Journal of Child Language (2021), **48**, 110–128
doi:10.1017/S0305000920000227

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Five-year-olds produce prosodic cues to distinguish compounds from lists in Australian English

Ivan YUEN[1]*, Nan XU RATTANASONE[1], Elaine SCHMIDT[2], Gretel MACDONALD[3], Rebecca HOLT[1], and Katherine DEMUTH[1]

[1]Macquarie University, Department of Linguistics, 16 University Avenue, Macquarie University, North Ryde, Sydney, NSW 2109, Australia, [2]University of Cambridge, UK, and [3]Yuendumu School, Australia
*Corresponding author: Macquarie University – Linguistics, 16 University Avenue, Macquarie University, North Ryde, Sydney, NSW 2109, Australia Email: ivan.yuen@mq.edu.au

(Received 11 September 2018; revised 27 June 2019; accepted 24 March 2020;
first published online 13 May 2020)

Abstract

Although previous research has indicated that five-year-olds can use acoustic cues to disambiguate *compounds* ($N_1 + N_2$) from *lists* (N_1, N_2) (e.g., '*ice-cream*' vs. '*ice, cream*') (Yoshida & Katz, 2004, 2006), their productions are not yet fully adult-like (Wells, Peppé & Goulandris, 2004). The goal of this study was to examine this issue in Australian English-speaking children, with a focus on their use of F_0, word duration, and pauses. Twenty-four five-year-olds and 20 adults participated in an elicited production experiment. Like adults, children produced distinct F_0 patterns for the two structures. They also used longer word durations and more pauses in lists compared to compounds, indicating the presence of a boundary in lists. However, unlike adults, they also inappropriately inserted more pauses within the compound, suggesting the presence of a boundary in compounds as well. The implications for understanding children's developing knowledge of how to map acoustic cues to prosodic structures are discussed.

Keywords: prosody; lists/compounds; speech production; prosodic boundary; child language

Introduction

Although children's ability to convey prosodic functions is by and large established at the age of 5, some functions, such as distinguishing compounds (e.g., '*ice-cream*') from lists (e.g., '*ice, cream*'), continue to develop until the age of thirteen (Wells, Peppé & Goulandris, 2004). This involves 'prosodic chunking', or the use of prosody to delimit units within an utterance. Wells and colleagues examined prosodic chunking in British English. Age-related differences in the comprehension suggested that children become better at utilizing the relevant acoustic cues between the ages of eight and thirteen. This is in line with reported age-related improvements in other

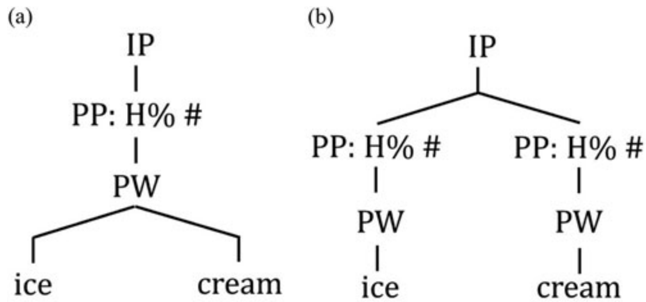

Figure 1. Prosodic structure of (a) compounds versus (b) lists. IP = intonational phrase, PP = phonological phrase, PW = prosodic word, H% = high boundary tone, # = phrasal boundary

prosodic phenomenon, such as differentiating compound vs. phrasal stress (Vogel & Raimy, 2002). An age-related change was also found in production, when Wells and colleagues analysed error data. Five-year-olds differed from older children in having more errors on two-item lists containing a compound than with three-item lists without one. Since the reported errors from Wells *et al.* (2004) were based on perceptual evaluation, it is not clear which acoustic cues the listener may have used. There is some evidence suggesting a correspondence between listener accuracy and acoustic realizations in other types of prosodic contrasts, such as distinguishing questions from statements (Patel & Grigos, 2006, Patel & Brayton, 2009). It is therefore necessary to examine children's acoustic realizations for a better understanding of how they implement acoustic cues during prosodic chunking. As children also exhibit individual variation in their acoustic implementations (cf. Dankovičová, Pigott, Wells & Peppé, 2004), it is also important to be able to replicate the observation about children's prosodic abilities (Wells *et al.*, 2004) (cf. discussions in Open Science Collaboration, 2015). With that in mind, the present study examined how the 5-year-old Australian English-speaking children acoustically realize the prosodic distinction between compounds and lists, and the extent to which the use of some acoustic cues might differ from the adults.

Adult use of acoustic cues to distinguish compounds from lists

Compounds are prosodically complex, composed of two prosodic words which combine to form a new prosodic word: [ice]$_{PW}$ [cream]$_{PW}$]$_{PW}$ with monomorphemic status (cf. Wheeldon & Lahiri, 2002; Wynne, Wheeldon & Lahiri, 2018). Wells *et al.* (2004) postulated two different structures for a noun + noun compound and two nouns respectively (Figure 1). The compound constitutes a single prosodic word and a phonological phrase, but the corresponding nouns constitute two prosodic words and two phonological phrases. These two structures differ in prominence pattern and temporal structure as reflected in different acoustic cues.

First, in terms of prominence, it is generally agreed that the first noun of the N + N compound is more prominent than the other (e.g., Liberman & Prince, 1977; Nespor & Vogel, 1986). Therefore, the compound '*ice-cream…*' will receive more prominence on '*ice*' than '*cream*' resulting in a strong-weak pattern; but in a list '*ice, cream…*' will receive equal prominence on both nouns resulting in a strong-strong pattern

https://doi.org/10.1017/S0305000920000227 Published online by Cambridge University Press

(Liberman & Prince, 1977; Nespor & Vogel, 1986; Hayes, 1995). Prominence is often associated with a pitch accent. When prominence is measured in terms of F_0 for every word in a compound, the strong-weak prominence pattern will be realised as an F_0 drop. However, the strong-strong prominence pattern will be realized as a relatively flat F_0 between words in a list.

Secondly, compounds and the corresponding two nouns will show different temporal structures, because of the absence vs. presence of a boundary. Due to its monomorphemic status, a compound is not likely to be inserted with a boundary within itself. Compound as a prosodic word also does not allow higher-level boundary (e.g., PHONOLOGICAL PHRASE BOUNDARY) to occur within itself because this will violate the Strict Layer Hypothesis (Nespor & Vogel, 1986). The absence of an internal boundary will foster tight temporal cohesion of a compound, resulting in short duration (cf. Farnetani, Torsello & Cosi, 1988). However, a list can be inserted with A PHONOLOGICAL PHRASE BOUNDARY. The presence of a postulated phonological phrase boundary in a list then will lead to pre-boundary lengthening. When a word occurs at the end of a sentence, as in 'He is a good boy', it is longer than when it occurs in utterance-medial position, as in 'The boy is happy'. Note that this is also pre-pausal. Pre-boundary lengthening at the end of an utterance is often referred to as utterance/phrase-final lengthening (Streeter, 1978; Lehiste, 1972; Klatt, 1975, 1976; Scott, 1982). Pre-boundary lengthening and pauses are also found to correlate with phonological phrase boundary and above (e.g., Price, Ostendorf, Shattuck-Hufnagel & Fong, 1991; Wightman, Shattuck-Hufnagel, Ostendorf & Price, 1992). The current study focused on pre-boundary lengthening in utterance medial position. Recent findings suggest that the scope of pre-boundary lengthening is not restricted to the final syllable of a word (Turk & Shattuck-Hufnagel, 2007), but can also extend as far as an initial unstressed syllable in a word such as '*guitar*' (Cho, Kim & Kim, 2013). Therefore, overall word duration, rather than simply syllable or rhyme duration, was used as an indicator of pre-boundary lengthening in the current study.

