Article:
Dankovicová, J., Pigott, K., Wells, B. et al. (2004) Temporal markers of prosodic boundaries in children's speech production. Journal of the International Phonetic Association, 34 (1). pp. 17-36. ISSN: 1475-3502

https://doi.org/10.1017/S0025100304001525

Temporal markers of prosodic boundaries in children's speech production

Jana Dankovičová
University College London
j.dankovicova@ucl.ac.uk

Kathryn Pigott
Greenwich Primary Care Trust, London

Bill Wells
University of Sheffield

Sue Peppé
Queen Margaret University College, Edinburgh

It is often thought that the ability to use prosodic features accurately is mastered in early childhood. However, research to date has produced conflicting evidence, notably about the development of children's ability to mark prosodic boundaries. This paper investigates (i) whether, by the age of eight, children use temporal boundary features in their speech in a systematic way, and (ii) to what extent adult listeners are able to interpret their production accurately and unambiguously. The material consists of minimal pairs of utterances: one utterance includes a compound noun, in which there is no prosodic boundary after the first noun, e.g. '*coffee-cake* and tea', while the other utterance includes simple nouns, separated by a prosodic boundary, e.g. '*coffee*, *cake* and tea'. Ten eight-year-old children took part, and their productions were rated by 23 adult listeners. Two phonetic exponents of prosodic boundaries were analysed: pause duration and phrase-final lengthening. The results suggest that, at the age of 8, there is considerable variability among children in their ability to mark phrase boundaries of the kind analysed in the experiment, with some children failing to differentiate between the members of the minimal pairs reliably. The differences between the children in their use of boundary features were reflected in the adults' perceptual judgements. Both temporal cues to prosodic boundaries significantly affected the perceptual ratings, with pause being a more salient determinant of ratings than phrase-final lengthening.

1 Introduction

Prosodic boundaries are used by speakers of English to segment speech for a range of grammatical and pragmatic purposes, e.g. to construct a list or to end a turn in talk. Research into marking prosodic boundaries in adult English reveals a complicated picture. This poses an interesting question whether children manage to convey the presence of prosodic boundaries to listeners, and – if they do – how they achieve this. In this paper we address these questions through the study of eight-year-old children.

Journal of the International Phonetic Association (2004) 34/1
DOI:10.1017/S0025100304001525

As a basis for our investigation of children's ability to mark prosodic boundaries, we first review the current state of knowledge about the most important phrase boundary markers in adult speech.

1.1 Phonetic realisation of prosodic phrase boundaries in adult English

Probably the most widely recognised prosodic phrasal unit is the intonation phrase. However, a number of researchers have pointed out that a single level of prosodic phrasing is inadequate, as it fails to capture the observation that not all prosodic phrase boundaries are equivalent. Various proposals have been put forward to allow for two levels of prosodic phrase structure. Within the American tradition we encounter a distinction between the intonation phrase and the intermediate phrase (e.g. Beckman & Pierrehumbert 1986), while in the British tradition we find a similar distinction between minor and major tone groups (e.g. Trim 1959), or the major phrase and tone group (Ladd 1986). One of the main reasons for postulating two levels of intonational phrasing has been the observation of different degrees of cohesion between some phrases (Grabe 1998). Another reason, partially related to the previous one, has been an observation of a mismatch that sometimes occurs in the distribution of melodic as opposed to rhythmical/temporal boundary markers (Ladd 1986). In more recent literature, there have also been suggestions of a richer prosodic hierarchy, a recursive intonational structure (Ladd 1986, Ladd & Campbell 1991). However, the whole issue about the hierarchy of prosodic phrases remains controversial.

As far as the quantitative research on markers of prosodic boundaries is concerned, the focus seems to have been almost exclusively on the intonation phrase. In the section below we review the main literature relating to the boundary markers of the intonation phrase and come back to the issue of the 'lower-level' phrasing later.

In general, intonation phrases are defined both in terms of their internal intonational structure and in terms of the location of phonetic boundaries (e.g. Ladd 1986) or, as Cruttenden (1997) calls them, 'internal' and 'external' criteria. Internal features involve mainly pitch characteristics, among them particularly 'coherent intonation contour' (Chafe 1987, Schuetze-Coburn, Shapley & Weber 1991) and declination (Ladd 1986). External features, or boundary markers, include both tonal characteristics (e.g. Stenström 1990, Bruce, Granström & House 1991), such as phrase-initial pitch reset (Schuetze-Coburn et al. 1991), pitch movement in the nucleus and in unstressed syllables following it (e.g. Cruttenden 1997), and temporal characteristics, namely, the presence of a pause, phrase-final lengthening and acceleration of tempo on initial unstressed syllables (e.g. Cruttenden 1997). However, it has been suggested that phrase-final lengthening may, in fact, be a part of a rhythmical pattern extending over the whole intonation phrase. Dankovičová (1999, 2001) demonstrates a regular pattern of rallentando within the intonation phrase, and hypothesises that this pattern might parallel the role of declination in serving as another internal criterion for delimiting intonation phrases.

Boundary markers may also involve changes in amplitude (e.g. Streeter 1978) and some aspects of voice quality, such as laryngealisation (e.g. Knowles 1991).

Pause (either filled or unfilled) is the most commonly mentioned feature. However, there has been considerable variation in the literature in deciding what the minimum duration is for a gap in speech to be considered as an encoding (cognitive) phenomenon, as opposed to an articulatory phenomenon (such as stop closure). In their literature overview, Kowal, Wiese & O'Connell (1983) show that the most usual value accepted has been between 200 and 300 ms. However, the tenability of such a high cut-off point is challenged in Heike, Kowal & O'Connell (1983), who demonstrate that pauses shorter than 250 ms are not primarily articulatory in origin and that pauses between 130 and 250 ms relate to psychological and textual factors. The value of 100 ms is the cut-off point used by Butcher (1981), who observes that breaks under 100 ms are mainly stop closures. Dankovičová's (1992) data, based on distributional analysis of pauses of different duration, suggest a minimum pause duration of 130 ms. A further complication is that the perception of pauses is affected by their location: breaks in

the acoustic signal have to be much longer at the boundary before they are perceived as pauses because listeners EXPECT pauses at phrase boundaries, but not within phrases (e.g. for English, Boomer & Dittmann 1962; for German, Butcher 1981); and conversely, sometimes a boundary pause may be identified even where there is no break in the acoustic signal: Butcher (1981) notes that this can happen when above average phrase-final lengthening precedes the perceived break.

Phrase-final lengthening, often also called 'final-syllable lengthening', is another phrase-boundary marker. The latter term may be somewhat misleading since it has been shown repeatedly that most boundary-related lengthening effects are concentrated on the rhyme of the phrase-final syllable, with the final-syllable codas showing proportionally more lengthening than final-syllable nuclei (e.g. Turk 1999). Phrase-final lengthening appears to extend to the syllable preceding the phrase-final syllable if this syllable bears primary lexical stress, with the left edge of this syllable's rhyme acting as a barrier to further spread (Turk 1999). As far as we know, very little research has been conducted on how much lengthening is needed for the signal to be perceptually discernible. With respect to segmental duration, Klatt & Cooper (1975) found a minimum 'just noticeable difference' (JND) of 25 ms for segments of about 100 ms in duration, when these segments were placed in a carrier sentence in a discrimination task. However, listeners expected a longer duration in phrase-final position and thus JND was noticeably larger for the phrase-final position.

The main pitch characteristics of phrase boundaries are pitch movement in the nucleus and in unstressed syllables following it (e.g. Cruttenden 1997), and phrase-initial pitch reset (Schuetze-Coburn et al. 1991), related to declination. In English, the phrase-final pitch movement generally involves falling tones in sentence-final intonation phrases when the sentence is a statement, but they are also used in *wh*-questions and imperatives. Rising tones (low or high rise and fall-rise) or mid-level tone are commonly used in sentence non-final intonation phrases and yes/no questions (Cruttenden 1997). Experiments in pitch discrimination indicate that JND for pitch is about 6% (Rosen & Fourcin 1986).

