

Running head: LEXICAL AND PHRASAL PROMINENCES IN CHILDREN

Lexical and Phrasal Prominence Patterns in School-Aged Children's Speech

Irina A. Shport & Melissa A. Redford

University of Oregon

Acknowledgments

This research was supported by Award Number R01HD061458 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development (NICHD). The content is solely the responsibility of the authors and does not necessarily reflect the views of NICHD. We are grateful to the 2009-2011 Speech and Language lab testing team for their help with data collection.

Abstract

This study investigated the integration of word- and phrase-level prominences in speech produced by 25 school-aged children (6;2 to 7;3) and 25 adults. Participants produced disyllabic number words in a straight count condition and in two phrasal conditions, namely, a stress clash and non-clash phrasal context. Duration and amplitude measures of syllable rhymes were used to assess the realization of lexical stress, and fundamental frequency (F0) measures were used to assess the realization of phrasal pitch accents across conditions. Results showed that the duration and F0 correlates varied independently of each other as a function of condition in child speech, but much less so in adult speech. The group differences were taken to indicate that 6-year-old children have yet to develop prosodic structures with integrated prominence. Structural and pragmatic interpretations of the results are discussed.

Introduction

Prominences encode the hierarchical relationships between prosodic units and signal the edges of phrases (*cf.* Beckman, Hirschberg, & Shattuck-Hufnagel, 2005; Cruttenden, 1986; Pierrehumbert, 1980). Prominences at the word level are due to lexical stress, and are conveyed via duration and amplitude changes (Fry, 1955; Kochanski, Grabe, Coleman, & Rosner, 2005; Lehiste & Fox, 1990; Mo, 2008; Turk & Sawusch, 1996). Vowel quality and a high fundamental frequency (F0) are also frequently associated with lexical stress, but these correlates vary across lexical items and prosodic environment (Beckman & Edwards, 1994; Huss, 1978). Prominences at the phrase level are due to default or context-dependent pitch accenting, and are conveyed by F0 peaks or valleys (Beckman, 1986; Cooper, Eady, & Mueller, 1985; Fry, 1958; Mo, 2009).

In adult speech, phrasal pitch accents land on stressed syllables (Bolinger, 1961; Hayes, 1984, 1995; Shattuck-Hufnagel, Ostendorf, & Ross, 1994; Vanderslice & Ladefoged, 1972) giving rise to prominence integration. The alignment of prominences at the phonological level is conveyed by the temporal alignment of acoustic features associated with lexical stress and pitch accenting. Prominence integration is especially striking under conditions where lexical stress is shifted from a default position, for example, in a stress clash context (Lieberman & Prince, 1977; Hayes, 1984, 1995). The stress clash context allows for the possibility of misalignments between lexical stress and pitch accent, which underscores the notion that word- and phrase-level prominences must be actively integrated at some level in phonological or phonetic planning. The current study focused on the questions of how word- and phrase-level prominences are organized in the language of school-aged children and whether the integration of prominences is as robust in child as in adult language.

Integrated Word- and Phrase-Level Prominences in Adult Language

Phrases where lexical stress shift is likely to occur may be used to examine how tightly word- and phrase-level prominences are integrated, because this stress shift may or may not be accompanied by a shift in the location of pitch accent. Stress shift reportedly occurs in clash contexts, that is, in contexts where word sequencing results in adjacent, lexically-stressed syllables. Consider, for instance, Liberman and Prince's (1977) classic example of leftward stress shift in the phrase *thirteen men*, which is represented in Figure 1 using metrical grid notation. In the grid, 'x's mark each syllable at the lowest level, then lexical stress at the next level, and phrasal prominence at the highest level. The height of an 'x' bar represents cumulative prominence for each syllable. Cumulative prominences are reflected in the phonetic realm following the general assumption, well-articulated by Cho and Keating (2009), that "phonological categories of prominence can be translated into a single prominence scaling, and that perceivers can differentiate varying degrees of prominence along such a scale" (p. 468). According to this assumption, then, the highest 'x' bar in the grid is perceived as the strongest prominence, and the strongest prominence is in turn a consequence of prominence integration, that is, the temporal alignment of lexical stress and phrasal accent.

Figure 1 approximately here

The arrow in panel A of Figure 1 shows only a word-level shift in prominence. In Metrical Stress Theory (Liberman & Prince, 1977), phrasal prominence (main stress) does not move. Rather, stress shift is viewed as a word-level phenomenon driven by a clash-avoidance strategy and by language-specific preferences; in this case, an English preference for a trochaic pattern. This view of stress shift can be contrasted with the approach taken in Intonational-Metrical Theory (Bolinger, 1986; Gussenhoven, 1991; Shattuck-Hufnagel et al., 1994). In this

theory, lexical stress shift at the word level is conditioned by phrase-level patterns as well as by language-specific preferences for some metrical patterns over others. The phrase-level patterns in English follow from a “template, with strong connections to rhythm, whereby a sort of annunciatory or attention-getting accent comes toward the beginning (of a phrase) and a ‘punch’ accent comes toward the end (Bolinger 1985: 85).” Panel B in Figure 1 shows how shifts in prominence at both word- and phrase-levels contribute to the accumulation of prominence under conditions of stress clash in Intonational-Metrical Theory.

Supporting evidence for the Intonational-Metrical Theory comes from the observation that stress shift is dependent on phrase position (Grabe & Warren, 1995; Shattuck-Hufnagel, 1992; Shattuck-Hufnagel et al., 1994). For example, a leftward stress shift will occur in the phrase *thirteen men* when it is in phrase-initial position (e.g., *Thirteen mónkeys escaped from the zoo this morning*), but not when it is in phrase-final position (e.g., *Media reports focused on the health of those thirtéen mónkeys*). In fact, the preference for a phrase-initial accent is so strong, that stress shift may occur even in the absence of clash (e.g., *Thirteen gorillas escaped from the zoo this morning*) (Shattuck-Hufnagel et al., 1994).

Word- and Phrase-level Prominence Patterns in Child Language

Previous research has established that English-speaking children realize lexical stress by age two, and control trochaic (strong-weak) patterns earlier than iambic (weak-strong) patterns (Allen & Hawkins, 1980; Gerken, 1991; Kehoe, Stoel-Gammon, & Buder, 1995; Schwartz, Petinou, Goffman, Lazowski, & Cartusciello, 1996), but very little is known about how or whether children subordinate word-level prominences in service of phrase-level patterns. Recent research suggests that children realize major pitch accents and boundary tones also by age two,

but the repertoire and accuracy of tonal targets are more adult-like in trochaic than iambic patterns. For example, Snow (2007) found that young children's production of adult-like falling and rising intonational patterns in polysyllabic utterances is more restricted in a weak-strong context compared to a strong-weak context. Similarly, Astruc, Prieto, Payne, Post, and Vanrell (2009) demonstrated that toddlers are less able to align pitch accent peaks and stressed syllables in words with ultimate stress compared to in words with penultimate or antepenultimate stress. These results could suggest that children's early preference for the trochaic pattern (i.e., at the one-word stage) affords them more practice with prominence integration in trochees than in iambs, and this then generalizes to short utterances.

Work with children between 5 and 10 years old suggests that phrase-level patterns of prominence integration in child language are immature well into middle childhood. Immature patterns include a strong bias towards an early accent in phrases such as *hot dóg*, perceived as compound stress (Allen & Hawkins, 1980; Atkinson-King, 1970; Vogel & Raimy, 2002); undifferentiated, and therefore unadult-like, prominence patterns in prosodic units with different morphosyntactic structures (Goffman, 2004); and difficulties incorporating two lexically stressed words into a single intonational phrase (Wells, Peppé, & Goulandris, 2004).

If immature phrase-level patterns in child language result in the poor integration of word- and phrase-level patterns for some structures, we might expect a mismatch between lexical stress placement and pitch accent placement. These immature phrase-level patterns may be most apparent under conditions that promote stress shift in that different contexts defined by word-level prominence patterns allow for the misalignment of word- and phrase-level prominences.