The other cue to the presence of a boundary is a *silent interval* (i.e., a pause) (e.g., Cutler & Butterfield, 1990). Pauses, like pre-boundary lengthening, are also associated with high-level prosodic boundary such as phonological phrases and above (Price *et al.*, 1991). The use of pauses in distinguishing compounds from lists was evident in Peppé, Maxim & Wells (2000), reporting the absence of pauses in 94% of compounds, and the presence of pauses in 88% of lists. Pause duration has also been observed to increase when the strength of prosodic boundary goes up from an 'intermediate phrase boundary' (e.g., phonological phrase in the current study) to an 'intonational phrase boundary' (Choi, 2003). Thus, given the structural difference between compounds and lists presented in Figure 1, pause duration (if there is one) is likely to be longer in lists than compounds. In short, adults can employ F_0, pre-boundary lengthening, and a pause to distinguish these two structures in utterance-medial position.

Child use of acoustic cues to compounds and lists

Some PERCEPTION studies have suggested that five-year-olds are sensitive to F_0 and word duration patterns when distinguishing conjoined structures such as '*pink and (green and white) socks*' from '(*pink and green) and white socks*' (Beach, Katz & Skowronski, 1996), and structures such as '*sunflower, pot*' from '*sun, flowerpot*' (Yoshida & Katz, 2004, Yoshida, 2007). In her thesis, Yoshida (2007) examined American English-speaking five- and seven-year-old children's use of prosody in

https://doi.org/10.1017/S0305000920000227 Published online by Cambridge University Press

processing ambiguous grammatical structure in the information-integration framework. She conducted a perception experiment using a test pair of sentences: '*sunflower, pot*' and '*sun, flowerpot*'. The main findings were that five- and seven-year-old children exhibited an adult-like cue-trading pattern between F_0 and word/pause duration, with an age-related developmental shift in the use of F_0 as a prosodic cue (cf. Yoshida & Katz, 2004). In other words, young children relied less on F_0 than duration, relative to other age groups. This observation differs from that of Beach *et al.* (1996) who reported a developmental shift in children's use of durational cues in American English. Perhaps this might be related to the fact that Yoshida (2007) manipulated both word and pause duration, whereas Beach *et al.* (1996) manipulated only word duration. Another study investigating compounds vs. phrasal stress has also reported that American English-speaking five-year-olds were less accurate than older children and adults in using acoustic cues to differentiate the two types of stress, partly because of the different prosodic domains to which stress is assigned (Vogel & Raimy, 2002). These differences in children's perceptual performance in different studies might be related to the different prosodic functions under investigation: prosodic disambiguation by means of the LOCATION of boundary in Beach *et al.* (1996) and Yoshida & Katz (2004), vs. PROSODIC DOMAINS for stress assignment in Vogel & Raimy (2002). These perception studies suggest that, by the age of five, children are sensitive to various prosodic cues, but that their ability to employ some acoustic cues is still not adult-like. The developmental patterns reported might thus be related to children's learning to map acoustic cues to specific prosodic functions in production.

Of relevance to the current study are a few PRODUCTION studies examining the acoustic realization of compounds vs. lists (cf. Yoshida, 2007; Yoshida & Katz, 2006; Dankovičová *et al.*, 2004). In Yoshida (2007), American English-speaking five- and seven-year olds described a picture prompt to a blindfolded experimenter in one of the two target sentences: '*sunflower, pot*' or '*sun, flowerpot*' in a total of seven repetitions. The F_0 contour and word duration of '*sun*' and '*flower*' were measured, as well as two inter-word pause durations (one between '*sun*' and '*flower*' and the other between '*flower*' and '*pot*'). The main findings were that these five- and seven-year-old children exhibited adult-like patterns for F_0, word and pause duration to distinguish the two prosodic structures: '*sunflower, pot*' vs. '*sun, flowerpot*' (cf. Yoshida & Katz, 2006). In addition, an age-related difference in pause duration was found: the seven-year-old children used shorter inter-word pause durations than the five-year-olds and the adults between '*sun*' and '*flower*'. Although pause was expected in this inter-word location in the target sentence '*sun*, *flowerpot*', most seven-year-olds did not use any. This suggests that the mapping of pause duration and incidence of pauses to the function of prosodic chunking takes time to master.

Despite the fact that the use of phrase-final lengthening appears in children around two years, at least at the ends of utterances (Snow, 1994), American English-speaking five-year-old children did *not* exhibit adult-like use of utterance medial word and pause durations to delimit items according to different phrasal groupings in the 'pink and green and white' example mentioned above (Katz, Beach, Jenouri & Verma, 1996). This contrasts with findings from Yoshida (2007), who argued that Katz *et al.* (1996) might have underestimated the five-year-old children's ability to use durational cues for prosodic grouping/chunking.

Examining children's use of durational cues (i.e., pre-boundary lengthening and pause duration) in British English, Dankovičová *et al.* (2004) focused on the data from the eight-year-olds who completed the Profiling Elements of Prosodic Systems – Child version (PEPS-C) in Wells *et al.* (2004). Although these children as a group utilized adult-like pre-boundary lengthening and longer pause durations as cues to the presence of phonological phrase boundary, there was individual variation in the use of two temporal/durational cues: while a third used both temporal/durational cues unambiguously, in line with the adult pattern, two thirds did not. This suggests that children are still not consistent in how they utilize durational cues at the age of eight.

Although these previous studies indicate that children generally can use prosodic cues to reflect phrasal groupings, there are variations in how these acoustic cues, particularly durational cues, are used from different age groups and English dialects. To disentangle children's ability to use acoustic cues in prosodic chunking from those potential confounds, it is necessary to be able to generalize the previous observations to other English dialects when age, stimuli and experimental design are similarly matched to those in previous studies (e.g., Yoshida, 2007; Yoshida & Katz, 2006; Wells *et al.*, 2004).

The current study therefore examined Australian English and investigated the generalizability of previous observations as to how five-year-old children use acoustic cues for producing compounds (N_1+N_2) and lists (N_1, N_2) in utterance-medial position, and the extent to which these children are adult-like in their use of acoustic cues. We tested this using an elicited production task where we examined the acoustic realisation of each of these relevant acoustic cues in children's productions, and compared this to that of adults.