In addition to the features mentioned above, some other phenomena have been claimed to function as boundary markers. Lehiste (1977), for instance, suggested that the rhythm of the phrase – an increase in the duration of a foot, relative to surrounding feet – can be used to signal the presence of a boundary. Some segmental effects can also function as cues to syntactic/prosodic structure: Scott & Cutler (1984) found that palatalisation and alveolar flapping are inhibited across phrase boundaries.

Finally there is the issue of how the boundary features are used in relation to one another. Scott (1982) claims that speakers vary in the way they mark prosodic boundaries and that boundary features can occur together, separately, or not at all. Schuetze-Coburn et al. (1991) suggest that the prototypical intonation phrase exhibits all of the main features, yet seldom are all actually present in any given instance. Related to this is a question of interaction between the features. Considering that phrase-final lengthening usually occurs together with pause and pitch movement, Cruttenden (1997) suggests that lengthening may actually be a by-product of these other features; it may occur as a pause substitute or the final syllable may be lengthened to carry the final pitch movement. This is, however, debatable. Although there is some evidence for an interaction between phrase-final lengthening and pauses (e.g. Fant, Kruckenberg & Nord 1991), others have not confirmed this (e.g. Horne, Strangert & Heldner 1995).

In terms of perception of boundaries, Schuetze-Coburn et al. (1991) claim that the relative importance of the features may differ. They further claim that no feature alone defines an intonation phrase boundary *per se* – rather, a conjunction of features is usually required for an intonation phrase to be perceived. Contrary to this, Scott (1982), in her experiment on syntactically ambiguous sentences, shows that the duration of the pause alone or the combined duration of a pause and phrase-final lengthening can provide the listener with a cue to the boundary location, even in the absence of a disambiguating pitch contour. A disturbance to the rhythmic pattern of speech also affected the perception of a phrase boundary and Scott

concludes that this affected rhythm of the beats was a preferred cue when determining the presence of a phrase boundary.

Independent manipulation of durational features and pitch features is, however, challenged by Streeter (1978), who points out that both pitch and duration exist in the time domain, and pitch is also a function of fundamental frequency. She further claims that manipulation of either alone distorts the melody of the utterance, and this melody cannot be perfectly reconstructed without changing both simultaneously. She hypothesises that pitch and duration may be inseparable cues, i.e. they form some sort of integrated percept. Beach (1991) provides evidence for this. She studied interaction between a duration cue (phrase-final lengthening) and a pitch cue in the perception of syntactically ambiguous sentences. She showed that (i) both duration and pitch are important in perception of prosodic boundaries, and (ii) they are processed interactively. The nature of this interaction is characterised as cue-trading relations, i.e. duration and pitch cues are perceived together as one integrated percept; the influence of one cue is greater when the other cue is weaker. Beach also demonstrated that listeners are able to use prosodic information online to predict upcoming syntactic structure.

In summary, research suggests that a combination of phonetic features is used by speakers to convey intonation phrase boundaries and by listeners to identify them. One question that therefore arises is that of individual differences, across speakers and listeners from the same speech community, in the use and combination of these features. There is evidence of considerable variability among adults in the particular combination of features produced to mark phrase boundaries (Peppé, Maxim & Wells 2000).

As we mentioned at the beginning of this section, some current theories recognise two (or more) levels of intonational phrasing. However, the quantitative data on prosodic units/phrases smaller than the intonation phrase are, to our knowledge, mostly lacking. An exception to be mentioned is an investigation of phrase-final lengthening in prosodic phrases at different levels of prosodic hierarchy by Ladd & Campbell (1991). This study demonstrates significant differences in phrase-final lengthening, correlating well with boundary strength. Quantitative research on pauses and pitch patterns in intermediate phrases/minor tone groups still awaits attention, as does research comparing the boundary markers in intonation phrases and intermediate phrases.

1.2 Prosodic boundaries in children's linguistic development

While a great deal of research has been carried out into children's phonetic and phonological development, this has mainly focussed on segmental aspects and on early childhood (Vihman 1996). It is sometimes assumed that learning to use prosodic features accurately is relatively unproblematic. However, the picture seems more complex. There is indeed some evidence, reviewed below, that shows children by the middle of the third year using adult-like features to mark intonation phrase boundaries prosodically. However, research on older children's marking of intermediate phrase/minor tone group boundaries (e.g. Katz, Beach, Jenouri & Verma 1996) suggests an as yet undeveloped ability to mark these types of prosodic phrases in an adult-like way. Thus, it seems possible that there are developmental differences in the ability to mark prosodically different levels of prosodic hierarchy.

Snow (1994) investigated the development of prosodic boundary production through a longitudinal study of the spontaneous speech of children embarking on multiword speech (16–25 months). Snow focussed on falling pitch and final lengthening at the end of utterances (i.e. at the end of intonation phrases) and found that pitch fall on final syllables was consistently greater than for non-final syllables. In the case of final lengthening, a durational difference between final and non-final syllables was only found towards the end of the period studied, once the children were producing combinatorial speech consistently. Snow suggested that, by the age of two, children are able to control both the pitch parameter and final lengthening appropriately to signal the presence versus absence of a phrase boundary. He further suggested that final lengthening is associated with syntactic development, since its onset coincides with the onset of grammatical combinations.

However, this evidence of early control over prosodic boundary features does not mean that children can necessarily deploy them in all appropriate contexts. Katz et al. (1996) investigated production of the contrast in prosodic phrasing that is found in the minimal pair: '[(pink and green) and white]' vs. '[pink and (green and white)]'. In other words, they examined intermediate phrases/minor tone groups. Three blocks (one pink, one green, one white) were grouped by the experimenter in different ways, and were described by the participants – groups of five- and seven-year-old children, and adults. For the adults, there was a strong effect of word lengthening and pause duration as a function of prosodic grouping. However, the children did not show the same effect: the children appeared to use neither pitch nor duration features in an adult-like way to convey grouping of objects. This result suggests that children as old as seven are unable to use prosodic cues to mark intermediate phrase boundaries in the way that adults use them, thus pointing to an association between prosodic phrasing and grammatical phrasing that has to be mastered by children learning English.

There are various possibilities as to why the children were unsuccessful on this task. One is that children of this age are unable to recognise and interpret the difference in the speech of adults; consequently they have no representation for this particular phonological distinction, and so do not signal it in their own production. However, this is unlikely in the light of the results of a parallel study into comprehension of the same contrast, by Beach, Katz & Skowronski (1996), in children aged five and seven years. Results showed that both groups of children behaved like adults in drawing on pitch and duration features to guide their interpretation. This suggests that children as young as five can use prosodic information to guide grammatical interpretation, and thus have an adequate input representation of the contrast.

Thus, in spite of the apparent ability of children as young as five to interpret adults' use of these features to indicate phrase boundary, in their own speech, children appear to use neither pitch nor duration features in an adult-like way to convey grouping of objects. It is possible, as we pointed out above, that marking smaller prosodic phrases, such as intermediate phrases in speech is more difficult for children, and that successful marking of this type of prosodic phrase is learnt at a later stage in development than marking of intonation phrases.

While one possibility is that the children were indeed failing to convey the relevant distinction in their speech, an alternative hypothesis is that they WERE conveying the distinction, but it was not picked up by the instrumental measures used by Katz et al. (1996), i.e. the children did not use the predicted prosodic features. A further possibility is that the negative group results presented by Katz et al. regarding children's use of prosodic features may conceal individual differences: that SOME of the children may in fact be using adult-like prosodic features. These issues are not resolved in the study by Katz et al. (1996), since no evidence is available of how adult listeners interpreted the children's utterances. Thus, the results point to the importance, in future research, of (i) incorporating data about adults' comprehension of children's productions, and (ii) looking at individuals, as well as at group performance.