Current Study

The current study investigated child and adult realizations of lexical stress and pitch accents in clash and non-clash contexts in an effort to better understand the acquisition of integrated word- and phrase-level prominence marking. Although this study may be the first to examine both lexical stress and pitch accent in these contexts, it is not the first to examine children's sensitivities to clash and non-clash contexts. Goffman, Heisler, and Chakraborty (2006) were interested in the possibility that children may exhibit a greater rhythmic readjustment in stress clash and gap contexts than adults due to a stronger preference for the strong-weak patterns in English. Goffman and colleagues elicited sentences with stress clash (e.g., *Bób's púppup is falling*) and stress gap (e.g., *Bóbb'y's puppúp is falling*) from 4- to 7-year old children using puppet props and skits. Kinematic data showed that children reproduced the strong-weak and weak-strong alternations in the nonword *puppup* (i.e., ['pʌpəp] and [pə'pʌp]), but did not reorganize prominence patterns at the phrase level to avoid stress clash or stress gap contexts. Then again, neither did the adults in the study.

It turns out that stress shift is not always detected in production studies. This is because stress shift is probabilistic: the likelihood of its occurrence is influenced by several factors including metrical context (Hayes, 1995; Liberman & Prince, 1977; Quené & Port, 2002), phrase and word structures (Bolinger, 1961; Gussenhoven, 1991), word position in the phrase (Grabe & Warren, 1995; Shattuck-Hufnagel, 1992; Shattuck-Hufnagel et al., 1994), and speech rate (Quené & Port, 2002). These factors have not always been controlled in experiments that investigate the rhythmic organization of speech. For example, the stress clash context in Goffman et al.'s (2006) study used the phrase-initial word *Bob*, which contains only one syllable, thus making a leftward stress shift in this word impossible.

Moreover, stress shift is not a stress pattern reversal from weak-strong to strong-weak. Production studies that find stress shift in clash contexts, like *thirtéen mén*, indicate that stress is not so much shifted as equalized across a word (Cooper & Eady, 1986; Grabe & Warren, 1995; Horne, 1990; Shattuck-Hufnagel et al., 1994; Vogel, Bunnell, & Hoskins, 1995). That is, the lexically unstressed (e.g., *thir-*) and stressed (e.g., *-teen*) syllables in a word are produced with more similar duration and amplitude values in stress clash contexts than in non-clash contexts. At least for duration patterns, the neutralization of relative differences may be attributed to the shortening of stressed syllables in stress clash contexts rather than to the lengthening of unstressed syllables (Vogel et al., 1995).

The current study was undertaken with full knowledge that stress shift can be difficult to detect. For this reason, we used a counting task that controlled for metrical context, phrase structure, information structure, target word position in the phrase, and speech rate. The task was designed to encourage highly rhythmic speech (i.e., speech with regularly timed intervals), albeit more naturally than in studies that have used continuous repetition of a particular phrase (i.e., cycling) or metronome-timing of speech (Cummins & Port, 1998; Quené & Port, 2002). Metrical context (clash vs. non-clash) was manipulated in our counting task by using intervening nouns with different lexical stress patterns. We also elicited a straight count (e.g., *thirteen, fourteen, fifteen*, etc.) to provide a baseline for the assessment of stress shift and prominence integration. We assumed that word- and phrase-level prominences would be temporally aligned in the lexically stressed *-teen* syllable in the straight count condition. This assumption was experimentally validated in the present study.

With respect to prominence integration, pitch accent placement was expected to be in the number word in both phrasal contexts, and not on the noun, because in a counting task the

number represents new information. This expectation was confirmed in a perceptual judgment experiment. Briefly, 11 adult judges were asked to decide which word was emphasized in the adult and child productions of the *N-teen barbeque* and *N-teen banana* phrases. On average, 7 out of 11 judges perceived the number word as the most prominent element in these phrases. This result held for both phrasal contexts and for adult and child speech, thereby validating our assumption that number words would be accented in number + noun phrases.

Two sets of predictions were made with regard to stress shifting and prominence integration in phrasal contexts. The first set of predictions was made with regard to adult speech (1 and 2 below), and the second set with regard to child speech (3 and 4 below):

1. In a clash context, the phrase-level prominence was expected to be on the initial syllable due to an early accent bias in a phrasal context (see the discussion of Intonational-Metrical Theory above). Word-level prominence was predicted to shift due to the clash context and the pressure to integrate word- and phrase-level prominences.
2. In a non-clash context, adults were also predicted to align prominences because these prominences are integrated during planning. The location of cumulative prominence was predicted to be on the second syllable according to Metrical Stress Theory, but on the first syllable according to Intonational-Metrical Theory.
3. In a clash context, children were predicted to behave like adults insofar as the conditioning context also interacts with their presumed preference for trochaic stress patterns and an early accent bias.
4. In a non-clash context, children were predicted to misalign word- and phrase-level prominences if they faithfully preserve the lexical stress pattern and yet still implement an early accent bias.

To summarize, we examined the integration of stress and pitch accent in a phrase, which would entail the alignment of lexical stress and pitch accent on a single syllable in the number word. Acoustic correlates of lexical stress (duration and amplitude) and pitch accents (F0) were measured to determine the extent to which word- and phrase-level patterns shifted in the clash and non-clash conditions, as compared to a straight count condition.

Methods

Participants

Twenty-five children and 25 adults participated in the study. At the time of the study, children ranged in age from 6;2 to 7;3. The adults were all University of Oregon undergraduate students. All participants were monolingual native speakers of American English. Seventeen of the 25 participants in each group were female and the rest were male. Child participants were developing typically according to parental report. Children also had average to above average receptive vocabularies, as determined by the *Peabody Picture Vocabulary Test* (Dunn & Dunn, 2007), and normal hearing, as determined by a pure-tone hearing screen. Adult participants also reported normal hearing and language.

A separate set of 12 University of Oregon undergraduates provided auditory judgments of syllable prominence in number words.

Counting Task

We asked children and adults to count from 1 to 20 under three conditions: a straight count (no phrasal context), a clash context, and a non-clash context. A straight count was elicited to provide information about the default location and baseline realization of lexical stress and phrasal accent during counting (e.g., *thirtéen, fourtéén...*). The task used to create the clash and

non-clash context conditions was modeled after the playground counting method of ‘one Mississippi, two Mississippi, three Mississippi...’ (note the customary singular form of the noun). American children use this method to slow their counting pace in games such as hide-and-seek. In our task, the clash context condition was created by inserting the noun *barbeque* in the count (e.g., *thirtéen barbeque, fourtéen barbeque...*). The non-clash context condition was created by inserting the noun *banana* in the count (e.g., *thirtéen banana, fourtéen banana...*). The order of the clash and non-clash context conditions was counterbalanced across participants.

The nouns *barbeque* and *banana* were chosen because these trisyllabic nouns are well-matched in segmental and syllabic structure. We chose to use trisyllabic nouns rather than disyllabic nouns in order to avoid an effect of final phrase-boundary tones on the realization of word-level prominence patterns. The selected nouns were also highly familiar to 6- and 7-year-old children, and so were easily depicted (see the task description below). While it is likely that our child participants had more life-time practice producing the word *banana* than the word *barbeque*, the counting task ensured that by the time children produced the first target phrase (*thirtéen barbeque*), they had repeated the noun at least twelve times.

Participants performed the counting task in a quiet laboratory room in the presence of a tester. Whereas the task was simply demonstrated orally for adult participants, testers used a number line and a picture of either a *barbeque* or a *banana* to support their oral explanation of the task to children. The props were then used to pace children’s counting by moving the picture across the number line. Pacing ensured that children maintained a consistent rate and rhythm across all numbers. Pilot work had indicated that children would otherwise race through the counting task and become unintelligible in so doing. Pacing also allowed the tester to back up on the number line so that a child could regain their counting rhythm if a disfluency occurred. The

counting task was trivial for adult participants: all maintained a steady, natural speech rate; and their speech was highly intelligible and fluent. All participants' productions were digitally recorded for later acoustic analyses.

Prominence Judgment Task

In addition to the acoustic analyses, auditory judgments of syllable prominence in number words were collected to validate the use of the straight count condition as a baseline. Our manipulation of clash and non-clash contexts to examine phrasal prominence structures is only valid if *-teen* is the default prominent syllable in numbers produced during counting. The counting sequences were excised, blocked by group, and presented to 12 adult listeners. The listeners were asked to decide whether the first or the second syllable was emphasized in the number words. They responded by clicking one of three buttons displayed on a computer screen: '*N*-syllable,' '*-teen* syllable,' or 'Unsure'.