Hypotheses

Given previous reports regarding children's ability to use adult-like acoustic cues to delimit a compound (e.g., Yoshida & Katz, 2006; Yoshida, 2007), three different hypotheses were thus formulated, one for each of the three cues.

- H1: Given that five-year-olds showed some degree of distinct F_0 patterns in Yoshida (2007), we expected that five-year-olds might also be able to use F_0 to reflect distinct prominence patterns for compounds and lists. A strong-weak prominence pattern within a compound would result in mean F_0 to fall, and a strong-strong prominence pattern in lists would result in mean F_0 to be relatively flat. These F_0 patterns would hold for children and adults, if both groups share similar prosodic structures.
- H2: Given that children as young as two years can use utterance-final lengthening to mark an utterance boundary (e.g., Snow, 1994), and that five-year-olds lengthen duration of a word preceding a phonological phrase boundary in '*sun, flowerpot*' (e.g., Yoshida & Katz, 2004, 2006), we expected the five-year-olds to implement pre-boundary lengthening in lists with a phonological phrase boundary. As a result, N_1 would be longer in lists than compounds. Therefore, words in a list would have an overall longer duration than the same two words in a compound. The pattern of pre-boundary lengthening would hold for children and adults, if both groups share similar prosodic structures.
- H3a: Based on Cutler & Butterfield (1990) and Yoshida (2007), we expected fewer pauses in compounds than in lists, because it is less likely for a pause to occur within a compound than between words in a list. This pattern of pause

Table 1. Seven sentence-medial noun-noun compound stimuli and their corresponding list stimuli embedded in a carrier sentence 'I can see *Stimuli* and Filler'

Compounds	Lists
ice cream	ice, cream
ice cubes	ice, cubes
goldfish	gold, fish
raincoats	rain, coats
jelly beans	jelly, beans
jellyfish	jelly, fish
waterslides	water, slides

distribution would hold for children and adults, if both groups share similar prosodic structures.

- H3b: Based on Yoshida (2007) and Dankovičová *et al.* (2004) we expected that pause duration between N_1 and N_2 would be longer in lists (N_1, N_2) than compounds (N_1+N_2). This pattern of pause duration would hold for children and adults, if both groups share similar prosodic structures.

Method

Participants

Twenty-four typically developing monolingual Australian English-speaking (AusE) children (9 M, 15 F) participated in the study (Mean age = 5;8 years, Range = 5;0–6;7). An additional seven children were excluded for failure to satisfy the criterion of producing 80% of the stimuli (n = 4), Autism Spectrum Disorder (n = 1), and experimental error (n = 2).

Twenty monolingual AusE-speaking undergraduates (6 M, 14 F) from the Sydney area formed the adult baseline for this experiment (Mean age = 21 years, Range = 18–30). They received course credit for participation in the study. A further 18 speakers were excluded, 12 due to exposure to an additional language at home, and six due to the heavy use of creaky voice, where F_0 could not be accurately tracked over the voiced region of each noun.

Stimuli

The target stimuli consisted of seven noun-noun items, where N_1 and N_2 occurred as compounds in one condition and as part of lists in another (see Table 1). In the compound condition, the compound (N_1+N_2) was embedded in a two-item list (**N_1+N_2**, N_3). In the list condition, the target stimuli (N_1, N_2) were part of a three-item list (**N_1**, **N_2**, N_3). All the nouns were identical in both experimental conditions. The target stimuli were vetted by all the authors to ensure they were culturally appropriate and familiar to children in Australia. We controlled the word frequency of the nouns using ChildFreq (Bååth, 2010). The N_1 items had a combined word frequency of 2199 per million with a mean word frequency of 314.1 per million. The N_2 items had a combined word frequency of 1076 per million with

https://doi.org/10.1017/S0305000920000227 Published online by Cambridge University Press

a mean word frequency of 153.7 per million. According to ChildFreq, frequency counts were not available for three compound words. The combined frequency for the remaining compound words was 370 per million with a mean frequency of 92.5 per million. The number of syllables in N_1 was varied to allow for generalisation across word length. All target stimuli occurred sentence-medially and were followed by another noun. This allowed us to dissociate utterance-final lengthening from any other boundary-related lengthening effects on the target stimuli. We also took care to use target words which were picturable and contained segments that five-year-old children could produce (Priester, Post & Goorhuis-Brouwer, 2011). The stimuli were presented as coloured cartoon pictures of objects.

Procedure

At the beginning of the testing session, participants were presented with pictures of the objects to be used in the experiment and were asked to name the objects. This ensured that participants were familiar with both the target nouns and the corresponding visual stimuli. Three practice items were then used to familiarise the participants with the task and the carrier sentence: 'I can see…'.

In the test phase, a female AusE native speaker played a language game with the participants where they were shown a set of two pictures for the compound condition or three pictures for the list condition and asked: 'What can you see here?' Participants were then instructed to respond by completing the carrier sentence using the names of the pictured objects in the order in which they appeared. No feedback was provided during elicitation. Thus, the response to a scene showing *ice-cream* and *juice* from left to right would be '*I can see ice-cream and juice*'. The response to a scene showing three items would be '*I can see ice, cream and juice*'. Compound and list items were pseudo-randomised to generate a test set. The item order of the test set was then reversed to generate a second test version. The two test versions were randomly assigned so that half the participants saw the pictures in one order, and the other half in the reverse order. The responses were audio-recorded in a sound insulated booth onto a computer using Audacity audio recording software at a sampling rate of 44.1 kHz, with a Behringer C2 condenser microphone.

Each participant produced 14 target sentences. This resulted in a possible total of 280 items from the adults and 336 items from the children. Nine items from the adult data and 36 items from the child data were discarded due to the insertion of 'and' between the two nouns of interest, misarticulation or naming errors (e.g., '*ice, cubes*' produced as 'ice, blocks'). Thus, a total of 271 items from the adults and 300 items from the children were included for acoustic coding.

Acoustic coding

All remaining productions were annotated and segmented in Praat (Boersma & Weenink, 2011). Since the stimuli contained various segment types in the nouns of interest, we adopted the following criteria to identify the onset and offset of N_1 and N_2. For onsets: (a) when the noun contained no onset consonant, the beginning of clear F2 and voicing were used as cues to the noun onset, (b) when the onset consonant was a plosive or an affricate, the beginning was indicated by the onset of the burst release, (c) in nouns beginning with an approximant /w/, we used the intensity minimum and the lowest formant transition in F2 as the word onset, (d) when the onset contained the approximant /ɹ/, the F3 minimum was used as the point of demarcation, and (e) in case of a fricative onset consonant, the beginning of high energy noise was used. The criteria for offsets were: (a) minimal high energy noise to identify the end of a fricative coda, (b) the beginning of the burst release to signal the end of words with a plosive coda, (c) the end of nasal formants and voicing to indicate the end of words with a nasal coda, and (d) cessation of F2 to identify the end of word-final vowels.