Some evidence relating to the latter two issues is found in a recent cross-sectional study of intonation development in the school years by Wells, Peppé & Goulandris (submitted). As part of a wide-ranging battery of tasks targetting different aspects of intonation ability, Wells et al. systematically investigated the production by groups of children, aged 5, 8, 10 and 13, of contrasts of the type: '/*coff*ee-cake /and *ho*ney/' vs. '/*coff*ee /*cake* /and *hon*ey/'. Here, prosodic phrasing serves to make a lexico-grammatical distinction between the compound noun, *coffee-cake*, and two successive nouns, *coffee* and *cake*. In terms of intonational phrasing, the former would be interpreted as two and the latter as three intermediate phrases/minor tone groups. The tasks involved picture naming, in which the child was assessed on whether or not s/he could realise his/her communicative intention. Test scores were based on perceptual judgement by trained phoneticians.

There was no statistically significant increase in group scores across age bands (though there was a non-significant trend showing improvement with age). In all age groups some

children were at ceiling while others did not score above chance. This suggests that some children are able to mark prosodic boundaries in such a way as to convey the correct meaning by the age of five. Conversely, there are children throughout the age range who do not consistently use the expected pattern. In general, the study by Wells et al. provides conflicting evidence to Katz at al. (1996) as to the ability of children to mark intermediate phrases/minor tone groups prosodically.

1.3 Aims of the present study

As our review of the literature shows, it is not clear how successfully children mark prosodic boundaries in their speech production. Possible reasons for this are not hard to find. Firstly, several prosodic features of phrase boundary have been identified in adult speakers, as discussed above, but how they interact, in terms of presence and magnitude, to convey phrase boundary is not well-established. Instrumental measurements may determine that relevant features are present in a child's utterances (e.g. pause), or occur in the right direction (e.g. longer rather than shorter final syllable duration before prosodic phrase boundary than elsewhere), but it does not, on its own, show how or whether the features impact on a listening adult's comprehension; in other words, whether their magnitude is sufficient for the listeners to perceive the boundary. However, it may not be necessary for children to be using prosodic boundary features in an adult-like way (i.e. both in the right direction and magnitude) for adults to interpret them correctly; the adults may, consciously or unconsciously, make allowances for incomplete command of prosodic production. Secondly, the prosodic features used for boundary signalling are also associated with the completion of a turn at talk, by young children (Corrin, Tarplee & Wells 2001) as well as by adults (Wells & Macfarlane 1998); the actual role these features play may therefore vary from study to study unless controlled for, and results may conflict because they are reporting on different communicative functions.

In this study, we attempt to find out whether eight-year-old children mark prosodic (intermediate phrase) boundaries to make the distinction between simple and compound nouns and if so, how they do it. We have chosen eight-year-old children mainly because the studies by Katz et al. (1996) and Wells et al. (submitted) suggest that not all children of this age have mastery of marking this kind of prosodic boundary. Our study is based on material used in the study by Wells et al., but uses acoustic measurements as well as perceptual scores. It aims to further our knowledge from Wells et al.'s study by (i) analysing the nature of variability in children's boundary marking, and (ii) investigating the role of certain phonetic parameters in perception of prosodic boundaries by adults.

As we have seen, research into adult English demonstrates that the role of each of the postulated prosodic phrase-boundary features remains unclear, as is their interaction. This applies to all levels of prosodic boundary.

We have limited our study to the consideration of two temporal features: pause and phrase-final lengthening, bearing in mind that (i) other features, particularly aspects of pitch, may have a role to play, and (ii) in a child's output, incomplete control of one feature may result in a signal which is at odds with that of another feature, but that the combination may yet be sufficient to convey the message to a listening adult.

The following research questions are addressed:

i. Do eight-year-old children use pauses and phrase-final lengthening to mark intermediate phrase boundaries?
ii. When listening to eight-year-old children, do adults use children's production of these prosodic features to guide their interpretation of intermediate phrase boundary location?
iii. To what extent do eight-year-old children mark intermediate phrase boundaries both accurately (i.e. in accordance with the target) and unambiguously (i.e. in such a way that adults correctly identify the phrase boundary location intended by the child)?

2 Method

2.1 Subjects

Ten children (five boys and five girls) were randomly selected from a group of thirty normally-developing eight-year-olds who had been tested as part of a study of prosodic abilities in school-age children (Wells et al. submitted). The children had no behavioural, learning or hearing difficulties, no identified speech and language problems and had English as their first language and the language of the home. They all fell within the normal range on the following tests: the Test for Reception of Grammar (TROG; Bishop 1989) for grammatical comprehension, and the sentence formulation subtest of the Clinical Evaluation of Language Fundamentals – Revised (CELF-R; Semel, Wiig & Secord 1987) for expressive language.

2.2 Material

The speech samples analysed consist of responses to one of the production tasks of the Profiling Elements of Prosodic Systems – Child version (PEPS-C) as devised for the study by Wells et al. (submitted). The task tests the child's ability to mark prosodic phrase boundaries in his or her own speech production.

For each test item, the child had a strip of coloured pictures. The picture-strips showed either two or three items of food, and children were asked to tell the tester what was on their picture-strip. Before the test, a vocabulary check was carried out: the children were shown all the food-item pictures and it was established that they knew what each one was to be called. Practice items were given, and the task took no longer than five minutes in total. Children's responses were recorded via tabletop microphone on DAT tape. In 2-item picture-strips, the first picture shows a food-item designated by a compound noun, e.g. *chocolate-biscuits*: there is, thus, unlikely to be a prosodic boundary between the first two words. The second picture (third word) is a separate food-item (e.g. *honey*). In 3-item picture-strips, three pictures and three separate nouns were involved, to elicit a prosodic boundary between the first two, e.g. *chocolate*, *biscuits* and *honey*.

A number of picture-strips were presented to each child in random order. We then selected those children whose trial included at least five minimal pairs. (The term 'minimal pair' refers here to pairs of utterances in which the three nouns are the same but one utterance contains a prosodic boundary between the first two nouns and the other does not, e.g. *milk, bottles and bread* vs. *milk-bottles and bread*.) In order to have a balanced set of data, we subsequently selected five minimal pairs in total for each child. These pairs were not necessarily identical across children. There were nine different minimal pairs in total (see the appendix), from which the five minimal pairs per child came. This yielded 100 utterances in total for the measurements.

2.3 Measurements

The recorded speech material from all children was digitised at a 44.1 kHz sampling rate, with 16-bit resolution and then segmented into individual utterances. Five utterances were excluded from the analysis since they contained hesitation, either in the form of an extremely long pause (three utterances) or a false start (two utterances). There were thus 95 utterances in total (46 minimal pairs and three unpaired utterances).

As mentioned in the Introduction (section 1), we examined two boundary features: the pause duration between the first and second noun, and the final syllable duration of the first noun. When this was a monosyllable, the duration of the whole noun was measured.

Pause durations and final syllable durations were measured with reference to sound pressure waveform and wideband spectrogram. When the second word began with a plosive, pause duration included the stop closure up to the burst, as it is generally impossible to determine the beginning of the closure after a pause. For the same reason, the syllable

duration of monosyllabic first words starting with a plosive was measured from the beginning of the burst.

2.4 Perceptual test

Two DAT tapes for perceptual presentation were created. On both tapes, all the utterances were randomised and each utterance was separated by three seconds of silence. The tapes differed only in the order of presentation – on one tape the utterances were in the reverse order. The purpose of this was to control for order effects.

Rating sheets were produced for the listeners, based on the two orders of presentation. A standard format was used with no punctuation marks, to avoid cueing the listeners. A 6-point rating scale was used with the following options (the numbers refer to the number of items of food):

DEFINITELY 2 PROBABLY 2 POSSIBLY 2 POSSIBLY 3 PROBABLY 3 DEFINITELY 3

Using a rating scale, rather than a binary decision, made it possible to measure a degree of listeners' certainty. Having an even number of ratings, rather than an odd number, forced listeners to make a decision and did not allow for a 'don't know' option.

