
Article:
Wells, B., Peppé, S. and Goulandris, N. (2004) Intonation development from five to thirteen. Journal of Child Language, 31 (4). pp. 749-778. ISSN 0305-0009

https://doi.org/10.1017/S030500090400652X

J. Child Lang. **31** (2004), 749–778. © 2004 Cambridge University Press
DOI: 10.1017/S030500090400652X Printed in the United Kingdom

Intonation development from five to thirteen*

BILL WELLS
University of Sheffield

SUE PEPPÉ
Queen Margaret University College, Edinburgh

AND

NATA GOULANDRIS
University College, London

(*Received* 30 *July* 2001. *Revised* 3 *March* 2004)

ABSTRACT

Research undertaken to date suggests that important developments in the understanding and use of intonation may take place after the age of 5;0. The present study aims to provide a more comprehensive account of these developments. A specially designed battery of prosodic tasks was administered to four groups of thirty children, from London (U.K.), with mean ages of 5;6, 8;7, 10;10 and 13;9. The tasks tap comprehension and production of functional aspects of intonation, in four communicative areas: CHUNKING (i.e. prosodic phrasing), AFFECT, INTERACTION and FOCUS. Results indicate that there is considerable variability among children within each age band on most tasks. The ability to produce intonation functionally is largely established in five-year-olds, though some specific functional contrasts are not mastered until C.A. 8;7. Aspects of intonation comprehension continue to develop up to C.A. 10;10, correlating with measures of expressive and receptive language development.

[*] This research was supported by award R000236696 from the Economic and Social Research Council (U.K.), to Bill Wells and Sue Peppé, both then at the Department of Human Communication Science, University College London. Preparation of this article was supported by an ESRC Research Fellowship to the first author (R000271063). We are grateful to the children who participated in this research, to the teachers and parents who facilitated their participation; to Jana Dankovičová for advice on prosodic hierarchies; and especially to Harriet Lang for assistance with recording and data collection. Address for correspondence: Professor Bill Wells, Department of Human Communication Sciences, University of Sheffield, 31 Claremont Crescent, Sheffield S10 2TA, UK. tel: 44 (0)114 2222429; fax: 44 (0)114 2730547; e-mail: bill.wells@sheffield.ac.uk

INTRODUCTION

While most research into child language development has focused on the preschool years, there is evidence that a wide variety of spoken language behaviours continue to change through the school years. At the phonetic level, there are changes in segmental and prosodic aspects of speech production (Ferrand & Bloom, 1996; Whiteside & Hodgson, 2000). In the area of discourse, school-age children only gradually develop the ability to produce accurate messages and to evaluate the adequacy of messages they hear (Lloyd, Mann & Peers, 1998). In the sociolinguistic domain, the ability to understand speakers of an unfamiliar accent has been shown to be superior in seven-year-olds, compared to four-year-olds (Nathan, Wells & Donlan, 1998). Such examples indicate that the speech and language development of children in the school years is of both theoretical and practical importance. In the United Kingdom, for example, teachers in primary schools are now provided with materials to develop the speaking and listening skills of their pupils, together with objectives to be met each term. In this situation, it is important for materials developers, educators and policy makers to have an understanding of what spoken language capabilities should be expected of children at different stages in their school career.

Research suggests that one area where skills continue to develop through the school years is intonation. Relevant evidence comes from a study of the spontaneous speech of children from Tyneside (N.E England) reported in Local (1980). During the course of the children's sixth year, there was a highly significant increase in number of words per tone unit over the period studied. There also were significant changes in the relative frequency of nuclear tones, with a decrease in the number of falls and an increase in the number of levels (boys and girls) and rises (girls only). Local concludes that these changes demonstrate how the children's intonation system is becoming more complex, and, in the case of the nuclear tones, closer to the particular adult intonation variety to which the children are exposed (the adult Tyneside system is characterized by rises and levels accompanying declaratives). This developmental shift is compatible with the view that children begin with simple intonation systems that become progressively more complex as new tones etc. are added (Crystal, 1986). Furthermore, the difference between boys and girls in relative frequency of rises indicates that there may be intonational differences within a group of children of the same age, differences that may be related to factors other than age – in this case, gender.