Analysis

Using Praat (Boersma & Weenink, 2011), we extracted F_0 and the duration values of the two nouns (N_1 and N_2), as well as the duration of the pause between them, in both the compound (N_1+N_2) and the list (N_1, N_2) conditions. F_0 values within the voiced region of N_1 and N_2 were examined for creak, glottalization and pitch errors (e.g., pitch doubling and pitch halving) in Praat. Pitch errors were corrected manually before automatically extracting the mean F_0 over the voiced region for each noun. Due to the difficulty in tracking F_0 with glottalization, nine of the 271 items from adults and one of 300 from children were removed from the data. This resulted in 262 items for the adults and 299 items for the children being used for statistical analysis.

Mean F_0 was employed to minimize any micro-prosodic perturbation at the beginning and the end of the voiced region for each noun. Due to differences in vocal tract size, F_0 differs between children and adults (e.g., Lee, Potamianos & Narayanan, 1999; Vorperian, Wang, Chung, Schimek, Durtschi, Kent, Ziegert & Gentry, 2009). Therefore, F_0 was transformed into normalized F_0 for each group according to the following: (F_0 – mean of $F_{0\ group}$) / Standard deviation of $F_{0\ group}$. Since child speech is often characterised by a slower speaking rate than adult speech, N_1 and N_2 durations were also transformed into normalized durations for each group according to the following: (Duration – mean of duration $_{group}$) / Standard deviation of duration $_{group}$. Normalization of F_0 and duration were further motivated by the findings in Aoyama, Akbari and Flege (2016) who found that 10-year-old American English-speaking children produced longer utterances and higher F_0 than adults in absolute terms, but these differences diminished using proportional metrics.

Since some items contained the closure duration of a plosive consonant between N_1 and N_2, a closure duration threshold was first factored in as a reference criterion to determine the presence or absence of a pause between N_1 and N_2. The reference closure duration was based on a recent corpus study of closure duration in English stops in TIMIT (Ghosh & Narayanan, 2009). The reference closure duration for voiced bilabials was 83 ms (Mean of 63 ms plus SD of 20 ms), and that for voiceless velars was 74 ms (Mean of 54 plus SD of 20 ms). If the temporal interval between the end of N_1 and the beginning of N_2 exceeded the reference closure duration by segment type, i.e., 83 ms for bilabial and 74 ms for velars, a pause was coded as present. A pause was considered absent otherwise. The presence of a pause was then tallied from the productions of children and adults in both compound and list conditions for analysis. Since the same nouns were included in both conditions (i.e., compounds and lists) and the same procedure of determining pauses was applied to both conditions, the effect of a stop consonant on the pause metric was kept constant across conditions. For items coded with presence of pause, pause duration was calculated according to the following: Temporal onset of N_2 – Temporal offset of N_1. Since speech rate might influence pause duration (e.g., Goldman Eisler, 1968; Fletcher, 1987; Trouvain & Grice, 1999), we also normalized pause duration

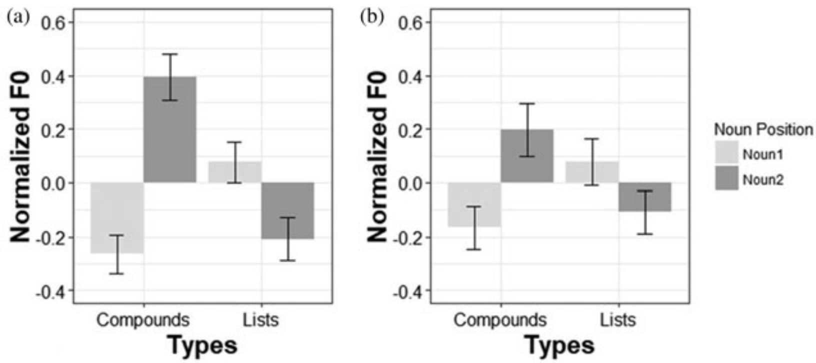

Figure 2. Normalized F_0 in Noun 1 (N_1) and Noun 2 (N_2) of Compounds (N_1+N_2) versus Lists (N_1, N_2) with +/- 1SE (standard error) in (a) Children and (b) Adults

according to the following: (Pause duration - Mean of pause duration $_{group}$) /Standard deviation of pause duration $_{group.}$ Normalized F_0, normalized duration, number of pauses and normalized pause duration were treated as dependent variables in subsequent statistical analysis.

Ten percent of the trials were randomly selected from children (n = 31) and adults (n = 36) for recoding by another annotator to check for coding consistency. These items covered a range of segment types in the target stimuli to ensure representative sampling. N_1 and N_2 durations were used as the dependent variable for correlation analysis, which showed high levels of reliability between the two coders for both the child ($r = .957$, $p < .0001$) and adult data ($r = .976$, $p < .0001$).

Results

Data from the 24 children and 20 adults were analysed. The normalized F_0, normalized duration and normalized pause durations were evaluated in the R package (R Core Team, 2015), using the lme4 package (Bates, Mächler, Bolker & Walker, 2015). The *anova* function, which provides Satterthwaite's approximation to degrees of freedom for estimating p-values, was used to test for statistical significance of linear mixed effects models in the lmerTest package (Kuznetsova, Brockhoff & Christensen, 2016). Pairwise comparisons of multi-level factors were performed with Tukey-HSD adjustments, using the lsmeans package (Lenth, 2016). The number of pauses was evaluated in a mixed effects logistic regression model (binomial), using the car package (Fox & Weisberg, 2011).

F_0

H1 predicted a strong-weak prominence pattern (i.e., F_0 fall) from N_1 to N_2 in compounds, but a strong-strong prominence pattern (i.e., relatively flat F_0) in lists. Figure 2 displays the overall pattern of normalized F_0 for N_1 and N_2 in compounds and lists from children and adults.

A linear mixed effects model was fitted to normalized F_0. The within-subject factors were Type (Compounds vs. Lists) and Noun position (N_1 vs. N_2), and the

Table 2. Statistical results of the linear mixed effects model # testing the effects of Type (Compounds vs. Lists), Noun position (N_1 vs. N_2), and Group (Children vs. Adults) on *normalized* F_0. *F*-values, degrees of freedom (*df*) and *p*-values are provided, where* indicates statistical significance with an alpha value of $p = .05$
F_0.model = lmer (normalized F_0 ~ type * noun position * group + (type * noun position | subjects) + (type * noun position | items))

Fixed Effects	*F*	*df*	*p*
Type	2.733	1	.1127
Noun position	5.106	1	*.0036***
Groups	.019	1	.892
Type: Noun position	32.47	1	*<.0001****
Type: Groups	1.629	1	.2079
Groups: Noun position	1.398	1	.2435
Type: Groups: Noun position	4.036	1	.0505

between-subjects factor was Group (Children vs. Adults). Random factors included by-subject and by-item intercepts and slopes. The results revealed significant main effects of Type and Noun position with a significant Type*Noun position interaction. Otherwise, no other interactions reached significance (Table 2). Figure 2 reveals how Type interacts with Noun position. While F_0 rose from N_1 to N_2 in compounds (N_1+N_2), the F_0 pattern was reversed in lists (N_1, N_2). That is, compounds exhibited a rising F_0 pattern from N_1 to N_2, but lists exhibited a falling F_0 pattern from N_1 to N_2. This result is counter to our predictions. The children and adults did not produce the expected F_0 pattern. The unexpected F_0 patterns might be related to the effect of intonation (cf. Morrill, 2011) to be discussed in Discussion. Despite that, compounds and lists exhibit DISTINCT F_0 patterns reflecting different structures.