Twenty-three listeners took part in the perceptual test. They were between 20 and 37 years old (mean age 23.6 years) and were undergraduates at University College London. All had English as their first language and had normal hearing. Subjects listened to the tape via headphones. The instructions asked them to decide how many items of food there were in the utterances and circle their choice on the rating sheet. The task supervisor provided a verbal example of what a 'DEFINITELY 2' and a 'DEFINITELY 3' utterance might sound like. The first item on the tape was used as an example to provide an opportunity for questions to be asked and the tape was then played, without stopping, to the end. The whole task took about 20 minutes. The subjects were naïve with respect to the purpose of the study. Thirteen listeners heard the presentation in its original order and ten listeners heard the reverse order.

3 Results

3.1 Children's production of prosodic boundary features

This part of the analysis concerns the question whether the children use pauses and phrase-final lengthening to mark prosodic phrase boundaries. In terms of statistical analysis, we attempt to establish whether the two temporal prosodic boundary features are significantly affected by the target. Repeated-measures analysis of variance (ANOVA) was used for the features separately, with the within-subject factor being 'target' (2- or 3-item). Thus the primary concern here is the presence/absence of a feature and/or the right direction of use (see the final section of the Introduction).

3.1.1 Pause duration

Since presence of a pause typically accompanies prosodic phrase boundaries, we would expect there to be no pause between the first and the second item in a list when these form a compound noun (i.e. in a 2-item utterance) and we would expect a pause when the two items are separate nouns in a list (i.e. in a 3-item utterance). It should be noted that, as we had to include stop closures in our pause measurements (see section 2.3 above), articulatory pauses of a certain duration were expected in those 2-item utterances which included stop consonants at the measurement point (i.e. at the end of the first noun and/or at the beginning of the second noun, e.g. 'chocolate-cake'). Although we mentioned a cut-off point of 130 ms between 'true' (cognitive) pauses and articulatory pauses in the Introduction, we did not consider it necessary to employ this cut-off point here for two reasons. Firstly, the focus of this part of the analysis was to see whether the children employ the feature in the right

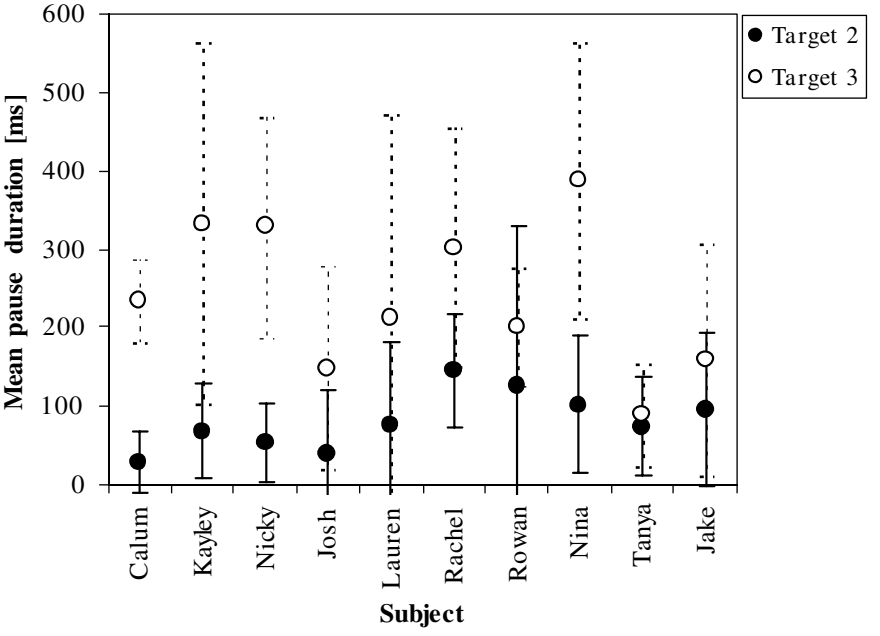

Figure 1 Mean pause durations (and standard deviations) in target-2 and target-3 utterances for individual subjects.

DIRECTION. Thus, the important issue is the RELATIVE behaviour in 2- vs. 3-item utterances. Secondly, the value of the minimum pause duration in children's speech may be different, reflecting, for example, likely slower speed of speech production in comparison with adults. We are not aware of any study investigating the minimum pause duration in children's speech.

In terms of the analysis of variance, we would expect a longer pause in target-3 than in target-2 utterances. The ANOVA analysis confirmed this tendency for the dataset as a whole – the 'target' was a highly significant factor [$F\,(1,45) = 32.2$, $p < 0.0001$]. The model accounts for 42% of the variance in pause duration.

Results for individual children can be seen in figure 1, which displays the mean pause durations and standard deviations with respect to target.

Figure 1 shows that although the expected pattern of a longer pause in the presence of a prosodic boundary was found, for some children there is considerable variation from the mean (as indicated by large standard deviations) and/or an overlap in the distribution of pause durations between target-2 and target-3 utterances. Only in the utterances produced by Calum, Nicky and Nina is the distribution clearly separated for the two targets. In most cases, the overlap reflects the fact that the pause durations were only marginally greater in target-3 than in target-2 utterances. For one child (Tanya) some minimal pairs had in fact the reverse pattern (pause in target-3 SHORTER than in target-2 utterances). This is an instance of a feature being used in the wrong 'direction'. The utterances containing these patterns may be expected to cause confusion in the perceptual test.

3.1.2 Final syllable duration

The expected pattern is for the final syllable in the first item in the list to be longer when a phrase boundary follows (target-3 utterances) than when a prosodic boundary is absent (target-2 utterances). This trend was confirmed for the dataset as a whole. The ANOVA revealed a highly significant effect of 'target' [$F\,(1,45) = 45.2$, $p < 0.0001$]. The model accounts for 50% of the variance in final syllable duration (FSD).

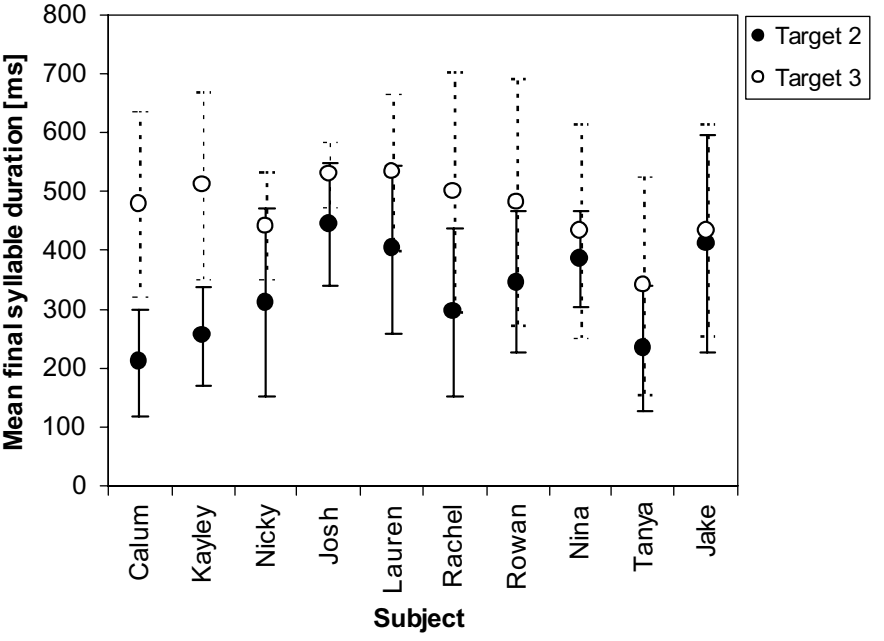

Figure 2 Mean final syllable durations (and standard deviations) in target-2 and target-3 utterances for individual subjects.

Figure 2 shows mean FSD values and standard deviations for individual children. Overall, the expected pattern is illustrated for all children. However, as for pause duration, the FSD data show large variations from the means, with considerable overlap for all children apart from Calum and Kayley, whose FSD distributions are separate for target-2 and target-3 utterances. For four children (Lauren, Rowan, Tanya and Jake) the values for target-2 and target-3 utterances were very close in some utterances though still in the right direction. In these utterances FSD is unlikely to be a very strong distinguishing feature in the perceptual test. Several minimal pairs (particularly in Jake's production) involved FSD being used in the wrong direction (i.e. FSD in target-3 SHORTER than in target-2 utterances).