Acoustic Measurements

A total of 300 words and 600 two-word phrases were analyzed in Praat (Boersma & Weenink, 2011). These were the six disyllabic stress-shiftable number words (i.e., 13-16, 18-19) produced by every child and adult participant in a straight count, and in phrases with an intervening strong-weak noun (*bárbeque*) or an intervening weak-strong noun (*banána*). Productions of phrases that contained pauses or some other disfluency were excluded from further analysis. One child produced the majority of his phrases disfluently, and so the entirety of his data was excluded. Acoustic measurements were made on the sonorant rhymes of 289 words (adults, $N = 149$; children, $N = 140$) and 558 phrases (adults, $N = 295$; children, $N = 263$). Sonorant rhymes were chosen for measurement instead of nucleus vowels to avoid problems associated with separating vowels from sonorant codas (Grabe & Warren, 1995).

Word-Level Prominence Patterns: To investigate the realization of lexical stress, rhyme duration and the average root-mean-square (rms) amplitude were measured in both sonorant rhymes of the number words (e.g., r1 and r2 in *fifteen*, Figure 2) and in the first two sonorant rhymes of the noun words (r3 and r4 in *barbeque*, Figure 2). Duration and amplitude ratios were then computed for each word by dividing the duration or amplitude of the first rhyme by the second. In this way, we were able to capture relative duration and amplitude patterns across the same number word produced in a straight count and with intervening nouns.

Figure 2 approximately here

Phrase-Level Prominence Patterns: To investigate the realization of pitch accents, F0 was recorded at three temporal points in each rhyme: 10%, 50%, and 90% into the rhyme duration. F0 ratios were computed for the number words by dividing the midpoint F0 values of the first syllable by the second. In addition, F0 values were normalized across a number word by subtracting a grand mean of all F0 values in the word from each individual value ($F0_{i_norm} = F0_i - \bar{M}$, where \bar{M} is an average of six F0 values). These normalized F0 values were used for a comparison of pitch accent locations in the number words produced by children and adults. A high-low pitch accent was identified as an F0 peak followed by a significantly lower F0 value in the following syllable.

Analyses

All data were analyzed using mixed effects modeling with speaker and item (number words) as random factors. Analyses of noun words were conducted to examine the characteristics of lexical stress in words with unambiguous lexical stress patterns. These analyses included group (children, adults) and stress pattern (strong-weak, weak-strong) as fixed factors. The analyses of number words included group (children, adults) and condition (straight count, clash,

and non-clash) as fixed factors. An additional fixed factor of rhyme position (r_1 , r_2) was included in an analysis of normalized F0 values. When the effect of condition was significant, differences between the straight count and two phrasal contexts were further investigated in pairwise comparisons. These tests were corrected for multiple comparisons using the Bonferroni method.

Results

Prominence Patterns in Nouns and Straight Count Numbers

The first set of analyses confirmed the expected duration and amplitude patterns for strong-weak and weak-strong nouns (*bárbeque* vs. *banána*). Duration and amplitude ratios in the noun words were much higher for the strong-weak pattern than for the weak-strong pattern [duration ratio, $F(1, 531) = 473.52, p < .001$; amplitude ratio, $F(1, 532) = 135.87, p < .001$]. The effect of group was not significant, nor was there an interaction between group and stress pattern. These results are evident in Table 1, which provides the means and standard deviations of ratios for child and adult productions of *bárbeque* and *banána*.

Table 1 approximately here

Next, duration and amplitude patterns of straight count numbers were examined to test the assumption that the *-teen* syllable is lexically stressed in this control condition. Table 1 shows that the patterns conformed neither to the canonical strong-weak or weak-strong pattern of the context nouns. Importantly, the duration ratios in the straight count number words were less than 1, which shows that the *-teen rhyme* was longer on average than the first syllable, thus the weak-strong lexical pattern was maintained. On the other hand, the duration (but not amplitude) ratios in the straight count words were significantly different from those of the weak-strong nouns

[$F(1, 571.32) = 480.40, p < .001$]. This difference can probably be attributed to the difference in the segmental composition of sonorant rhymes in the first syllable of the *N-teen* words compared to the highly reduced schwa rhymes in the *banana* nouns. Note that it is exactly this property that makes the *N-teen* words ‘stress-shiftable.’

Native listener perceptual judgments provided further support for the assumption that word- and phrase-level prominences would align in the second syllable in the straight count condition. The averaged proportion of ‘*teen* syllable prominent’ judgments was .66 in this condition (children: 64%; adults: 68%), but dropped to .32 in the phrasal context conditions (children: 35%; adults: 29%). For each category of auditory judgments, the difference between the child and adult productions was not significant (Mann-Whitney tests, $p > .150$). Taken together, the acoustic and auditory results validated the use of the straight count condition to assess prominence shifting in the phrasal conditions.

Word-Level Prominence Patterns Across Conditions

The second set of analyses tested the effect of condition on duration ratios in number words. The first analysis showed a significant interaction between group and condition [$F(2, 824.07) = 4.05, p = .018$], and a nearly significant simple effect of condition [$F(2, 824.07) = 2.93, p = .054$] on duration ratios. The significant interaction between group and condition led us to split the data by group to test for an effect of condition in child and adult productions, respectively. These results are shown in Figure 3.

Figure 3 approximately here

The second analysis showed that condition had no effect on adults’ realization of lexical stress. This result indicates that adults did not stress shift in either of the phrasal contexts. Condition had an effect on children’s realization of number words [$F(2, 384.01) = 7.21, p =$

.001]. The difference between ratios in the clash condition and the straight count was not significant, and neither was the difference between the clash and non-clash conditions. This result indicates that children did not stress shift in either of the phrasal contexts. The significant effect of condition was therefore due to a difference between the straight count and the phrasal contexts. Mean comparisons indicated only significantly smaller duration ratios in the non-clash condition compared to the straight count [$p = .001$].

The finding that children had significantly smaller duration ratios in the non-clash condition compared to the straight condition context was due to a lengthening of the second rhyme compared to the first in the phrasal context. In particular, the mean durations of the first rhyme were similar across the two conditions (r1: straight count = 109 ms, $SD = 46$; non-clash condition = 106 ms, $SD = 45$), but those of the second rhyme were quite different (r2: straight count = 133 ms, $SD = 50$; non-clash condition = 166 ms, $SD = 54$). This lengthening suggests a strengthening of second syllable prominence. Although not significant, the mean duration ratio of number words in the clash condition was also smaller than the mean duration ratio of number words in the straight count condition. Again, this was due to a longer second rhyme in the phrasal context (r2: straight count = 133 ms, $SD = 50$; clash condition = 146 ms, $SD = 59$). The mean durations of the first rhyme were similar across the two conditions (r1: straight count = 109 ms, $SD = 46$; clash condition = 108 ms, $SD = 45$). Note also that the mean duration ratio of number words in the straight count condition was more similar to that of the adult mean duration ratio than to the mean duration ratios of number words in both of the phrasal contexts. Altogether these results suggest that children's realization of lexical stress changed as a function of phrase length rather than as a function of metrical context per se.

Like the analysis on duration ratios, the analysis on amplitude ratios indicated a significant interaction between group and condition [$F(2, 823.31) = 4.25, p = .015$], and no significant simple effects. These results are shown in Figure 4.

Figure 4 approximately here

When the data were split by group, the effect of condition was found only for adults [$F(2, 439.01) = 4.43, p = .012$]. A comparison of amplitude ratios in the straight count and two phrasal contexts indicated somewhat lower ratios in the straight count condition compared to the non-clash and clash context, but these differences were not significant after corrections for multiple comparisons. Nonetheless, the result suggests some effect of phrase length on adult speech, just like in child speech. In contrast to the child data, though, the slight increase in prominence was in the first syllable of the number word (i.e., absolute phrase-initial position).

Phrase-Level Prominence Patterns Across Conditions

The next set of analyses investigated the effect of condition on the realization of pitch accents. Results from the omnibus analysis showed significant main effects of condition [$F(2, 803.81) = 51.35, p < .001$] and group [$F(1, 24.51) = 5.41, p = .029$] on F0 ratios in number words production, but no interaction between the factors. These results are shown in Figure 5.