Pre-boundary lengthening

According to H2, children, like adults, would use pre-boundary lengthening to indicate the absence vs. presence of a phonological phrase boundary in differentiating compounds (N_1+N_2) from lists (N_1, N_2), with longer duration for the N_1 in the list condition. Figure 3 displays the patterns of normalized duration for compounds and lists in children and adults. Normalized duration of N_1 and N_2 in lists have positive values, suggesting lengthening. However, normalized duration of N_1 in compounds has negative values, suggesting shortening (probably due to polysyllabic shortening within a compound). Normalized duration constituted the dependent variable in the linear mixed effects model, with Type, Noun position and Group as factors. The random structure of the model included by-subject and by-item intercepts and slopes.

The results showed significant main effects of Type and Noun position, with a significant Group*Noun position interaction (Table 3). As predicted, the overall normalized duration of N_1 and N_2 in lists was longer than that in compounds, and this pattern held for both children and adults, suggesting pre-boundary lengthening in lists. There was also a Group*Noun position interaction: the effect was due to a statistically significant difference between children and adults in the normalized duration of N_1 in the compounds, with children showing less shortening of N_1 than adults (Table 4).

https://doi.org/10.1017/S0305000920000227 Published online by Cambridge University Press

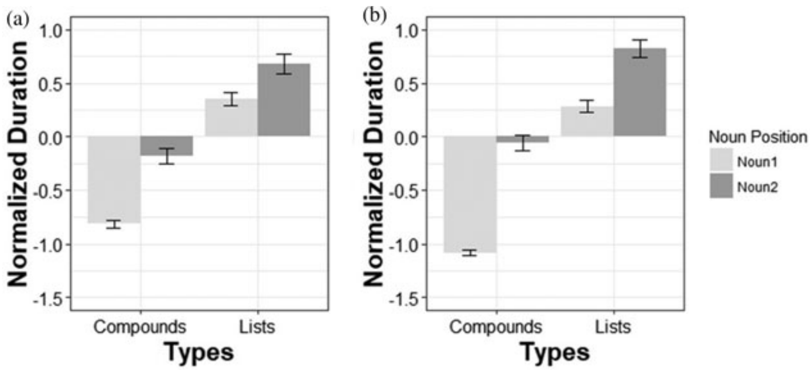

Figure 3. Normalized duration in Noun 1 (N_1) and Noun 2 (N_2) of Compounds (N_1+N_2) and Lists (N_1, N_2) with +/- 1SE (standard error) in (a) Children and (b) Adults

Table 3. Statistical results of the linear mixed effects model # testing the effects of Type (Compounds vs. Lists), Noun position (N_1 vs. N_2) and Group (Children vs. Adults) on *normalized duration* of stimuli. *F*-values, degrees of freedom (*df*) and *p*-values are provided, where* indicates statistical significance with an alpha value of $p = .05$
duration.model = lmer (normalized duration ~ type * noun position * group + (type * noun position | subjects) + (type * noun position | items))

Fixed Effects	*F*	*df*	*p*
Type	61.939	1	*<.0001****
Noun position	29.732	1	*<.0001****
Groups	.05	1	.8244
Type: Noun position	3.02	1	.0987
Type: Groups	2.049	1	.1563
Groups: Noun position	10.954	1	*.0016***
Type: Groups: Noun position	1.157	1	.2869

Table 4. Contrasts in normalized duration between Adults and Children for N_1 and N_2 in Compounds and Lists

Contrast	Types	Noun position	*Estimates*	*SE*	*df*	*t*	*p*
Adults – Children	Compounds	N_1	−.2783	.0865	42.9	−3.22	*.0127**
		N_2	.1115	.1535	44.6	.727	.0886
	Lists	N_1	−.0673	.1411	44.7	−.477	.9638
		N_2	.1329	.1587	44.6	.837	.8364

Pauses

There were two H3 predictions. First, we expected a difference in the incidence of pauses between compounds and lists, because it is less likely for a pause to occur

https://doi.org/10.1017/S0305000920000227 Published online by Cambridge University Press

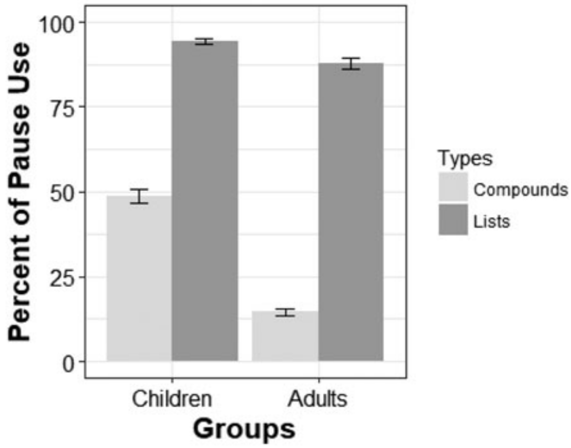

Figure 4. Percent of pause use in Compounds and Lists by Children and Adults, with +/−1SE (standard error)

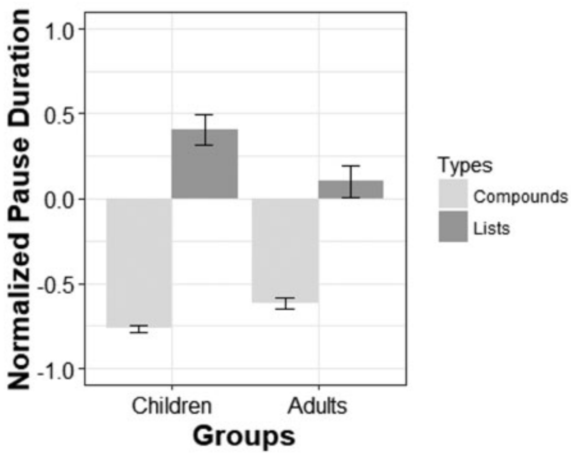

Figure 5. Normalized pause duration in Compounds and Lists from Children and Adults, with +/− 1SE (standard error)

within a compound than in a list. Second, we expected pause duration within a compound to be shorter than that within a list (if there is any pause in a compound), because N_1 in compounds is not in a pre-boundary position, whereas its N_1 counterpart in lists is. Figure 4 displays the incidence of pauses in compounds and lists for both children and adults, and Figure 5 shows the respective normalized pause duration.