In summary, the analysis shows that the expected pattern in relation to target was highly significant for both of the boundary features that we examined. Thus, the first part of the analysis suggests that, OVERALL, children's use of temporal boundary features is in the expected direction. However, a more detailed analysis revealed a considerable degree of variation in the data and, for most children, a noticeable overlap in the distribution between the features in target-2 utterances and target-3 utterances, partially caused by some minimal pair utterances involving either or both features being used in the wrong direction. This, together with the closeness of the mean values in target-2 and target-3 utterances, suggests a source of possible confusion in perceptual judgements. Perceptual salience of the boundary features will be considered in the next two sections.

3.2 Adults' comprehension of children's production

The second question we address is whether adults use children's production of the temporal boundary features to guide their interpretation of phrase boundary location. We investigate to what extent they are able to perceive unambiguously the presence/absence of a prosodic boundary. The issue of magnitude comes into focus here. As we mentioned earlier, it is possible that a child might produce boundary features in the right direction, but with magnitude insufficient for the adults to perceive the boundary.

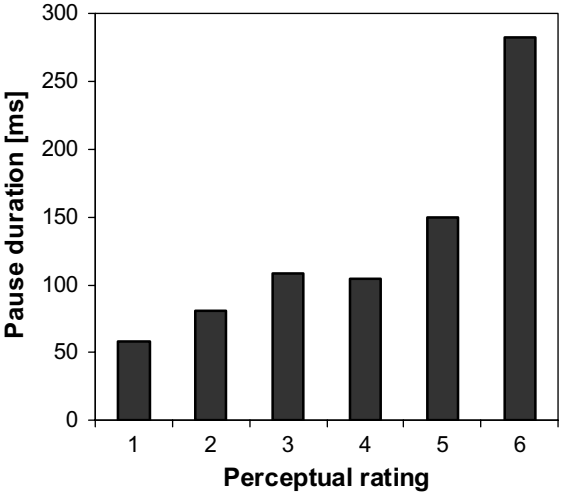

Figure 3 Median pause duration per individual ratings across the listeners (1 = definitely 2-item utterance; 6 = definitely 3-item utterance).

In terms of statistics, we are interested in finding out (i) whether listeners' judgements are affected significantly by both, or only one of the prosodic features investigated, and (ii) if the former is the case, what the relative importance of these prosodic features is for the adult perceptual judgements. To make the statistical analysis possible, the judgements were coded numerically as follows:

1	2	3	4	5	6
DEFINITELY 2	PROBABLY 2	POSSIBLY 2	POSSIBLY 3	PROBABLY 3	DEFINITELY 3

Prior to the main statistical analysis of the perceptual judgements we examined the cross-subject consistency in perceptual judgements for individual utterances and all 23 adult listeners. A very good overall level of agreement was found. The coefficient of concordance (Kendall's W) was 0.74, $N = 95$, $p < 0.0001$. This coefficient can be viewed as a function of the average Spearman correlation (Howell 2001). In our case its corresponding value to the Kendall's W is 0.73.

In order to see a more detailed picture of the consistency of rating between the listeners, we also conducted Spearman's correlation analysis, separately for target-2 and target-3 utterances, correlating each listener's scores with every other listener's scores. The results showed that the absolute majority of the pairwise correlations were highly significant ($p < 0.0001$). There was only one listener who clearly stood out. His ratings in target-2 utterances differed from most other listeners (resulting in a non-significant correlation). This listener was therefore excluded from the further analysis of the perceptual ratings.

We predicted that the longer the pause and the final syllable duration, the more likely the listeners would perceive the prosodic boundary (i.e. judge that the utterance was 3-item rather than 2-item). In order to examine the relationship between the ratings and the two boundary markers, the median values per each rating category across the listeners were calculated for both pause and FSD. Figures 3 and 4 present the results.

Overall the graphs confirm the expected tendency. This tendency is particularly clear for pause duration, where the gradual increase in median pause duration with the increasing rating is evident throughout the whole rating scale. For FSD, the middle ratings (2–5) show little distinction from each other.

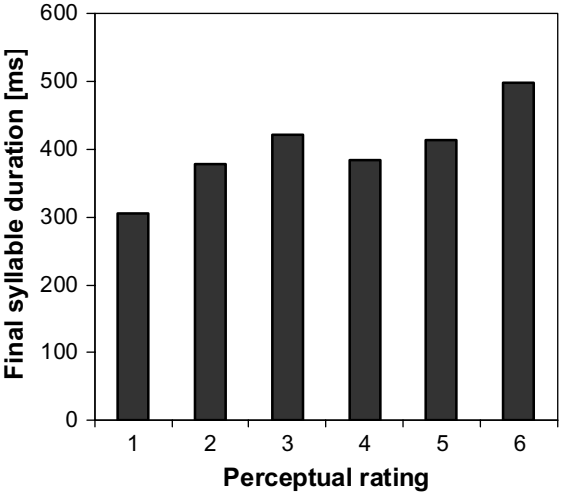

Figure 4 Median final-syllable duration per individual ratings across the listeners (1 = definitely 2-item utterance; 6 = definitely 3-item utterance).

Figures 3 and 4 suggest that the listeners' judgements were in a clear relationship with the two parameters. Further analysis was carried out in order to establish whether the contribution of the two boundary features to the perceptual ratings was statistically significant. We used a multinomial logistic regression, as this type of regression analysis allows a categorical variable with more than two levels to be used as the dependent variable. In our analysis the dependent variable was the frequency of each rating category across 22 listeners and the independent variables were pause duration and final syllable duration. The analysis was conducted using the statistical package SPSS, version 11. The results showed that both parameters were highly significant predictors of the perceptual ratings, but pause duration was a stronger predictor than FSD; for pause duration χ^2 (df 5) = 936.4, $p < 0.001$; for FSD χ^2 (df 5) = 396.4, $p < 0.001$.The model accounts for 54% of variance in the data, as indicated by the coefficient of determination – pseudo-R^2 [Nagelkerke].

3.3 Individual children's efficiency in marking prosodic boundaries

In the first part of the analysis we established that, on average, the children performed well, i.e. as a group they seemed relatively accurate in realising the target – both of the temporal boundary features were mostly used in the expected direction (longer FSD and pause when a prosodic boundary was present). The second part of the analysis suggested that the adults used these features to guide their perceptual judgements of prosodic boundaries.

In this section we consider the data in more detail and examine the degree of prosodic 'efficiency' of individual children. By 'prosodic efficiency' we understand a combination of accuracy (presence/correct direction of use of the features) and perceptual unambiguousness (sufficient magnitude of used features for correct perceptual judgements). The reason for analysing the results of the individual children is that statistical analysis shows mainly general trends and may leave some more subtle but interesting individual differences hidden.

The graphs in figure 5 show, for each child individually, the frequency distribution of perceptual ratings according to the intended target. Since point 1 on the rating scale represents 'definitely 2 items' and point 6 represents 'definitely 3 items', the perceptual judgements should be 100% for 1-rating of target-2 utterances and 100% for 6-rating of target-3 utterances if the child performed the task perfectly.