While these findings indicate that important developments in children's intonation output continue into the 'school-age' period (most UK children start school during their fifth year), our present knowledge of intonation developments in the school years is patchy, in terms of what specific developments take place, at what age, and the degree of variability across children of the same age. Findings from the literature with regard to these

issues tend to conflict, and there appear to be two main reasons for this: differences in aims and differences in methodology. We now discuss the aims, methods and results of the most relevant studies; before presenting our own study, which has as its goal to contribute to a more comprehensive picture of the development of functional prosodic ability in the school years.

Theoretical framework

Any investigation of intonation development is necessarily predicated on a phonological description of the adult intonation system that serves to relate meanings to their prosodic exponents. There is little consensus as to how this should be achieved, due in part to disagreements as to how to identify the meaning functions that intonation realizes, and in part to differences of view about the organization of prosodic structure (Ladd, 1996). For the purposes of this study, meanings of intonation are grouped into four communicative areas, described in detail below: chunking, affect, interaction and focus. Within each of these areas, a specific contrast in meaning is investigated in detail in the study. This meaning contrast is related to its phonetic exponents by means of phonological structures and systems that enter into a prosodic hierarchy (House, Dankovičová & Huckvale, 1999). The largest constituent is the Intonational Phrase (IP), the domain for a complete, well-formed intonation contour. This has initial and final boundaries (%), which can each be high (H) or low (L). Each IP comprises one or more Accent Groups (AGs), defined as the domain for a pitch accent configuration. The pitch accent notation represents pitch accent configurations as a sequence of high (H) or low (L) levels, the central one of which is marked * (Ladd, 1996, p. 79). AGs contain one or more Feet, each of which consists of a strong initial syllable and any following weak syllables. In the following section, in the context of a review of studies of the development of intonational functions in children, it will be shown how intonational oppositions at various levels of the hierarchy can express meaning contrasts in each of the four communicative areas.

Communicative areas of intonation: structure and development

Chunking. 'Chunking' refers to prosodic delimitation of the utterance into units. Prosodic boundary features are associated with interactional as well as grammatical units (Ford & Thompson, 1996) but it is grammatical chunking which has been the focus of attention in developmental studies, not least because these grammatical distinctions are more amenable to formal testing. For the present study, it was decided to investigate chunking in the context of a grammatical distinction between compound nouns, such as

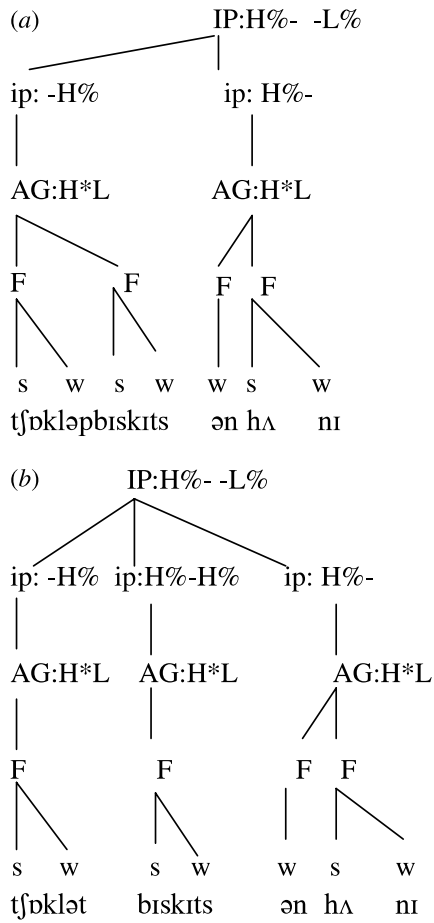

Fig. 1. Prosodic structure of *chocolate biscuits and honey*: (a) compound noun plus simple noun, (b) three simple nouns.

chocolate biscuits (Figure 1a) and strings of two nouns, such as *chocolate, biscuits* .... (Figure 1b).

Here, the difference is in the number of small intonation phrases and consequently, accent groups. In the second utterance, there are three small intonational phrases, each with its own accent, whereas in the first utterance there is no separate small intonational phrase and accent for *biscuits* (see Dankovičová, Piggott, Wells & Peppé, 2004, for further discussion of small intonational phrases and their development).