Figure 5 approximately here

Comparisons of F0 ratios within each group confirmed that these were significantly lower in the straight count condition than in the phrasal context conditions. The relatively low F0 ratio associated with the straight count suggests an initial low-high contour (LH), which could signal the presence of a low pitch accent on the first syllable (L*), or a high pitch accent on the second (H*). Of course, since the ratios were calculated using values averaged across the rhyme, the question of peak location cannot be directly addressed. However, the previously reported finding

that listeners heard the second syllable as more prominent than the first in the straight count condition favors the LH* interpretation of the initial low-high contour for both child and adult speech.

In order to more accurately assess pitch accent location in number words, normalized F0 contours were reconstructed from the F0 measures to compare intonation patterns across the three conditions and across the two groups of speakers. These contours are shown in Figure 6.

Figure 6 approximately here

Visual inspection of children's contours suggests the possibility of a leftward shift of the high tone in both phrasal contexts, whereas the adult low-high pattern appears to remain more or less constant across conditions. Analyses on normalized F0 values confirm the observation that intonation patterns differ across conditions more in child speech than in adult speech. An analysis on the F0 midpoint values in the two rhymes of the number words indicated a significant three-way interaction between group, rhyme position, and condition [$F(2, 1453.41) = 4.47, p = .012$]. When the data were split by group, the interaction between rhyme position and condition was significant in the child [$F(2, 704.74) = 10.88, p < .001$] and adult data [$F(2, 748.55) = 3.80, p = .023$], but for different reasons. Figure 6 shows that the interaction in the child data was due to a change in pattern from LH in the straight count to HL in the non-clash condition and to HH in the clash condition. We will refer to the change from LH to HL as a shift in pitch accent from the second to the first syllable, and to the change from LH to HH as the spreading of a high tone from the second to the first syllable. In contrast to the changes observed for children's speech, the interaction in the adult data was due to changes in the excursion from low to high, which at their maximum approximated the spreading of a high tone in the clash condition but never resulted in a pitch accent shift.

In sum, both children and adults produced an initial low-high contour in the straight count condition. Group difference emerged in the non-clash condition where adults continued to produce an F0 pattern more similar to that of the straight count, whereas children shifted the high tone from the second to the first syllable.

Discussion

The current study examined the integration of word- and phrase-level prominence patterns in children's speech by investigating the temporal alignment of lexical stress and pitch accents in two-word phrases. By varying metrical context, we created opportunities for the misalignment of word- and phrase-level prominence on *N-teen* number words. The predictions were that adults would always align lexical and phrasal prominences, but that the location of cumulative prominence would certainly be on the first syllable in the clash context (prediction 1), but possibly on the second in the non-clash context (prediction 2). Children were predicted to behave like adults in the clash context; albeit for the additional reason that they have a preference for trochaic patterns (prediction 3). We expected children to preserve the repeating iambic lexical stress pattern in the non-clash context, but predicted that a misalignment in prominences might occur if children also preserved a preference for early accent (prediction 4). Even though prediction 1 was not upheld, in that prominences were temporally aligned with the second syllable across conditions in adult speech, word- and phrase-level prominences were always integrated. By contrast, word- and phrase-level prominences were never fully integrated in phrasal contexts in children's speech. Relative to the straight count condition, children strengthened the lexical prominence of the second syllable regardless of the conditioning context. This result was at odds with prediction 3. The F0 results indicated, however, that

children used an initial high pitch accent in the phrasal contexts, albeit more consistently in the non-clash context than in the clash context. This result suggests an early accent bias in children's speech, but one that does not precipitate stress shift.

The misalignment of prominences in child speech is illustrated in Figure 7. This figure shows how children organized prominence patterns in two-word phrases (Panels B and C) as compared to one-word phrases (Panel A). In longer phrases, children (1) strengthened lexical prominence on the second syllable of the number words, and (2) either shifted or spread the high pitch accent to the phrase-initial syllable. By contrast, lexical and phrasal prominence patterns in adults' productions were more invariant (robust) across all types of phrases as indicated by the duration and F0 correlates. Taken together, these results suggest that prominence patterns are less stable in child language compared to adult language. In particular, the results suggest that the prosodic structures that condition prominence integration are still immature in 6-year-old children.

Figure 7 approximately here

Below, we discuss these findings, first with reference to the notion of stress shifting, and then in broader terms.

Prominence Shifting and Prosodic Templates

Stress shifting from non-initial stressed syllables to phrase-initial syllables in English is viewed as an adjustment of the prosodic template in stress clash contexts (Metrical Stress Theory; Liberman & Prince, 1977) or as the realization of a preferred prosodic template (Intonational-Metrical Theory; Shattuck-Hufnagel et al., 1994). Previous studies investigating stress shift have compared clash and non-clash contexts to each other (e.g., Goffman et al., 2006;

Vogel et al., 1995). Across context differences in such studies are interpreted as an effect of stress clash on prominence location. Stress shifting is said to have occurred when an unstressed syllable is strengthened or a stressed syllable is neutralized through weakening (e.g., Goffman et al., 2006; Vogel et al., 1995). We have already noted that, in our study, there was little effect of metrical context in adult speech, and the expected stress shift was not observed. Given that context effects were interpreted with reference to a baseline condition, the near invariant patterning in adult speech likely represents a basic prosodic template associated with number word production in the specific functional linguistic context of counting. Although first syllable amplitudes in adult speech were slightly higher in the phrasal conditions compared to the straight count condition, word-level stress and the location of the high pitch accent was invariantly on the second syllable, as indicated by the duration ratios and F0 values across conditions. If the near invariant pattern in fact reflects a prosodic template for counting, then the findings fall outside of what is predicted by Metrical Stress Theory or Intonational-Metrical Theory with regard to stress shifting and phrase-initial pitch accenting.

As for the increased amplitude of phrase-initial syllables in two-word phrases in adult speech, we note that while amplitude typically covaries with duration (e.g., Fry, 1955; Kochanski et al., 2005; Mo, 2008), it also independently marks phrase-initial position (Cho & Keating, 2009). For example, Cho and Keating found higher amplitude in phrase-initial vowels compared to phrase-medial vowels. This pattern of so-called initial strengthening corresponds to our results.

Phrasal Context Effect on Children's Prominence Patterns

Adults exhibited minimal duration and F0 changes across counting conditions in our study, suggesting that they have a robust prosodic template for counting. By contrast, children exhibited substantial differences in their realization of prominence across conditions. In this section, we consider several explanations for this variability in child speech.

The 6- and 7-year-old children in our study produced the same lexical prominence patterns as adults for trochaically- and iambically-stressed nouns and for number words in the straight count condition. The absence of between-group differences here is consistent with prior work showing that children phonetically distinguish lexical stress patterns using both duration and amplitude as early as age 3 or 4 years (Goffman & Malin, 1999; Kehoe et al., 1995; Pollock, Brammer, & Hageman, 1993). Where children differed from adults was in the production of number words in the phrasal context conditions. In particular, children strengthened the second syllable in number words when these were followed by nouns. This strengthening cannot be explained in terms of immature control over the correlates of lexical stress per se, or in classical terms of stress clash avoidance. Rather, the prominence adjustments observed in the current study suggest that children may focus on the production of lexical stress patterns at the expense of word- and phrase-level prominence integration. This possibility seems especially likely for the context with an iambic (weak-strong) alternation as described below.

It is possible that children were especially focused on word-level patterns in the weak-strong condition (e.g., non-clash context in this study) because the weak-strong pattern in a word with two full vowels (e.g., *N-teen* number words) may be more difficult for English-speaking children than the strong-weak pattern (e.g., *bárbeque*). As a consequence, children may 'over do' the weak-strong pattern in contexts where adults do not. This suggestion is consistent with the

documented preference for the trochaic metrical pattern in early childhood (Allen & Hawkins, 1980; Gerken, 1991; Kehoe et al., 1995; Snow, 2007), and with the developmental assumption that patterns that are acquired earlier are more easily accessed and controlled. Note also that this assumption is fully consistent with work by Goffman and colleagues who describe trochaic forms as requiring less modulation and iambic forms as requiring more, that is, a greater degree of articulatory weakening or strengthening of one of the syllables involved (e.g., Goffman et al., 2006; Goffman & Malin, 1999). In a phrasal context, the speaker can either default to the less modulated pattern through prominence neutralization (e.g., duration ratios closer to 1) or strive to maintain the weak-strong pattern (e.g., duration ratios closer to 0.5). Our suggestion is that in middle childhood children strive to maintain lexically-specified patterns and, in a sense, end up ‘over-doing’ them, especially under the challenging condition of integrating the pattern into a phrase.

