: 10 six-year-olds (mean = 6 years; 7 months), 8 eight-year-olds (mean = 8 years; 10 months), and 13 eleven-year-olds (mean = 11 years; 1 month). In addition, there were 8 adult speakers of a standard north-east variety of American English (mean age = 25 years). All participants are from mid-SES families, and reported having no language, hearing or cognitive problems.

3.3. Stimuli

In order to compare their prosody, novel compounds and phrases were elicited with the same target elements, 10 mono-syllabic verb roots suffixed with either *–ing* or *–er*, appearing as the first and second words in both types of structure. The stimuli were further restricted to CVC roots where V is either a tense vowel or diphthong, as illustrated in Table 1, where C1 and C2 indicate the first and second word targets of compounds, and P1 and P2 the corresponding items in phrases.

Table 1: *Examples of Stimuli. Targets are underlined.*

Category	Structure	Position	N	Example
C1	Compound	Word 1	10	b**a**king shoes
C2	Compound	Word 2	10	muffin b**a**ker
P1	Phrase	Word 1	10	b**a**king slowly
P2	Phrase	Word 2	10	the b**a**ker

3.4. Procedure

Experiments designed as games were developed so that the desired compounds and phrases would be uttered as naturally as possible even by the youngest children. A separate game corresponded to each of the categories in Table 1 (in total there were 4 games). In all cases, the participant had to help an experimenter complete a poster by placing a particular picture in the correct location on the poster – corresponding to the view the participant saw on a computer screen (e.g., "*Put the baking shoes in the ocean.*"). The target compounds and phrases always appeared sentence medially to avoid possible sentence-final prosodic confounds. Moreover, before producing the desired sentences, the participants had to name the targets, making them "old information" and thus reducing the possibility of confounding prosodic properties due to focus. Prior to beginning the actual experiments, the participants were instructed with two familiarization items and four training items (2 with familiar compounds/phrases; 2 with novel ones).

4. Analysis

4.1. Phonetic Analysis

Praat was used to segment the nuclei of the target syllables of the verb roots and to measure their Duration, mean Intensity, mean F0, ΔF0 (F0 of last quarter minus F0 of first quarter of vowel), and vowel centralization. All phonetic measurements were normalized for each speaker by converting them to z-scores.

4.2. Statistical Analysis

4.2.1. Acoustic Properties

In order to compare the properties of compounds and phrases among the age groups, a Mixed-Design MANOVA was conducted, with dependent variables: Duration, Intensity, F0, ΔF0, and Centralization. The between-groups independent variable was Age Group (adults, 6-, 8-, and 11-year-olds) and the within-groups independent variables were Structure (Phrase and Compound) and Position (Word 1 and Word 2).

4.2.2. Variability

Variability was determined for each acoustic measure in each structure using the ratio of word 1 and word 2 (W2/W1), based on their standard deviations. A Mixed-Design MANOVA was used again, but with the dependent variables being the standard deviation of the acoustic properties, and the only within subjects variable being Structure.

5. Results

We focus here on the three main phonetic properties of stress, (mean) F0, Duration and Intensity.

5.1. Phonetic Properties of Compounds and Phrases

The mixed-design MANOVA showed a significant interaction between Age Group, Structure, and Position: $F(15,759) = 3.362$, $p < .001$. The interaction was further explored with ANOVAs and paired t-tests; the results are summarized below.

5.1.1. Adult Baseline

We first present the adult production patterns since we take these as our baseline. In the **compounds,** we only found a significant drop in intensity from C1 to C2 ($t(78) = 4.37$, $p < .001$). By contrast, in the **phrases** we observed: a) lengthening of P2 compared to P1 ($t(77) = -5.49$, $p < .001$), b) lowering of F0 from P1 to P2 ($t(75) = 2.19$, $p < .05$) and c) decrease in

Intensity from P1 to P2 ($t(77) = 3.83$, $p < .001$), as shown in Figure 1. Error bars, here and below, show standard error.

In comparing the two structures, it thus appears that adults differentiate compounds from phrases by producing the former with greater Intensity on the first word, combined with similar duration and pitch properties on both words; the latter have longer Duration and lowered F0 and Intensity on the second word, consistent with the typical English phrasal pattern.

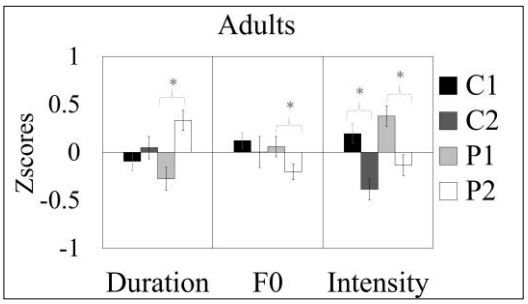

Figure 1: *Adult production: Duration, F0, Intensity.*

We next examine the children's productions to determine to what extent the different age groups show the adult patterns.