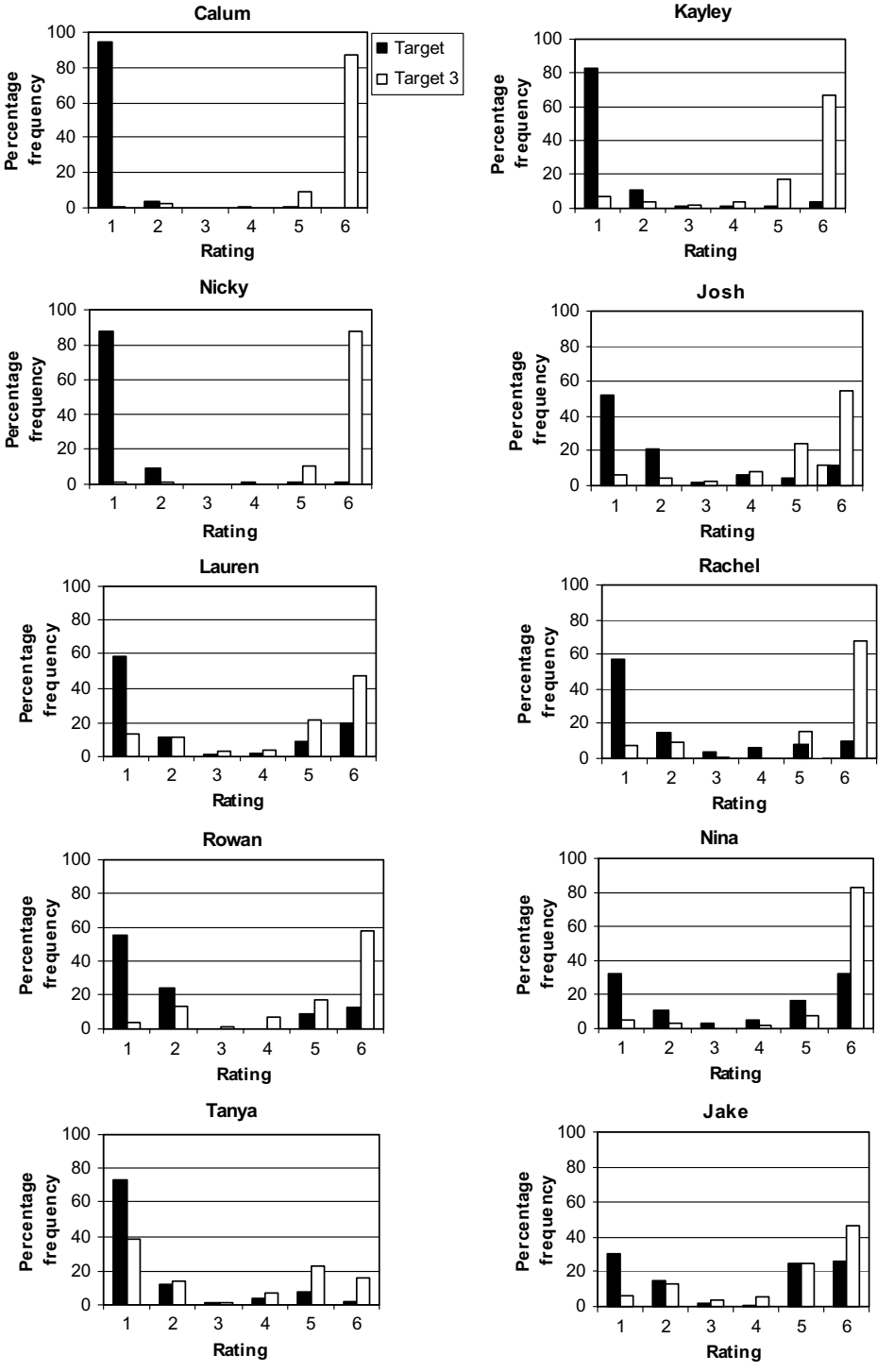

Figure 5 Frequency distribution of perceptual ratings for individual children.

Table 1 Mean values for pause duration and final syllable duration (FSD) for individual children (values from figures 1 and 2).

		2-item utterances		3-item utterances	
Group	Subject	Mean pause duration [ms]	Mean FSD [ms]	Mean pause duration [ms]	Mean FSD [ms]
1	Calum	28.7	209.6	234.0	479.3
	Kayley	67.7	253.8	333.5	510.0
	Nicky	52.2	312.0	328.9	442.2
2	Josh	40.3	444.9	148.4	529.4
	Lauren	75.6	402.4	211.5	533.2
	Rachel	144.3	295.9	302.5	499.6
	Rowan	126.5	346.0	234.0	483.1
3	Nina	101.1	384.3	387.3	433.5
	Tanya	73.4	233.0	88.2	340.2
	Jake	95.3	411.6	157.9	434.5

The following observations can be made on the basis of figure 5. Three children – Calum, Kayley and Nicky – were very efficient in their prosodic ability. They are closest to the 'perfect case'. Most of their utterances were produced according to the required target and in a perceptually unambiguous way, i.e. with a 1-rating for over 80% of 2-item utterances and/or a 6- rating for over 80% of 3-item utterances. Four other children – Josh, Lauren, Rachel and Rowan – were also able to contrast their 2-item and 3-item utterances, but not as efficiently as the previous group: unambiguous ratings were used for 50–70% of their utterances. The remaining three children differed from each other and from the rest in the rating distribution of their utterances. Nina excelled at producing unambiguously perceived 3-item utterances (over 80% were judged as 'definitely 3 items'), but of her 2-item utterances only 33% were rated as 'definitely 2 items' and a further 33% were 'definitely 3 items', indicating that Nina was poor on 2-item utterances. For Tanya, the pattern of rating distribution is opposite to Nina's: over 70% of her target-2 utterances but also nearly 40% of her target-3 utterances were rated as 'definitely 2 items'. A mere 16% of target-3 utterances were perceived as 'definitely 3 items'. Generally, Tanya manifests a slight bias towards 2-item utterances, i.e. less marked prosodic boundaries. Jake, unlike Tanya or Nina, showed no particular bias either to 2-item or 3-item utterances, but he was rather poor on both targets. Only 31% of his target-2 utterances were judged correctly and unambiguously as 2-item utterances and 46% of his target-3 utterances as 3-item utterances.

In general, we can conclude that all the children were able to produce relatively unambiguous utterances in terms of the presence/absence of prosodic boundary: no child had a majority of perceptual ratings of 3 or 4. Moreover, most of them produced prosodic boundaries appropriate to the target. Although there were no cases where there was a complete switch between 2- and 3-item utterances (i.e. all 2-item utterances judged unambiguously as 3-item utterances and vice versa), two children (Nina and Tanya) seemed to be unable to contrast the two well and most utterances they produced were perceived as EITHER containing a prosodic boundary (3-item – Nina), OR not containing a boundary (2-item – Tanya).

In the final section we shall return to the prosodic features and examine, for the groups of children discussed above, what prosodic features (individually or in combination) may account for the distribution of the perceptual ratings. For this purpose, we shall refer to the mean values and distribution of the temporal features as displayed in figures 1 and 2 above, and as detailed in table 1.

GROUP 1: MOST EFFICIENT BOUNDARY MARKING (Calum, Kayley and Nicky)
Figures 1 and 2, and table 1 show very clear differentiation for target-2 and target-3 utterances for both pause duration and final syllable duration (with the exception of Nicky's FSD).

GROUP 2: LESS EFFICIENT BOUNDARY MARKING (Josh, Lauren, Rachel and Rowan)
For two children in this group (Josh and Lauren) pause duration in their target-2 utterances is on average very short and in their target-3 utterances about 3–4 times longer, yet noticeably shorter than for Group 1. Final syllable duration for these two children's target-3 utterances is similar to Group 1 values, but longer than for Group 1 in target-2 utterances, making the difference between target-2 and target-3 utterances less marked than in Group 1.

The other two children in this group (Rachel and Rowan) produced pause duration and FSD in target-3 utterances and FSD in target-2 utterances close to the values of Group 1, but the mean pause duration in their target-2 utterances was noticeably longer than in Group 1, perhaps thus contributing to the perception of a prosodic boundary.

Although the two pairs of children in Group 2 produced a similar pattern of perceptual rating, there seems to be a trading relationship between pause duration and FSD, with Rachel and Rowan relying mainly on short FSD in target-2 utterances and Josh and Lauren on the absence of a pause.