With regard to phrasal prominence patterns (i.e., pitch accenting), both children and adults in our study produced a low-high F₀ contour in the straight count condition. The high tone was in the second syllable of the number words as illustrated in Panel A of Figure 7. In the clash context, this high tone tended to spread from the second to the phrase-initial syllable as illustrated in Panel B. Group differences emerged in the temporal alignment of the high tone in the non-clash context. In children’s productions, the high tone shifted leftward to the phrase-initial syllable as illustrated in Panel C. In contrast, adults produced the high tone on the second syllable of the number words. The similarity of children’s productions across the two contexts suggests that F₀ contours in child speech may be more influenced by the presence of a phrasal context than by the type of context (clash or non-clash).

There are at least two types of explanation – one structural and the other pragmatic – for why a tone might shift to the phrase-initial syllable in children’s productions of phrases. The first type of explanation concerns intonational structure, which can be decomposed into phrase boundary tones and pitch accents (Cruttenden, 1986; Pierrehumbert, 1980; Snow, 2007). A phrase-initial rise in F0 may represent an interpolation between a low boundary tone and a high-tone pitch accent (L% and H* in ToBI transcription conventions; Beckman et al., 2005; Pierrehumbert, 1980). It is possible that the boundary tone is overridden by the pitch accent in children’s productions of longer or more complex phrases because the F0 rise is more difficult for children to produce than other contours (Patel & Grigos, 2006; Snow, 2007). The articulatory difficulty of an initial F0 rise may have to do with its slow, non-iconic F0 change. Alternatively, the rise could represent an immature understanding of intonational phonology, resulting in the collapse of boundary tones and pitch accents.

Along these lines, it is also possible that children have acquired intonational phonology, and merely have difficulty with the implementation. That is, it could be that children correctly represent the pitch accent on the second syllable, but inadvertently spread the tone to the first syllable in production. While we acknowledge that it is difficult to completely ascertain the location of pitch accent placement, we assume contra the possibility of a representation-production mismatch, that the underlying contour is H*L or H*H contour in the child data because the F0 values were higher in the first syllable. Moreover, a H*L or H*H target contour is consistent with perceptual and acoustic studies that suggest a strong early accent bias in child language (Allen & Hawkins, 1980; Atkinson-King, 1970; Vogel & Raimy, 2002).

A second type of explanation for children’s pitch accent placement on the phrase-initial syllables is pragmatic. Note that the only new information in the counting task was the number word itself and, more specifically, the initial syllable of the *N-teen* number words (i.e., *thir.teen*, *four.teen*, *fif.teen*, and so on). The nouns *barbeque* and *banana* were repeated throughout the task

and so always represented given information. According to the analysis in Pierrehumbert and Hirschberg (1990), a high-low pitch accent (falling F0 contour) is used to highlight information that is new to the discourse. A low-high pitch accent (F0 rise) is used to mark information selected from a small well-known domain of alternatives. It could be that children used the first type of pitch accent and adults the second in our study. Then again, it could be that only children's behavior reflected pragmatic influences: adults may have found the counting task so mundane that phonological structure was privileged over information structure.

Phrasal Prosody and Compound Noun Prosody

Although we have been treating target productions in the phrasal contexts as generic numeral-noun phrases, one could also imagine them as compound noun phrases. In fact, there is a documented preference in early and middle childhood for compound stress patterns (Atkinson-King, 1970; Vogel & Raimy, 2002). This preference is observed in English-speaking children who, given a task of prosodic disambiguation between a compound noun interpretation and an adjective-noun interpretation of two-word collocations produced by adults (e.g., *hót dog* vs. *hot dóg*, *gréen house* vs. *greenhóuse*, respectively), tend to favor the compound noun interpretation. In other words, children tend to misperceive the first word in two-word phrases as most prominent, which is reminiscent of a trochaic bias and an early accent bias in young children's productions. The acoustic analysis of stimulus materials in Vogel and Raimy (2002) showed an F0 rise in the first word of compound nouns, whereas F0 fell slightly in the first word of adjective-noun phrases. It is possible that children's misperception of the falling contour as also indicative of a compound noun is due to an immature understanding of boundary tones and pitch accents. This immature understanding may lead children to collapse the two tone types and treat

the initial rises and falls as variable realizations of an underlying initial high tone. If this is the case, then one could imagine that children might also default to producing initial high tones in two word phrases. Thus, a final explanation for children's behavior in the current study is similar to the first: children may collapse boundary tones and pitch accents in a phrasal context. This could be due to an early accent bias, as previously suggested, which may itself be due to the misapprehension of %L+H* as another instance of H*. Future studies of prominence patterns in different types of tasks with different kinds of materials is clearly needed to further investigate the reason for the phrasal context effects found in the present study.

Integrated Prominence and Speech Rhythmicity

As a final matter, let us consider the integration between word- and phrase-level prominence patterns in relation to speech rhythmicity. We conceived of prominence integration as the temporal alignment of lexical stress and pitch accent in a single syllable, which is indexed by coordinated changes in duration, amplitude, and F0. Since amplitude also has the function of marking prosodic boundary strength (Cho & Keating, 2009), we will consider the question of prominence integration in terms of its other two correlates.

Focusing on duration and F0, we find that lexical stress and pitch accents were aligned in all adult productions in our study. Prominences in child speech were not integrated in this way. The variation suggests that word- and phrase-level prominence patterns were poorly aligned and the structures only loosely integrated. Note that these group differences in prominence integration have consequences for rhythmicity. This is because rhythm is largely defined by the spacing between cumulative prominences (e.g., Arvaniti, 2009). Since adults' pitch accents were always aligned with the lexically stressed syllables, the well-defined cumulative prominences

created a strong sense of alternating rhythm (strong-weak or weak-strong) in adult speech. Not so for child speech. The absence of well-defined cumulative prominences, due to the temporal misalignment of word- and phrase-level prominences, means that each prominence was weaker and more evenly distributed across the phrase. The more even distribution of weak prominences results in the percept of a more evenly-timed rhythm. Thus, the local patterns of prominence production observed in the current study likely provide at least a partial explanation for developmental differences in global rhythm that have been remarked on in a separate, but related literature (e.g., Allen & Hawkins, 1980; Grabe, Watson, & Post, 1999; Payne, Post, Astruc, Prieto, & Vanrell, 2012; Prieto, Vanrell, Astruc, Payne, & Post, 2012).

Conclusion

To conclude, word- and phrase-level prominence patterns are temporally aligned in adult speech, and invariant across phrases varying in length and metrical structure, suggesting a robust prosodic template for counting. School-age children have yet to fully acquire prosodic structures that are functionally independent of lexical constituent patterns. The development of adult-like phrasal prosody includes the use of both duration and F0 to signal prominence, and the structural coordination of pitch accents with phrase boundary tones. It could be that children acquire word- and phrase-level prominence patterns separately and then integrate them slowly over a long period of time. While lexical stress is acquired relative early by children, the acquisition of adult-like phrasal prosody appears to require more substantial linguistic experience.

References

- Astruc, L., Prieto, P., Payne, E., Post, B., & Vanrell, M. M. (2012). Tonal targets in early child English, Spanish, and Catalan. *Language and Speech*, 0(0), 1-25. doi: 10.1177/0023830912460494
- Allen, G. D., & Hawkins, S. (1980). Phonological rhythm: Definition and development. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology. Vol. I: Production*, 227-256. New York: Academic Press.
- Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica* 66, 46-63. doi: 10.1159/000208930
- Atkinson-King, K. (1970). Children's perception and identification of stress contrast. *Speech Communication* 9, 113-119. Retrieved from <http://hdl.handle.net/1721.1/56095>
- Beckman, M. E. (1986). *Stress and non-stress accent*. Dordrecht: Foris.
- Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P. A. Keating (ed.), *Phonological structure and phonetic form. Papers in Laboratory Phonology* 3, 7-33. Cambridge: Cambridge University Press.
- Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (Ed.), *Prosodic typology: The phonology and intonation of phrasing* (pp. 9-54). New York: Oxford University Press.
- Boersma, P., & Weenink, D. (2011). *Praat: doing phonetics by computer* (Version 5.2) [Software]. Available from <http://www.fon.hum.uva.nl/praat/>
- Bolinger, D. (1961). Contrastive accent and contrastive stress. *Language* 37(1), 83-96.
- Bolinger, D. (1985). Two views of accent. *Journal of Linguistics* 21(1), 79-123.