A mixed effects logistic regression model (binomial) was fitted to the number of pauses, with Type and Group as factors. The reported model included by-subject intercepts and slopes and by-item intercepts. There were significant effects of Type, Group and Type*Group interaction (Table 5).

https://doi.org/10.1017/S0305000920000227 Published online by Cambridge University Press

Table 5. Statistical results of the logistic regression mixed effects model # testing the effects of Type (Compounds vs. Lists) and Group (Children vs. Adults) on *pause occurrence*. χ^2, degrees of freedom (*df*), and p-values are provided, where * indicates statistical significance with an alpha value of $p = .05$
#Model: glmer(pause occurrence ~ type * group + (type | subjects) + (1 |items), family = binomial, nAGO = 1)

Fixed Effects	χ^2	*df*	*p*
Type	27.36	1	<*.0001****
Groups	20.47	1	<*.0001****
Type: Groups	5.06	1	*.0245**

Table 6. Statistical results of the linear mixed effects model # testing the effects of Type and Group on *normalized pause duration*. *F*-values, degrees of freedom (*df*) and *p*-values are provided, where* indicates statistical significance with an alpha value of $p = .05$
#Model: lmer (pause duration ~ type * group + (type | subjects) + (1 | items))

Fixed Effects	*F*	*df*	*p*
Type	32.15	1	<*.0001****
Groups	.31	1	.5772
Type: Groups	2.7	1	.1044

Table 7. Mean N_1 durations (SD) of Compounds and Lists from Children and Adults with and without pauses

	Mean N_1 duration (SD) in milliseconds			
	Compounds		Lists	
Pauses	Yes	No	Yes	No
Children	348 (111)	315 (66)	567 (146)	368 (110)
Adults	211 (46)	267 (40)	456 (89)	386 (111)

Consistent with the prediction in H3a, children used significantly fewer pauses in compounds (N_1+N_2) than lists (N_1, N_2). Adults exhibited a similar pattern, with significantly fewer pauses in compounds (15%) compared to lists (88%). However, children employed three times more pauses than adults for compounds. The Type-by-Group interaction arose because children had a much higher use of pauses for compounds (48%) than adults.

To evaluate H3b, the normalized pause duration from items coded for the presence of pause was fitted in a linear mixed effects model, with Type and Group as factors. The model included by-subject intercepts and slopes and by-item intercepts. There was a significant effect of Type (Table 6). Thus, as predicted in H3b, pause duration in compounds was shorter than that in lists, for both children and adults.

As Wightman *et al.* (1992) reported pre-boundary lengthening to be larger for boundaries with pauses than those without, it is possible that the high usage of pauses in children's productions of compounds might be due to the presence of a

https://doi.org/10.1017/S0305000920000227 Published online by Cambridge University Press

major boundary. To explore this issue, we examined N_1 durations in compounds and lists with vs. without pauses. Table 7 displays the means and standard deviations of the N_1 durations in compounds and lists for children and adults. Children generally showed longer N_1 durations in compounds and lists with pauses than those without. However, adults exhibited shorter N_1 durations in compounds with pauses than those without. Since the N_1 is part of a monomorphemic compound and is not in a pre-boundary position, pauses are not expected. If pauses occur in compounds, adults seem to compensate by shortening N_1 duration. This suggests a compensatory relationship between the incidence of pauses and pre-boundary lengthening for adults, but this is not the case for children. Instead, children appear to have a longer N_1 duration with a pause in compounds than without – possibly suggesting the presence of a major boundary.

Discussion

In this study we investigated whether the five-year-olds can use prosodic cues to distinguish compounds from lists in their speech productions, and the extent to which their use of cues is acoustically adult-like. Regarding the use of F_0 as a cue, our results show that, like adults, these children can associate distinct F_0 patterns with compounds and lists, with F_0 rising from N_1 to N_2 in compounds and F_0 falling in lists. However, these distinct F_0 patterns did not correspond to the expected prominence patterns, namely the strong-weak pattern for compounds vs. the strong-strong pattern for lists. The compound stimuli in the current study were selected to have lexical stress assigned to N_1, whereas both N_1 and N_2 in the list attracted stress. If F_0 is used as one of the acoustic correlates of stress (e.g., Fry, 1958; Morrill, 2011), it was expected to be higher in N_1 than N_2 for compounds, but equal between N_1 and N_2 for lists. As Morrill (2011) has shown, the intonational/prosodic context can affect how robustly the F_0 cue serves as a correlate of compound stress in American English. For instance, a compound with primary stress on 'Red' in 'Red Sox' had a higher F_0 on 'Sox' than 'Red' when the compound was produced with question intonation which has a H% boundary tone. In our study, the five-year-old children and adults also exhibited a similar F_0 pattern as reported in Morrill (2011). This suggests that a H% boundary might have influenced how the expected prominence patterns of compounds could be realized. The short word length and the nature of segments (for example, voiceless stops selected for ease of segmentation) in our study made it difficult to examine the continuous F_0 contours of compounds and lists to shed light on the use of boundary tone, which deserves future investigation. Recall that Vogel and Raimy (2002) found that children had difficulty using stress information to differentiate compounds from phrases during a LISTENING task. Perhaps the difficulty children faced in relying on F_0 as a cue to compound stress in Vogel and Raimy (2002) might be related to the ambiguity of F_0 as a stress cue in different intonation contexts.

Similar to findings of American English in Yoshida and Katz (2006) and Yoshida (2007), our pre-boundary lengthening results indicate that, like the adults, Australian English-speaking children CAN employ duration to differentiate compounds from lists, with overall longer duration in lists. This suggests pre-boundary lengthening of a phonological phrase in lists. Children did NOT DIFFER from adults in the duration of N_1 and N_2 in lists, indicating that the pattern of pre-boundary lengthening is adult-like by five years in both English dialects. This is also consistent with previous

https://doi.org/10.1017/S0305000920000227 Published online by Cambridge University Press

reports that five-year-old children around the same age can use duration as a cue to demarcate other types of prosodic units in perception (Beach *et al.*, 1996).

Results from pause occurrence and pause duration indicate that both children and adults used more pauses and had longer pause durations in lists than compounds. The long pause duration patterns in the current study replicated those in Yoshida (2007). Children, like the adults, generally employed pauses and pause duration to signal a phonological phrase boundary in lists, consistent with the findings in Price *et al.* (1991) and Wightman *et al.* (1992). However, children differed from adults in their use of pauses in compounds, with pauses used 48% of the time after the N_1 in compounds, compared to only 15% for adults, though both groups used comparable rates of pauses in the lists. In other words, there is an age-related difference in the incidence of pauses within a compound.