GROUP 3: LEAST EFFICIENT BOUNDARY MARKING (Nina, Tanya and Jake)

Nina
As we pointed out, perceptual judgements of the vast majority of Nina's target-3 utterances were correct and unambiguous. Factors which might contribute to this are her markedly long pauses (even longer than in Group 1) and appropriately long FSD (similar to that of Group 1). Nina's target-2 utterances were interesting in that over half of them were judged wrongly, but unambiguously, as 3-item utterances. In these utterances Nina's pause duration is longer than the Group 1 values and longer than those of Josh and Lauren. It was shorter than Rachel's pause duration (Group 2), but while Rachel and Rowan (also Group 2) compensated for longer pause duration in target-2 utterances with shorter FSD, Nina did not.

Tanya
Tanya's production seems to have a bias towards 2-item utterances. Pause duration was relatively short in both targets, and FSD in her target-3 utterances was markedly shorter than for Group 1 (in fact, shorter than any other child's). This is likely to account for the bias. Her good performance on target-2 utterances can be explained by these same features: lack of pause and short FSD.

Jake
Jake's distribution of perceptual ratings for target-2 utterances is as poor as Nina's, reflecting values for pause duration and FSD similar to hers (for both children the main problem seems to be too long FSD; for Jake the mean FSD for target-2 and target-3 is nearly the same). His target-3 utterances are, unlike Nina's, poor as well. The responsible factor for this appears to be too short pause duration in some utterances.

The implications of these groupings will be discussed further below.

4 Discussion

In this study we set out to address three research questions. We will now consider each question in turn, in the light of the results.

i. Do eight-year-old children use pauses and phrase-final lengthening to mark intermediate phrase boundaries?

Two candidate prosodic boundary features (pause duration and phrase-final lengthening) were analysed, in order to establish whether their occurrence is determined by the target (i.e. 2-item

or 3-item utterance). For both features there was a significant effect of target – pause and longer final syllable duration tended to occur in those utterances in which a prosodic boundary was expected (target-3), as opposed to the utterances without a prosodic boundary (target-2). This result indicates that, when the children were analysed as a group, they seemed to use the features in the expected direction to mark prosodic boundaries.

However, a more detailed analysis showed that, for some children, the means for target-2 and target-3 utterances were close together and, also, there was a large standard deviation, indicating that their use of these features varied considerably across different items in the test. Moreover, some utterances in some of the children proved to have a reverse pattern, suggesting that eight-year-old children are not consistent as a group in the use of these prosodic features across individual utterances, and that some children are more consistently accurate than others. The statistical analysis was not affected by the presence of these reverse patterns as their number was relatively small and the actual durational differences involved were also relatively small.

It has been widely reported that adults reliably use pauses and phrase-final lengthening to mark intonation phrase boundaries (e.g. de Pijper & Sanderman 1994), and the rare studies on intermediate phrases/minor tone groups (Katz et al. 1996) report similar findings. Our detailed analysis suggests that eight-year-old children's prosodic production abilities may not yet be completely adult-like. That said, our results suggest a more positive picture than Katz et al. (1996) did for seven-year-olds, who, according to their analysis, did not use adult features at all.

It should be pointed out that our observations about the variability in the use of temporal boundary features among children are to be treated as tentative since they are based on a limited dataset. The degree of variability in children's marking of prosodic boundaries remains to be verified in future research. Similarly, comparison between the variation in adults' and children's use of pauses and phrase-final lengthening needs to be investigated further, since, to our knowledge, no quantitative data are available on the degree of variability in the adult production of prosodic boundaries, based on material similar to ours.

ii. When listening to eight-year-old children, do adults use children's production of these prosodic features to guide their interpretation of intermediate phrase boundary location?

The listeners were overall highly consistent with each other in their perceptual judgements of the prosodic boundary. Both pauses and final-syllable duration related to the adult perceptual ratings in the predicted way, i.e. the longer the pause and FSD, the more unambiguous the perception of the prosodic boundary. They were both highly significant predictors of the perceptual ratings, but pause duration seemed to be a stronger cue than FSD.

The finding that both pauses and phrase-final lengthening have a significant effect on the perception of prosodic boundaries is in line with the literature on adult perception of adult speech (e.g. Streeter 1978, Scott 1982, Beach 1991, de Pijper & Sanderman 1994, Sanderman & Collier 1997).

iii. To what extent do eight-year-old children mark intermediate phrase boundaries both accurately (i.e. in accordance with the target) and unambiguously (i.e. in such a way that adults correctly identify the phrase boundary location intended by the child)?

Effective communication requires that the child be able to mark prosodic phrase boundaries by using appropriate phonetic resources in the right direction with respect to target (Question i), and, at the same time, in sufficient magnitude for the listeners to be able to interpret those features unambiguously as boundary markers (Question ii). The analysis of the results for the individual children indicated that they fell into three subgroups:

a. ACCURATE AND UNAMBIGUOUS. This group consistently marked prosodic phrase boundaries in accordance with the target, using both pause duration and final syllable duration consistently to differentiate between 2-item and 3-item targets, and attracted mainly unambiguous ratings. It seems reasonable to conclude that the children in this group have acquired the adult system.

b. ACCURATE BUT AMBIGUOUS. While children in this group were rated as unambiguous for over half their utterances, the others were less clear-cut to listeners. Each of these children relies on just one key phonetic boundary feature, whereas the accurate and unambiguous children used both. It can be hypothesised that the 'accurate but ambiguous' children are making a relevant contrast between 2-item and 3-item targets, but that this contrast is sometimes not available to the listener. It may be that the accurate but ambiguous children are developmentally at a slightly earlier phase than the 'accurate and unambiguous' ones – a phase where the target phonological contrast is made by the speaker but is not always conveyed to the listener because a relevant phonetic exponent is missing.

There is some evidence for cue-trading between pause duration and phrase-final lengthening in the group of 'accurate and ambiguous' children, who attracted similar perceptual ratings although their phonetic exponency varied considerably. However, further research is needed to establish exactly how these two boundary features inter-relate, and how they interact with pitch factors.

c. INACCURATE AND AMBIGUOUS. The combination of inaccurate phonetics and ambiguous message suggests that the children in this group may be at a yet more immature stage of development than the children who were 'accurate but ambiguous'. Why are these children inaccurate and ambiguous when attempting to signal prosodic phrase boundaries? One possibility is that they do not all appreciate that there is a distinction to be made. In such cases, it is possible that they do not perceive this distinction themselves, when listening to adults. As the children described in the present paper were also tested on the comprehension of this distinction (see Wells et al. submitted), we were able to assess this possibility. The comprehension data suggest that the above hypothesis may hold for Jake, but it is unlikely for the other two, who performed no worse than the rest of the group. In the cases of Nina and Tanya, it therefore seems more likely that they simply have inaccurate/immature output.

5 Conclusions

The present study offers a new perspective on children's prosodic organisation, by combining objective acoustic analysis of children's productions with data on adults' perception of those productions.

With regard to prosodic organisation, we have shown that the temporal features for marking prosodic boundaries – pause duration and phrase-final lengthening – are instrumental in forcing listener's interpretation of the message. However, this does not mean that these were the only features oriented to by the adult listeners. It remains to be established whether the eight-year-old children produce boundary pitch cues, which, alongside with pause and phrase-final lengthening, are also prominent boundary markers in adult speech production. If the children produced pitch cues systematically, it is possible that the listeners in our experiment used them when judging the presence/absence of prosodic boundaries. Further analysis would also be needed to investigate possible cue-trading relationships between pauses, phrase-final lengthening and pitch cues.

We have also demonstrated that there is considerable variability among children at this age, in their use of temporal parameters to mark prosodic phrase boundaries. At the present time, it is not possible to gauge the extent to which some children are closer to the adult 'norm' than others. It seems likely that a substantial amount of the variability in the child data might be accounted for by lack of maturity, and in our discussion of subgroups of children in the present study, we have suggested three possible phases of development. However, further quantitative research is needed to establish the extent of variability in prosodic phrase boundary production in comparable structures in the adult population, since a possibility remains that some of the variability we have found among eight-year-olds parallels variability among adults.