- Bolinger, D. (1986). *Intonation and its parts: Melody in spoken English*. Stanford, California: Stanford University Press.
- Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, 37, 466-485. doi:10.1016/j.wocn.2009.08.001
- Cooper, W. E., & Eady, S. J. (1986). Metrical phonology in speech production. *Journal of Memory and Language* 25(3), 369-384.
- 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(6), 2142-2156.
- Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.
- Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics* 26, 145-171.
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test 4: Manual* (4 ed.): NCS Pearson.
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America* 27(4), 765-768.
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech* 1, 126-152.
- Gerken, L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language* 30, 431-451.
- Goffman, L. (2004). Kinematic differentiation of prosodic categories in normal and disordered language development. *Journal of Speech, Language, and Hearing Research* 47, 1088-1102.

- Goffman, L., Heisler, L., & Chakraborty, R. (2006). Mapping of prosodic structure onto words and phrases in children's and adults' speech production. *Language and Cognitive Processes* *1/2/3*, 25-47. doi: 10.1080/01690960400001820
- Goffman, L., & Malin, C. (1999). Metrical effects on speech movements in children and adults. *Journal of Speech, Language, and Hearing Research* *42*, 1003-1015.
- Grabe, E., & Warren, P. (1995). Stress shift: do speakers do it or do listeners hear it? In B. Connell & A. Arvaniti (eds), *Phonology and phonetic evidence. Papers in Laboratory Phonology 4*, 95-110. Cambridge: Cambridge University Press.
- Grabe, E., Watson, I., & Post, B. (1999). The acquisition of rhythmic patterns in English and French. *Proceedings ICPHS 1999* (pp. 1201–1204), San Francisco.
- Gussenhoven, C. (1991). The English rhythm rule as an accent deletion rule. *Phonology* *8(1)*, 1-35.
- Hayes, B. (1984). The phonology of rhythm in English. *Linguistic Inquiry* *15(1)*, 33-74.
- Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.
- Horne, M. (1990). Empirical evidence for a deletion formulation of the rhythm rule in English. *Linguistics* *28(5)*, 959-982.
- Huss, V. (1978). English word stress in postnuclear position. *Phonetica* *35*, 86-105.
- Kehoe, M., Stoel-Gammon, C., & Buder, E. H. (1995). Acoustic correlates of stress in young children's speech. *Journal of Speech and Hearing Research* *38(2)*, 338.
- Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America* *118(2)*, 1038-1054. doi: 10.1121/1.1923349

- Lehiste, I., & Fox, R. A. (1990). Contribution of duration and intensity to the perception of prominence. *Journal of the Acoustical Society of America* **87**(S1), S72.
- Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry* **8**(2), 249-336.
- Mo, Y. (2008). Acoustic correlates of prosodic prominence for naïve listeners of American English. Paper presented at the *34th Meeting of the Berkeley Linguistic Society*.
- Mo, Y. (2009). F0 max and formants (F1, F2) as perceptual cues for naïve listeners' prominence perception. Paper presented at the *35th Annual Meeting of the Berkeley Linguistic Society*.
- Patel, R., & Grigos, M. I. (2006). Acoustic characterization of the question–statement contrast in 4, 7 and 11 year-old children. *Speech Communication* **48**, 1308-1318.
doi:10.1016/j.specom.2006.06.007
- Payne, E., Post, B., Astruc, L., Prieto, P., & Vanrell, M. M. (2012). Measuring child rhythm. *Language and Speech* **55**(2), 203-229. doi: 10.1177/0023830911417687
- Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Bloomington, Indiana: Indiana University Linguistics Club.
- Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, & P. Pollack (Eds.), *Intentions in communication*, 271-311. Cambridge, Mass.: MIT Press.
- Pollock, K. E., Brammer, D., & Hageman, C. (1993). An acoustic analysis of young children's productions of word stress. *Journal of Phonetics* **21**, 183–203.

- Prieto, P., Vanrell, M. M., Astruc, L., Payne, E., & Post, B. (2012). Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish. *Speech Communication* **54**, 681-702. doi:10.1016/j.specom.2011.12.001
- Quené, H., & Port, R. F. (2002). *Rhythmical factors in stress shift*. Paper presented at the 38th meeting of the Chicago Linguistic Society: The main session.
- Schwartz, R. G., Petinou, K., Goffman, L., Lazowski, G., & Cartusciello, C. (1996). Young children's production of syllable stress: An acoustic analysis. *Journal of the Acoustical Society of America* **99**(5), 3192-3200.
- Shattuck-Hufnagel, S. (1992). Early pitch accent placement within words in American English. *Journal of the Acoustical Society of America* **92**(4), 2443.
- Shattuck-Hufnagel, S., Ostendorf, M., & Ross, K. (1994). Stress shift and early pitch accent placement in lexical items in American English. *Journal of Phonetics* **22**, 357-388.
- Snow, D. (2007). Polysyllabic units in the vocalizations of children from 0;6 to 1;11: Intonation-groups, tones and rhythms. *Journal of Child Language* **34**, 765-797. doi: 10.1017/S030500090700815X
- Turk, A., & Sawusch, J. (1996). The processing of duration and intensity cues to prominence. *Journal of the Acoustical Society of America* **99**, 3782-3790.
- Vanderslice, R., & Ladefoged, P. (1972). Binary suprasegmental features and transformational word-accentuation rules. *Language* **48**(4), 819-838.
- Vogel, I., Bunnell, H. T., & Hoskins, S. (1995). The phonology and phonetics of the Rhythm Rule. In B. Connell & A. Arvaniti (Eds.), *Phonology and phonetic evidence: Papers in Laboratory Phonology 4*, 111-127. Cambridge: Cambridge University Press.

Vogel, I., & Raimy, E. (2002). The acquisition of compound vs. phrasal stress: The role of prosodic constituents. *Journal of Child Language* **29**, 225-250. doi:

10.1017/S0305000902005020

Wells, B., Peppé, S., & Goulandris, A. (2004). Intonation development from five to thirteen.

Journal of Child Language **31**(4), 749–778. doi: 10.1017/S030500090400652X

Table 1.

Acoustic correlates of lexically-specified stress patterns in context count nouns (barbeque, banana) and straight count numbers (N-teen)

Group	Sonorant rhyme ratios	<i>Bar:be(que)</i>	<i>Ba:na(na)</i>	<i>N:teen</i>
	r1:r2	\bar{M} (SD)	\bar{M} (SD)	\bar{M} (SD)
Child	Duration	2.38 (1.34)	.28 (.12)	.83 (.41)
	RMS	1.06 (.08)	1.01 (.04)	.99 (.09)
Adult	Duration	2.80 (1.17)	.29 (.09)	.84 (.41)
	RMS	1.08 (.07)	1.01 (.04)	1.03 (.08)

Figure Legends

Figure 1. Models of lexical and phrasal prominences in two-word phrases produced by adults.

Circles indicate prominence location, and arrows indicate prominence shifting.

Figure 2. An oscillogram and a spectrogram of one child's production of the phrase *fifteen barbeque*. Measurements of duration, amplitude, and F0 were extracted based on a segmentation of sonorant rhymes (r).

Figure 3. Duration ratios associated with number words across the three conditions. The number was followed by *bárbeque* in the clash context and by *banána* in the non-clash context. Error bars indicate $\pm 1 SE$ around the mean.

Figure 4. Amplitude ratios associated with number words across the three conditions. The number was followed by *bárbeque* in the clash context and by *banána* in the non-clash context. Error bars indicate $\pm 1 SE$ around the mean.

Figure 5. F0 ratios associated with number words across the three conditions. The number was followed by *bárbeque* in the clash context and by *banána* in the non-clash context. Error bars indicate $\pm 1 SE$ around the mean.