The children's high usage of pause in compounds is inconsistent with their pre-boundary lengthening pattern, whereby N_1 and N_2 in compounds are shorter than their counterparts in lists. On the one hand, the high incidence of pauses in compounds suggests some kind of boundary; on the other, N_1 and N_2 duration in compounds suggest no boundary. Price *et al.* (1991) and Wightman *et al.* (1992) pointed out that pre-boundary lengthening increased with the presence of pause for high-level prosodic boundaries. When we compared the N_1 duration in compounds with and without a pause, we found that children lengthened only N_1 durations in the compounds with pauses, a temporal pattern opposite to that of adults who shortened N_1 duration in compounds with pauses. This, together with the incidence of pauses, suggests that children might treat N_1 in compounds as constituting a separate unit with a boundary. This inconsistent use of pauses and pre-boundary lengthening is in line with the inconsistent use of temporal cues from the eight-year-olds in Dankovičová *et al.* (2004), which used stimuli that were similar to that used in the current study. Recall that the correct prosodic structure of a compound is a holistic prosodic word. Adults process compounds as a single prosodic unit during phonological encoding (Wynne *et al.*, 2018). Yet the five-year-olds in the current study did not seem to use word duration and incidence of pause in a coherent manner to reflect that.

Our interpretation of these findings is that children are somewhat uncertain about the mapping between durational cues and the prosodic structure of compounds (N_1+N_2) in their acoustic realization. Perhaps they have a problem in suppressing lexical word structure in compounds. If children construct the high-level complex prosodic word (i.e., PW) of compound by building from embedded prosodic words in a BOTTOM-UP manner, this will lead them to insert a boundary between '*ice*' and '*cream*' in ($[[\text{ice}]_{PW}[\text{cream}]_{PW}]_{PW}$. This tendency for five-year-olds to insert a pause in compounds then suggests that lexical structure might have (mis)guided children to formulate prosodic structure (cf. Vogel & Raimy, 2002). Perhaps children first acquire the prosodic cues for simple prosodic words, and only later learn to use and weigh the word duration and pause cues consistently to reflect high-level prosodic units (e.g., PW) and different prosodic structures/domains (see Gerken, 2006; Demuth & McCullough, 2009, and Tang, Yuen, Xu Rattanasone, Gao & Demuth, 2019, for similar proposals for younger children in other domains). This interpretation is based on the observation that children do not compensate the presence of pause by shortening N_1 duration for compounds. Since unfilled pauses might also reflect verbal planning functions in five-year-olds (cf. MacWhinney & Osser, 1977), the high incidence of pauses might also suggest that it is cognitively demanding for children to encode the recursive structure of compounds during planning in lab speech, resulting in the inconsistent use of durational cues. In spontaneous speech, it would then be even harder for children to do so.

Although five-year-old children can use F_0, word duration, pause and pause duration to indicate the different prosodic structures of compounds and lists in Australian English, they are not yet adult-like in their use of pauses. Our findings are in partial agreement with previous findings. On the one hand, the findings in American English from Katz *et al.* (1996) does not fit neatly with our findings in Australian English. Perhaps this might be related to the different linguistic structures of the stimuli and tasks. It might be more difficult and cognitively demanding to map durational cues to reflect the three phrasal groupings/structures of '*pink and green and white*': (*pink*) (*and green*) (*and white*) vs. (*pink*) (*and green and white*) vs. (*pink and green*) (*and white*) in a spontaneous speech task (Katz *et al.*, 1996). In contrast, in our study there are only two structures to be disambiguated: '(*ice-cream*) (*and juice*)' vs. '(*ice*), (*cream*), (*and juice*)'. Yoshida (2007) eliciting similar linguistic structures in American English, found a similar ability of five-year-old children to use acoustic cues for prosodic chunking.

On the other hand, the age-related difference in the use of pauses in Australian English was not observed in American English (Yoshida, 2007). Perhaps this might be related to the number of test items. Yoshida (2007) tested a pair of phrases: '*sun, flowerpot*' vs. '*sunflower, pot*'; whereas the current study examined seven different test pairs (see also Dankovičová *et al.* (2004) who reported individual variation in the use of durational cues by eight-year-old children, using nine different test pairs).

Our data show that children do not have problems using acoustic cues to reflect different structures, but may have problems planning WHERE and WHAT KIND of boundary to use for compounds. This raises further questions as to what kinds of structural frames (prosodic or lexical) children generate and use to guide their speech production and planning and when this become adult-like. The current findings therefore have implication for assessing the prosodic abilities of atypical populations as well, such as children with autism spectrum disorder (ASD) who encounter problems in processing prosody (Peppé & McCann, 2003).

Conclusion

This study found that five-year-old children can utilize different acoustic cues to distinguish compounds from lists. Like adults, they employ pre-boundary lengthening and pauses to signal the presence of a phonological phrase boundary in lists. In other words, they do not have problems in using temporal cues to signal a boundary. However, these temporal cues were used INCONSISTENTLY in compounds, suggesting that children do not have an adult-like mapping of acoustic cues to the prosodic structure of compounds during planning. It seems that children tend to preserve a lexical word representation of N_1 in a compound, in competition with the status of a compound as a holistic complex prosodic word. This suggests that the challenge of prosodic chunking for five-year-old children may be related to the recursive prosodic word structure of compounds. These findings raise further questions regarding how and when children can construct adult-like prosodic structure for compounds, and how and when they can use these acoustic cues to distinguish compounds from lists. These issues can be tested in language comprehension as well, comparing the results to production in the same children.

Acknowledgments. This research was supported by Macquarie University and the following grants from the Australian Research Council (ARC): ARC CE110001021, awarded to Crain *et al.*; ARC FL130100014, awarded to Katherine Demuth. We would like to thank our participants for taking part in the research, Amit German for data collection and coding, and the Child Language Lab and Phonetics Lab for helpful feedback and suggestions.

References

Aoyama, K., Akbari, C., & Flege, J. E. (2016). "Prosodic characteristics of American English in school-age children," Proceedings of Speech Prosody 2016, 572–576.

Bååth, R. (2010). "ChildFreq: an online tool to explore word frequencies in child language," *LUCS Minor*, *16*, 1–6.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*; Vol *1*, Issue 1 (2015). Retrieved from https://www.jstatsoft.org/index.php/jss/article/view/v067i01

Beach, C. M., Katz, W. F., & Skowronski, A. (1996). "Children's processing of prosodic cues for phrasal interpretation," *Journal of the Acoustical Society of America*, *99*, 1148–1160.

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer. http://www.praat.org/ (Last viewed May 4, 2011).

Cho, T., Kim, J., & Kim, S. (2013). "Preboundary lengthening and preaccentual shortening across syllables in a trisyllabic word in English," *Journal of the Acoustical Society of America*, *133*, EL384–390.

Choi, J. Y. (2003). "Pause length and speech rate as durational cues for prosody markers," *Journal of the Acoustical Society of America*, *114*, 2395–2395.

Cutler, A., & Butterfield, S. (1990). "Durational cues to word boundaries in clear speech," *Speech Communication*, *9*, 485–495.

Dankovičová, J., Pigott, K., Wells, B., & Peppé, S. (2004). "Temporal markers of prosodic boundaries in children's speech production," *Journal of the International Phonetic Association*, *34*, 17–36.