5.1.2. Duration

Figure 2 provides the Duration results for the children and the adults. Comparing first the youngest children with the adults, we see that the 6-year-olds have the opposite pattern from adults. While for adults Duration is very similar in C1 and C2, for the 6-year-olds C2 is longer than C1 ($t(91) = 2.38$, $p < .05$). By contrast, the significant lengthening in P2 in the adults is absent in the children ($t(64) = 0.14$, $p = .9$), and in fact, although not significant, the Duration relationship between the two words tends in the direction opposite that of the adults.

There is little difference between the 6- and 8-year-olds (C1 vs. C2, $t(91) = 2.38$, $p < .05$); however, in the latter, the Duration relationship between P1 and P2 tends in the same direction as adults, with a longer P2, though not significantly so ($t(70) = -0.58$, $p = .6$). The 11-year-olds show essentially the same Duration pattern as adults in compounds (i.e., no distinction, $t(126) = -1$, $p = .3$). Moreover, although it is not significant ($t(122) = -1.93$, $p = .056$), the relationship between P1 and P2 is also approaching the pattern seen in adults.

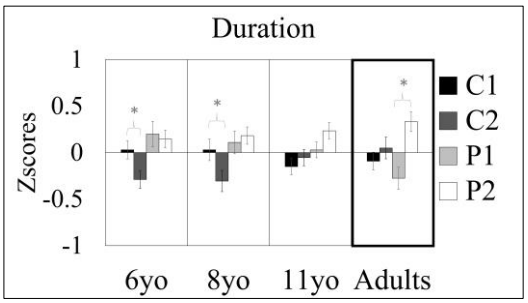

Figure 2: *Duration patterns of each age group.*

5.1.3. F0

The F0 results are shown in Figure 3. Beginning with the youngest children, we see that, like the adults, they do not have a difference in F0 between C1 and C2 ($t(89) = 0.31$, $p = .8$). Also like adults, they exhibit a significant F0 difference between P1 and P2 ($t(61) = -4.48$, $p < .001$). The relationship between the two words, however, is the opposite of that found in adults, with P2 being shorter rather than longer than P1.

Again, the 8-year-olds show essentially the same pattern as the youngest group (Comp: $t(77) = 0.15$, $p = .9$; Phr: $t(69) = -2.69$, $p < 0.01$). The 11-year-olds, while still not exhibiting adult behavior in both structures, show progress. Although they do not exhibit significant differences between the two words of either compounds ($t(116) = 0.11$, $p = .9$; like adults) or phrases ($t(120) = -0.94$, $p = .4$), the fact that the incorrect phrasal pattern used by the 6- and 8-year-olds has been suppressed indicates a development in the right direction.

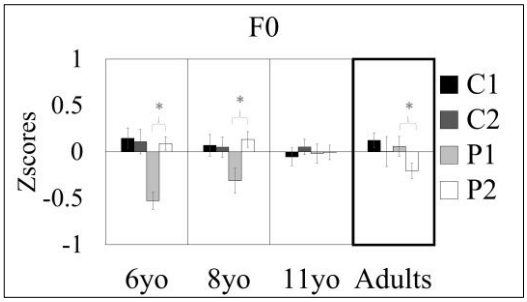

Figure 3: *F0 patterns of each age group.*

5.1.4. Intensity

Figure 4 shows the results for Intensity. All three groups of children use the same adult-like pattern in compounds: greater Intensity in C1 than C2 (6-year-olds: $t(91) = 2.89$, $p < .005$; 8-year-olds: $t(78) = 4.33$, $p < .001$; 11-year-olds: $t(126) = 6.59$, $p < .001$). With regard to phrases, however, we see that the 6-year-olds exhibit higher Intensity in P2 than P1 ($t(64) = -4.06$, $p < .001$), the opposite of the adult pattern.

Unlike with Duration and F0, where the 8-year-olds pattern with the youngest group, when it comes to Intensity, it is the 8- and 11-year-olds who exhibit the same patterns. They still differ from the adults, however, in showing similar Intensity on P1 and P2 (8-year-olds: $t(70) = 0.31$, $p = .76$; 11-year-olds: $t(122) = 0.69$, $p = .5$), rather than a decrease on P2. Again, although they are not yet adult-like in phrasal prosody, 8- and 11-year-olds have progressed beyond the first group by suppressing the incorrect pattern with greater Intensity on P2.