Finally, we have provided data on the type of prosodic phrase for which there is a marked lack of quantitative research reported in the literature. Our study supports the findings that

intermediate phrases/minor tone groups are marked in a similar way as intonation phrases, at least as far as the temporal boundary markers are concerned. In regard to children, it remains to be investigated whether there are developmental differences between the ability to mark intonation phrase boundaries, as opposed to intermediate phrase boundaries.

Acknowledgements

We would like to thank the three anonymous reviewers for their helpful comments, Andrew Faulkner and Chris Donlan for their advice on statistical analysis, and Esther Grabe for pointing us to a number of useful literature sources.

Appendix

Minimal pairs

Target-2	Target-3
chocolate-icecream and honey	chocolate, icecream and honey
chocolate-biscuits and tea	chocolate, biscuits and tea
chocolate-cake and cream	chocolate, cake and cream
fruit-salad and milk	fruit, salad and milk
cheese-sandwiches and cake	cheese, sandwiches and cake
cream-buns and jam	cream, buns and jam
fish-fingers and buns	fish, fingers and buns
coffee-cake and tea	coffee, cake and tea
milk-bottles and bread	milk, bottles and bread

References

BEACH, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: evidence of cue trading relations. *Journal of Memory and Language* **30**, 644–663.

BEACH, C. M., KATZ, W. F. & SKOWRONSKI, A. (1996). Children's processing of prosodic cues for phrasal interpretation. *Journal of Acoustical Society of America* **99.2**, 1148–1160.

BECKMAN, M. E. & PIERREHUMBERT, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook* **3**, 255–310.

BISHOP, D. (1989). *Test for Reception of Grammar (TROG)* (2nd edn.). Department of Psychology, University of Manchester.

BOOMER, D. S. & DITTMAN, A. T. (1962). Hesitation pauses and juncture pauses in speech. *Language and Speech* **5**, 215–220.

BRUCE, G., GRANSTRÖM, B. & HOUSE, D. (1991). Strategies for prosodic phrasing in Swedish. In *Proceedings of the 12th International Congress of Phonetic Sciences* (vol. 4), 182–185. Aix-en-Provence.

BUTCHER, A. (1981). Aspects of the speech pause: phonetic correlates and communicative function, *Arbeitsberichte* **15**. Institut für Phonetik, Universität Kiel.

CHAFE, W. (1987). Cognitive constraints on information flow. In Tomlin, R. S. (ed.), *Coherence and Grounding in Discourse*, 21–51. Amsterdam: Benjamin.

CORRIN, J., TARPLEE, C. & WELLS, B. (2001). Interactional linguistics and language development: a conversation analytic perspective on emergent syntax. In Selting, M. & Couper-Kuhlen, E. (eds.), *Studies in interactional linguistics*, 199–225. Amsterdam: John Benjamins.

CRUTTENDEN, A. (1997). *Intonation* (2nd edn.). Cambridge: Cambridge University Press.

DANKOVIČOVÁ, J. (1992). The minimum pause duration in spontaneous speech. *PROPH – Progress Reports from Oxford Phonetics* **5**, 17–24.

DANKOVIČOVÁ, J. (1999). Articulation rate variation within the intonation phrase in Czech and English. In *Proceedings of the 14th International Congress of Phonetic Sciences*, 269–272. San Francisco.

DANKOVIČOVÁ, J. (2001). *The Linguistic Basis of Articulation Rate Variation in Czech* (Forum Phoneticum **71**). Hector: Frankfurt am Main. (Published Ph.D. dissertation.)

DE PIJPER, J. R. & SANDEMAN, A. (1994). On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *Journal of the Acoustical Society of America* **96.4**, 2037–2047.

FANT, G., KRUCKENBERG, A. & NORD, L. (1991). Prosodic and segmental speaker variations. *Speech Communication* **10**, 521–531.

GRABE, E. (1998). Intonational phonology: English and German. Ph.D. dissertation, Max-Planck-Institut for Psycholinguistics and University of Nijmegen.

HEIKE, A. E., KOWAL, S. & O'CONNELL, D. C. (1983). The trouble with "articulatory" pauses. *Language and Speech* **26**, 203–214.

HORNE, M., STRANGERT, E. & HELDNER, M. (1995). Prosodic boundary strength in Swedish: final lengthening and silent interval duration. In *Proceedings of the 13th International Congress of Phonetic Sciences* (vol. 1), 170–173. Stockholm.

HOWELL, D. C. (2001). *Statistical Methods for Psychology* (5th edn.). Pacific Grove, CA: Brooks/Cole-Thomson Learning.

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** (5), 3179–3191.

KLATT, D. H. & COOPER, W. E. (1975). Perception of segments duration in sentence contexts. In Cohen, A. & Nooteboom, S. (eds.), *Structure and Process in Speech Perception*, 69–89. Heidelberg: Springer Verlag.

KNOWLES, G. (1991). Prosodic labelling: the problem of tone group boundaries. In Johansson, S. & Stenström, A.-B. (eds.), *English Computer Corpora: Selected Papers and Research Guide*. Berlin: Mouton de Gruyter.

KOWAL, S., WIESE, R. & O'CONNELL, D. C. (1983). The use of time in storytelling. *Language and Speech* **26**, 377–392.

LADD, D. R. (1986). Intonational phrasing: the case for recursive prosodic structure. *Phonology Yearbook* **3**, 311–340.

LADD, D. R. & CAMPBELL, W. N. (1991). Theories of prosodic structure: evidence from syllable duration. In *Proceedings of the 12th International Congress of Phonetic Sciences* (vol. 2), 290–293. Aix-en-Provence.

LEHISTE, I. (1977). Isochrony reconsidered. *Journal of Phonetics* **5**, 253–263.

PEPPÉ, S., MAXIM, J. & WELLS, B. (2000). Prosodic variation in southern British English. *Language and Speech* **43**, 309–334.

ROSEN, S. M. & FOURCIN, A. J. (1986). Frequency selectivity and perception of speech. In Moore, B. C. J. (ed.), *Frequency selectivity in hearing*, 373–487. London: Academic Press.

SANDERMAN, A. A. & COLLIER, R. (1997). Prosodic phrasing and comprehension. *Language and Speech* **40**, 391–409.

SCHUETZE-COBURN, S., SHAPLEY, M. & WEBER, E. G. (1991). Units of intonation in discourse: a comparison of acoustic and auditory analyses. *Language and Speech* **34**, 207–234.

SCOTT, D. R. (1982). Duration as a cue to the perception of a phrase boundary. *Journal of the Acoustical Society of America* **71.4**, 996–1007.

SCOTT, D. R. & CUTLER, A. (1984). Segmental phonology and the perception of syntactic structure. *Journal of Verbal Learning and Verbal Behaviour* **23**, 450–466.

SEMEL, E., WIIG, E. & SECORD, W. (1987). *Clinical Evaluation of Language Fundamentals – Revised*. London: The Psychological Corporation.

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

STACKHOUSE, J. & WELLS, B. (1997). *Children's Speech and Literacy Difficulties: A Psycholinguistic Framework*. London: Whurr Publishers.

STENSTRÖM, A.-B. (1990). Pauses in monologue and dialogue. In Svartvik, J. (ed.), *The London-Lund Corpus of Spoken English: Description and Research*. Lund: Lund Univesity Press.

STREETER, L. A. (1978). Acoustic determinants of phrase boundary perception. *Journal of the Acoustical Society of America* **64.6**, 1582–1592.

TRIM, J. L. M. (1959). Major and minor tone-groups in English. *Le Maître Phonétique* **112**, 26–29.

TURK, A. E. (1999). Structural influences on boundary related lengthening in English. In *Proceedings of the 14th International Congress of Phonetic Sciences*, 237–240. San Francisco.

VIHMAN, M. (1996). *Phonological Development: The Origins of Language in the Child*. Oxford: Blackwell.

WELLS, B. & MACFARLANE, S. (1998). Prosody as an interactional resource: turn-projection and overlap. *Language and Speech* **41** (special issue on Prosody and Conversation), 265–298.

WELLS, B., PEPPÉ, S. & GOULANDRIS, A. (submitted). Intonation development in children aged five to thirteen. Ms., University of Sheffield.