Demuth, K., & McCullough, E. (2009). "The prosodic (re)organization of children's early English articles," *Journal of Child Language*, *36*, 173–200.

Farnetani, E., Torsello, C. T., & Cosi, P. (1988). "English compound versus non-compound noun phrases in discourse: an acoustic and perceptual study," *Language and Speech*, *31*, 157–180.

Fletcher, J. (1987). "Some micro and macro effects of tempo change on timing in French," *Linguistics*, *25*, 951–967.

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression*, 2nd edition. Thousand Oaks CA: Sage. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Fry, D. B. (1958). "Experiments in the perception of stress," *Language and Speech*, *1*, 126–152.

Gerken, L. A. (2006). "Decisions, decisions: infant language learning when multiple generalizations are possible," *Cognition*, *98*, B67–B74.

Ghosh, P. K., & Narayanan, S. S. (2009). "Closure duration analysis of incomplete stop consonants due to stop-stop interaction." *Journal of the Acoustical Society of America Express Letters*, *126*, EL1–EL7.

Goldman Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London and New York: Academic Press.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago, IL: University of Chicago Press.

Katz, W. F., Beach, C. M., Jenouri, K., & Verma, S. (1996). "Duration and fundamental frequency correlates of phrase boundaries in productions by children and adults," *Journal of the Acoustical Society of America*, *99*, 3179–3191.

Klatt, D. H. (1975). "Vowel lengthening is syntactically determined in a connected discourse," *Journal of Phonetics*, *3*, 129–140.

Klatt, D. H. (1976). "Linguistic uses of segmental duration in English: acoustic and perceptual evidence," *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). "lmerTest: tests in linear mixed effects models," R package version 2.0-33, URL https://cran.r-project.org/web/packages/lmerTest

Lee, S., Potamianos, A., & Narayanan, S. (1999). "Acoustics of children's speech: Developmental changes of temporal and spectral parameters," *Journal of the Acoustical Society of America*, *105*, 1455–1468.

Lehiste, I. (1972). “The timing of utterances and linguistic boundaries,” *Journal of the Acoustical Society of America*, *51*, 2018–2024.

Lenth, R. V. (2016). “Least-square means: the R package lsmeans,” *Journal of Statistical Software*, *69*, 1–33. <doi: 10.18637/jss.v069.i01>

Liberman, M., & Prince, A. (1977). “On stress and linguistic rhythm,” *Linguistic Inquiry*, *8*, 249–336.

MacWhinney, B., & Osser, H. (1977). “Verbal planning functions in children’s speech,” *Child Development*, *48*, 978–985.

Morrill, T. (2011). “Acoustic correlates of stress in English adjective-noun compounds,” *Language and Speech*, *55*, 167–201.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Riverton, NJ: Foris Publications.

Open Science Collaboration. (2015). “Estimating the reproducibility of psychological science,” *Science*, *349* (6251), 943–951.

Patel, R., & Brayton, J. T. (2009). “Identifying prosodic contrasts in utterances produced by 4-, 7-, and 11-year-old children,” *Journal of Speech, Language and Hearing Research*, *52*, 790–801.

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.

Peppé, S., Maxim, J., & Wells, B. (2000). “Prosodic variation in Southern British English,” *Language and Speech*, *43*, 309–334.

Peppé, S., & McCann, J. (2003). “Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version,” *Clinical Linguistics and Phonetics*, *17*, 345–354.

Price, P. J., Ostendorf, S., Shattuck-Hufnagel, S., & Fong, C. (1991). “The use of prosody in syntactic disambiguation,” *Journal of Acoustical Society of America*, *9*, 2956–2970.

Priester, G. H., Post, W. J., & Goorhuis-Brouwer, S. M. (2011). “Phonetic and phonemic acquisition: normative data in English and Dutch speech sound development,” *International Journal of Pediatric Otorhinolaryngology*, *75*, 592–596.

R Core Team. (2015). “R: A language and environment for statistical computing,” R foundation for statistical computing, Vienna, Austria. URL http://www.R-project.org/.

Scott, D. R. (1982). “Duration as a cue to the perception of a phrase boundary,” *Journal of the Acoustical Society of America*, *71*, 996–1007.

Snow, D. (1994). “Phrase-final syllable lengthening and intonation in early child speech,” *Journal of Speech and Hearing Research*, *37*, 831–840.

Streeter, L. A. (1978). “Acoustic determinants of phrase boundary perception,” *Journal of the Acoustical Society of America*, *64*, 1582–1592.

Tang, P., Yuen, I., Xu Rattanasone, N., Gao, L., & Demuth, K. (2019). “The acquisition of phonological alternations: The case of the Mandarin tone sandhi processes,” *Applied Psycholinguistics*, *40*, 1495–1526.

Trouvain, J., & Grice, M. (1999). “The effect of tempo on prosodic structure,” In Proceedings of the XIVth International Congress of Phonetic Sciences, San Francisco, CA, Vol 2, pp. 1067–1070. Martinez: East Bay Institute for Research and Education, Inc.

Turk, A., & Shattuck-Hufnagel, S. (2007). “Multiple targets of phrase-final lengthening in American English words,” *Journal of Phonetics*, *35*, 445–472.

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

Vorperian, H. K., Wang, S., Chung, M. K., Schimek, E. M., Durtschi, R. B., Kent, R. D., Ziegert, A. J., & Gentry, L. R. (2009). “Anatomic development of the oral and pharyngeal portions of the vocal tract: An imaging study,” *Journal of the Acoustical Society of America*, *125*, 1666–1678.

Wells, B., Peppé, S., & Goulandris, N. (2004). “Intonation development from five to thirteen,” *Journal of Child Language*, *31*, 749–778.

Wheeldon, L. R., & Lahiri, A. (2002). “The minimal unit of phonological encoding: prosodic or lexical word,” *Cognition*, *83*, B31–B41.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). “Segmental durations in the vicinity of prosodic phrase boundaries,” *Journal of the Acoustical Society of America*, *92*, 1707–1717.

Wynne, H. S. Z., Wheeldon, L., & Lahiri, A. (2018). “Compounds, phrases and clitics in connected speech,” *Journal of Memory and Language*, *98*, 45–58.

Yoshida, M., & Katz, W. F. (2004). "Children's use of prosody to identify ambiguous sets of compound nouns," *Journal of the Acoustical Society of America*, *116*, 2645–2645.

Yoshida, M., & Katz, W. F. (2006). "Children's production of prosody: Disambiguating sets of compound nouns," *Journal of the Acoustical Society of America*, *119*, 3421–3421.

Yoshida, M. (2007). *Children's use of prosodic information to understand and produce phrasal distinctions in American English*. (Doctoral dissertation, University of Texas at Dallas).

Cite this article: Yuen I, Xu Rattanasone N, Schmidt E, Macdonald G, Holt R, Demuth K (2021). Five-year-olds produce prosodic cues to distinguish compounds from lists in Australian English. *Journal of Child Language* **48**, 110–128. https://doi.org/10.1017/S0305000920000227