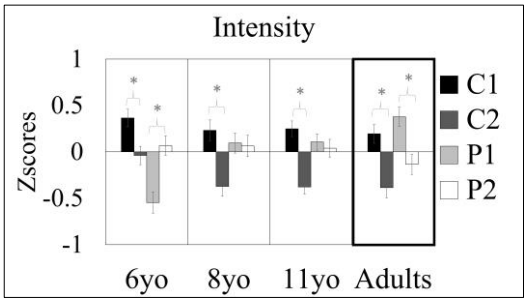

Figure 4: *Intensity patterns of each age group.*

5.2. Variability

Taking the adult patterns as a baseline again, we next consider the variability present in the use of the three acoustic properties. The asterisks indicate a significant difference between the child and adult variability measures for the W2/W1 ratios in the compounds and phrases. The results from

a mixed-design MANOVA showed a significant effect of Age Group, $F(4,33) = 6.96$, $p < .05$ which was explored with further tests (ANOVAs and T-tests), as discussed below.

With regard to Duration, as can be seen in Figure 5, all the groups of children show greater variability in the compounds than adults (6-year-olds: $t(16) = -2.34$, $p < .05$; 8-year-olds: $t(14) = -3.23$, $p < .01$; 11-year-olds: $t(19) = -2.88$, $p < .01$). Even though no statistical difference was found in phrases, the trend is the same, greater variability in children.

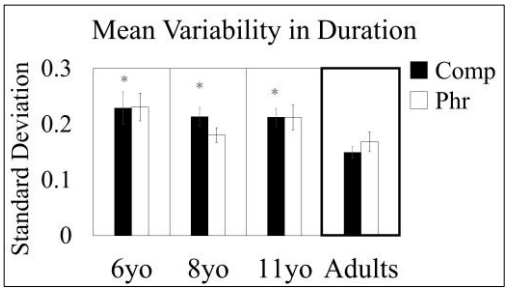

Figure 5: *Mean Variability in Duration for all ages.*

With regard to F0, no statistically significant differences were found, however, overall the trend indicates somewhat more variability in the children, except for the unusually high variability in the adult compounds. Since this may be due to the use of creaky voice, HNR [18] was measured and it showed that for the adults, C2 was more creaky than C1 leading to greater variability in the C2/C1 ratio of F0. As mentioned above, however, it is difficult to interpret this in relation to acquisition since such sociolinguistic variation does not fall within the realm of variation due to mastery of articulatory control.

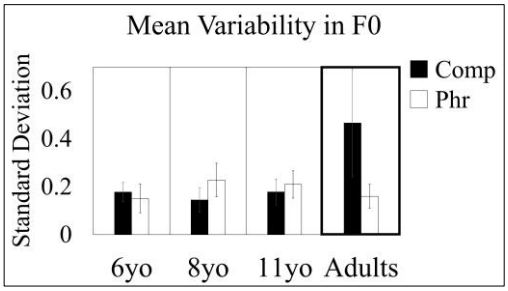

Figure 6: *Mean Variability in F0 for all ages.*

Finally, Figure 7 shows that the variability in Intensity is stable across the age groups and word positions.

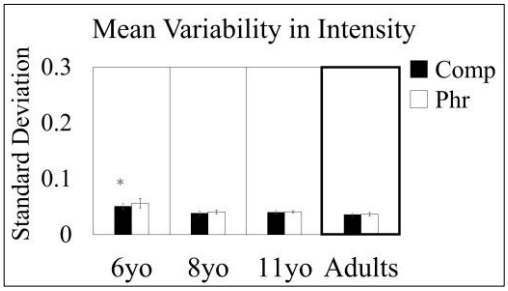

Figure 7: *Mean Variability in Intensity for all ages.*

Only the youngest children exhibit more variability than the adults in the compounds ($t(16) = -2.2$, $p < .05$), and the same trend in phrases ($t(15) = -2.1$, $p = .053$). The variability differences are very small, however, and thus most likely not meaningful.

6. Discussion

Our investigation focusing on the acoustic properties of Duration, F0 and Intensity in compounds and phrases revealed a developmental progression between 6 and 11 years; however, none of the properties was completely acquired even by age 11, if we consider acquisition to entail adult-like use of a given property in both compounds and phrases. While the 6-year-olds' behavior is the most distinct from the adults, and the 11-year-olds' behavior the most similar, the 8-year-olds pattern with either of the other age groups depending on the property. Thus, Hypothesis 1 is partially confirmed, since the oldest participants show adult-like use of some of the acoustic properties in both constructions, but not all of them.

The variability results also indicate that adult-like behavior is not fully achieved by age 11. While F0 variability appears to be fairly adult-like even at age 6, this property may involve additional sociolinguistic factors that are not relevant here. Duration shows a gradual reduction in variability with age, but still substantially greater variability in the 11-year-olds than the adults. The variability in Intensity is minimal across the age groups. Thus, Hypothesis 2 is confirmed since, where there are differences between the child and adult behaviors, the variability is greater at the younger ages.