Figure 6. Reconstructed F0 contours of the number words in the straight count, the following *bárbeque* context (solid lines) and *banána* context (dashed lines). Error bars indicate $\pm 1 SE$ around the mean.

Figure 7. Prominence patterns in children's productions of number words (word-level: duration correlate; phrase-level: F0 correlate). Arrows indicate the spread in phrase-level prominence compared to the straight count.

(A) Metrical Stress Theory

(B) Intonational-Metrical Theory

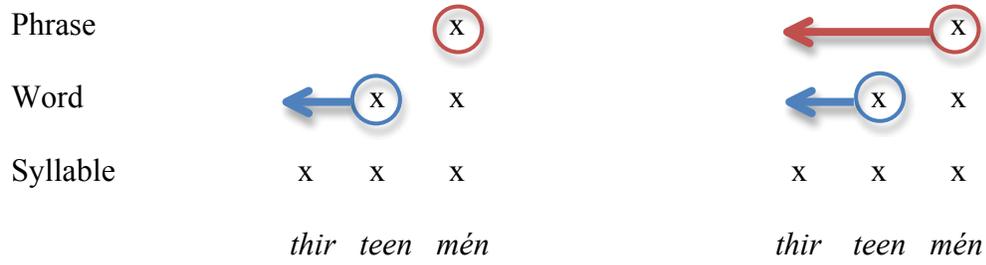


Figure 1

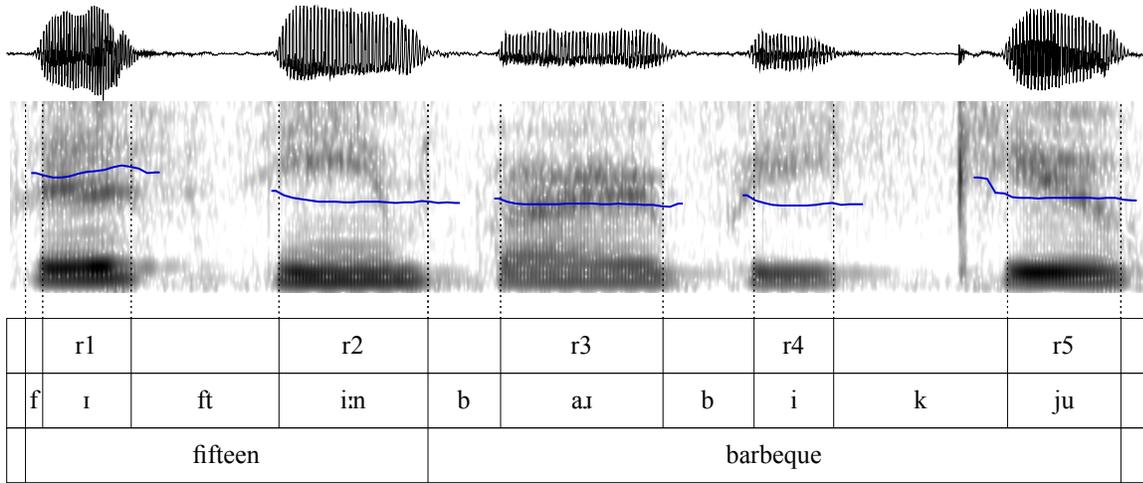


Figure 2

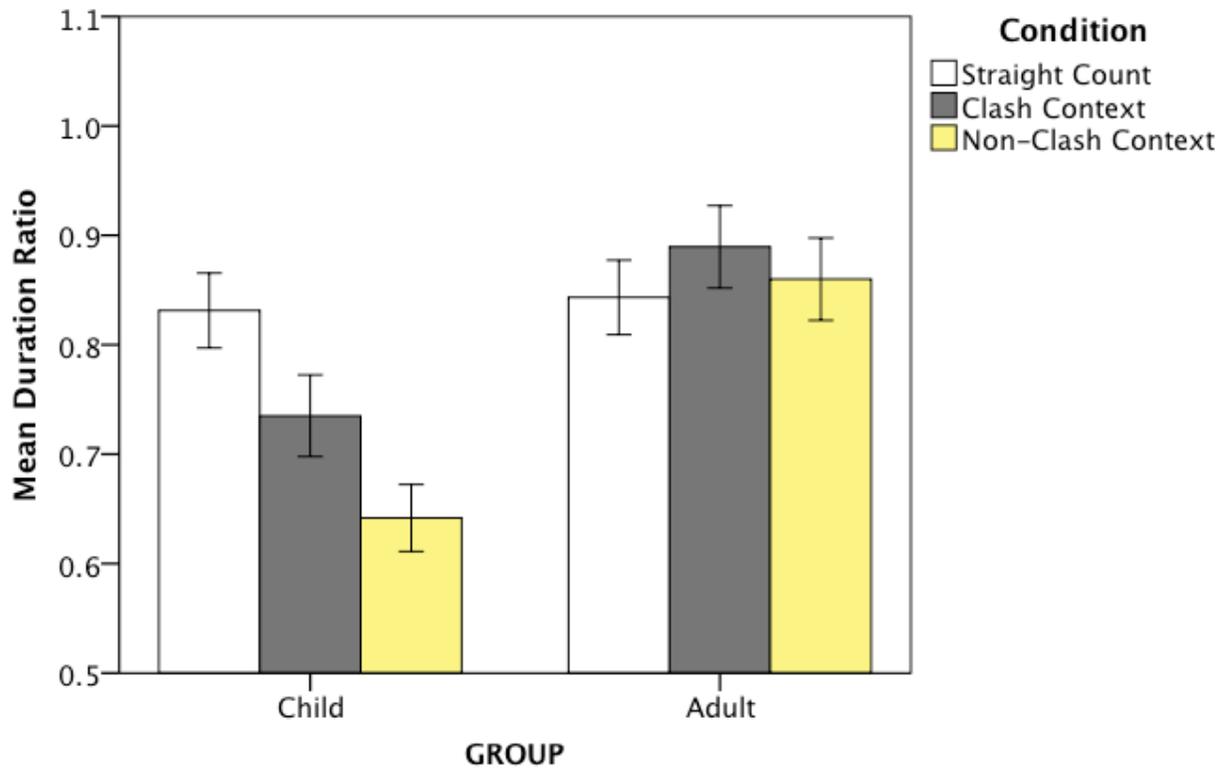


Figure 3

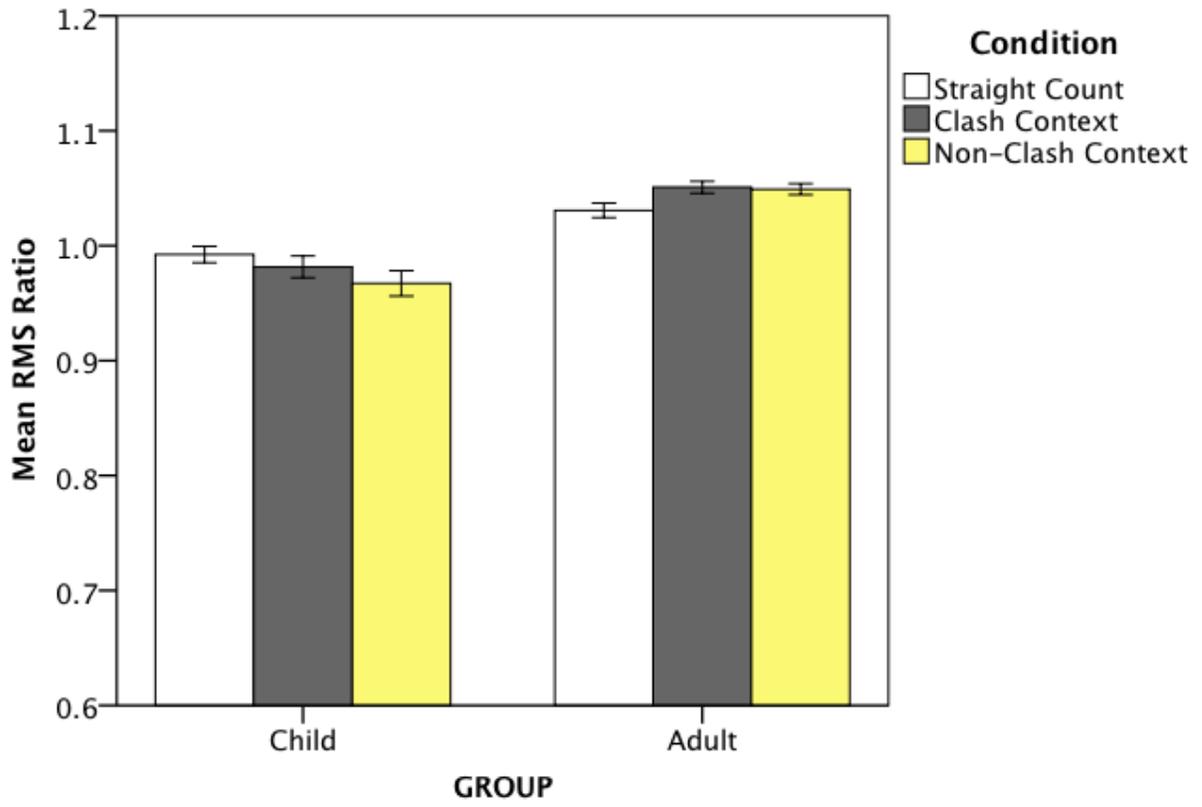


Figure 4

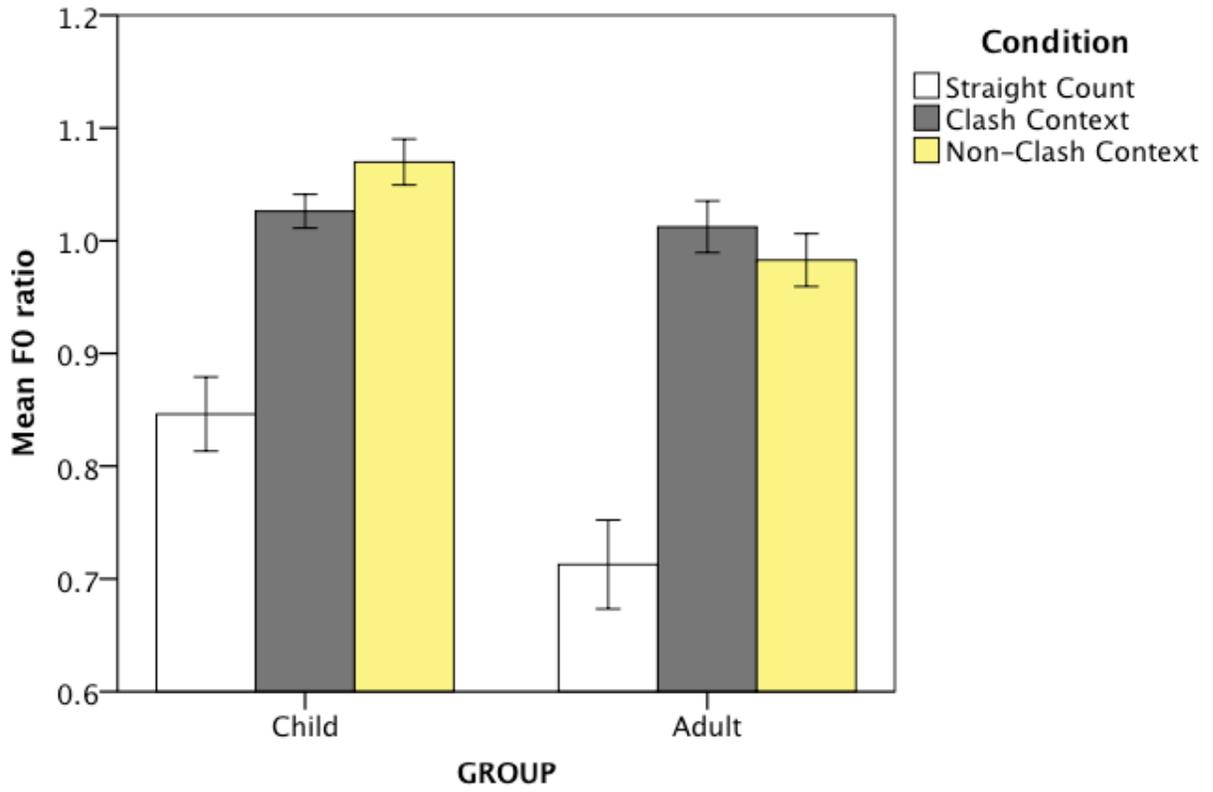


Figure 5

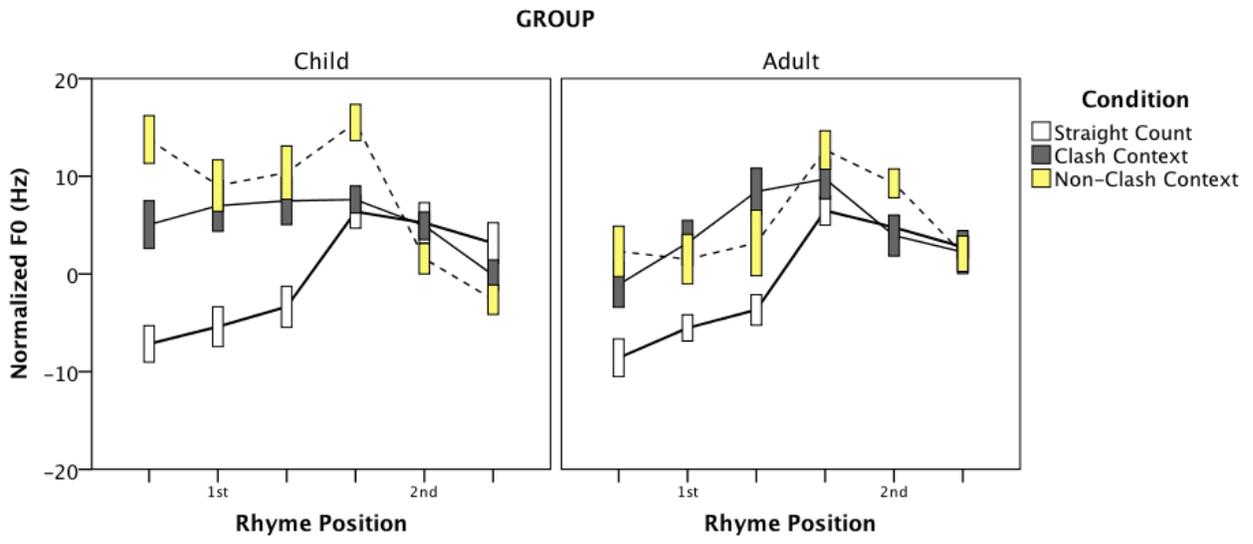


Figure 6

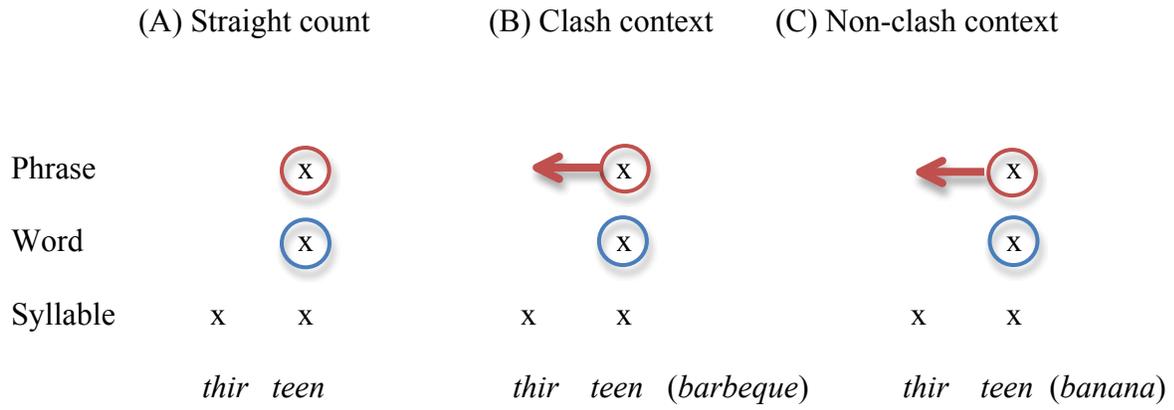


Figure 7