7. Conclusions

Taken together, our acoustic and variability analyses provide insight into the complex process of acquiring prosody. While all participants in some way distinguish their compounds and phrases, possibly explaining claims based on impressionistic assessment that compound prosody is acquired early, examination of the acoustic properties shows that even at age 11 they do not match those of adults. In fact, this would account for other observations that seemed to conflict with the early acquisition claims – the fact that 11-year-olds are not able to perceptually distinguish compound vs. phrasal stress like adults [9, 10]. That is, while children make some difference between the structures, since it is not the same as that made by adults, they do not yet consistently map the adult acoustic patterns to the relevant morpho-syntactic structures. The fact that in general the compound properties are mastered earlier than the phrase patterns, moreover, suggests that the phonology of smaller (compound word) constituents is acquired earlier than that of larger (phrasal) constituents.

What both analyses show, furthermore, is that the process of phonological acquisition reaches into the early teen years when fine-tuning of acoustic properties is concerned, as has been found for some segments as well. The fact that we found minimal variability in Intensity from the youngest age is consistent with the fact that the fine articulatory control needed for Duration, and possibly F0, is not a limiting factor in the mastery of Intensity patterns.

8. References

[1] J. Morgan and K. Demuth, Signal to Syntax: Bootstrapping From Speech To Grammar in Early Acquisition, Mahwah, NJ: Lawrence Erlbaum Associates, 1996.

[2] M. van Heugten, I. Dautriche and A. Christophe, "Phonological and prosodic bootstrapping," in *Encyclopedia of Language Development*, Thousand Oaks, CA, SAGE Publications Inc, 2014, pp. 447-451.

[3] B. Höhle, "Bootstrapping mechanisms in first language acquisition," *Linguistics*, vol. 47, no. 2, pp. 359-382, 2009.

[4] F. Ramus, "Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues," *Annual Review of Language Acquisition*, vol. 2, pp. 85-115, 2002.

[5] J. Mehler and A. Christophe, "Maturation and learning of language during the first year of life," in *The Cognitive Neurosciences*, Bradford Books / MIT Press, 1995, pp. 943-954.

[6] T. Nazzi, J. Bertoncini and J. Mehler, "Language discrimination by newborns: towards an understanding of the role of rhythm," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 24, no. 3, pp. 756-766, 1998.

[7] J. Mehler, P. Jusczyk, G. Lambertz, N. Halsted, J. Bertoncini and C. Amiel-Tison, "A precursor of language acquisition in young infants," *Cognition*, vol. 29, pp. 143-178, 1988.

[8] M. Kehoe, C. Stoel-Gammon and E. H. Buder, "Acoustic correlates of stress in young children's speech," *Journal of Speech and Hearing Research*, vol. 38, no. 2, pp. 338-350, 1995.

[9] I. Vogel and E. Raimy, "The acquisition of compound vs. phrasal stress: The role of prosodic constituents," *Journal of Child Language*, vol. 29, p. 225–250, 2002.

[10] H. Shilling, Compound and phrasal stress acquisition: When a greenhouse becomes different to a green house, University of Birmingham: Unpublished doctoral dissertation, 2010.

[11] A. Smith, "Speech motor development: Integrating muscles, movements, and linguistic units," *Journal of Communication Disorders*, vol. 39, p. 331–349, 2006.

[12] S. Lee, A. Potamianos and S. Narayanan, "Analysis of children's speech: Duration, pitch and formants," in *EUROSPEECH*, 1997.

[13] R. M. Golinkoff, W. Ma, L. Song and K. Hirsh-Pasek, "Twenty-Five Years Using the Intermodal Preferential Looking Paradigm to Study Language Acquisition: What Have We Learned?," *Perspectives on Psychological Science* , vol. 8, no. 3, pp. 316-339, 2013.

[14] E. Farnetani, C. Torsello and P. Cosi, "English compound versus non-compound noun phrases in discourse: An acoustic and perceptual study," *Language and Speech*, vol. 32, no. 2, p. 157–180, 1988.

[15] I. Plag, "The variability of compound stress in English: Structural, semantic and analogical factors," *English Language and Linguistics*, vol. 10, p. 143–172, 2006.

[16] P. Štekauer, J. Zimmermann and R. Gregová, "Stress in Compounds: An Experimental Research," *Acta Linguistica Hungarica*, vol. 54, no. 2, pp. 193-215, 2007.

[17] E. V. Clark, S. A. Gelman and N. M. Lane, "Compound nouns and category structure in young children," *Child Development*, vol. 56, pp. 84-94, 1985.

[18] Y.-L. Shue, P. Keating, C. Vicenik, and K. Yu, "VoiceSauce: A program for voice analysis," *Proceedings of the ICPhS XVII*, pp. 1846-1849, 2011.