Research suggests that some aspects of prosodic phrasing continue to develop in the school years. Cruttenden (1985) investigated the following

contrast in prosodic phrasing, which accompanies a contrast in grammatical constituent structure:

(1a) she *dressed% and fed the *baby%
(1b) she dressed and fed the *baby%

As part of a wider study of intonation comprehension, Cruttenden compared a group of twenty ten-year-olds to twenty adults on the ability to interpret 'dressed' as transitive or intransitive according to its prosodic phrasing. Three interpretation options (provided as pictures) were offered: the minimal pair of meanings distinguished by the presence or absence of prosodic boundary within the utterance, and a distracter. Participants heard each sentence three times, and on that basis made a judgement as to the most appropriate picture. The results showed a significant difference between adults and children for sentence (1b). This suggests that the ability of ten-year-olds to map prosodic phrasing onto these different coordinated structures is not yet adultlike.

However, it appears that in other circumstances, younger children are able to use prosodic phrasing to interpret grammatical structure. Beach, Katz & Skowronski (1996) investigated children's processing of coordinated adjectival phrases that are differentiated by prosodic phrasing: [(*pink and green*) *and white*] vs. [*pink and* (*green and white*)]. Subjects were adults, and groups of seven- and five-year-old children (twenty per group). In an identification task, subjects had to choose the correct picture according to the stimulus. For [(*pink and green*) *and white*], pink and green rabbits were depicted close together, with a white rabbit on its own; for [*pink and* (*green and white*)], the pink rabbit was on its own and the green and white rabbits grouped together. Results showed that both groups of children behaved like adults in drawing on pitch and duration features to guide their interpretation. This suggests that even five-year-olds can use prosodic boundaries to guide grammatical interpretation – a finding strikingly at variance with Cruttenden's. If, as might reasonably be expected, the phonetic exponency of prosodic phrasing in the two experiments is similar, the different results suggest that children's success in interpreting prosodic boundaries as cues to grammatical structure may be in part a function of the complexity of the grammatical structure in question.

In a parallel study, Katz, Beach, Jenouri & Verma (1996) investigated the same three age groups' production of the same contrast: [(*pink and green*) *and white*] vs. [*pink and* (*green and white*)]. Three blocks (one pink, one green, one white) were grouped by the experimenter in different ways, and the subject was asked to tell the experimenter what s/he saw. The adults manipulated the length of *pink* and *green*, the accompanying intonation contour and the pauses following them, to indicate the grouping of the blocks. However, the children did not. Thus, in spite of the apparent ability of children

of this age to interpret adults' use of prosodic boundaries in an adultlike way (Beach *et al.*, 1996), in their own speech the children appeared to use neither pitch nor duration features in an adultlike way to convey grouping of objects. This result suggests that the production of this intonational function may still be developing in the school years.

Affect. As an instantiation of the use of intonation to convey affective or attitudinal meaning, we use the distinction between expressing strong liking as opposed to reservation, on a single syllable *M*. This distinction can be expressed by using rise–fall vs. fall–rise pitch movement respectively (see also the task used by Cruttenden, 1985, described below). In terms of the prosodic hierarchy, this opposition is reflected in the boundary tones of the Intonational Phrase, and in the accent type at the level of Accent Group, as illustrated in Figure 2.

Some studies have reported a development in children's use of intonation to interpret speaker affect. Van der Meulen, Janssen & den Os (1997) had three lexically neutral sentences recorded by an actor attempting to convey each of four different emotions (fear, happiness, anger, sadness). Participants (four-, five- and six-year-old Dutch children, and adults) had to associate a particular sentence heard on tape with one of four pictures depicting the four emotions. Adults were almost 100% accurate on the task, and the older groups of children were significantly better than the younger children at identifying emotion. However, it is not clear exactly what intonational or other phonetic/vocal differences distinguished the four different emotions.

Cruttenden (1985) provides more phonetic detail. One of his tasks consisted of the sentences:

(2a) it's a very nice H*garLdenH%
(2b) it's a very nice H*gardenL%

The three pictures depicted:

i. a nice garden but the house falling down
ii. house and garden both very nice
iii. garden overgrown but house very nice.

Sentence (2a) is regarded as a 'marked' intonational option here, since (2b) has a falling tone, generally regarded as the most common and neutral tone. Sentence (2a) is associated with picture (i), since the fall–rise tone carries a meaning of 'reservation' here, as in a wide number of contexts in many varieties of British English. This meaning of reservation is absent from the tune in (2b) with falling tone, which is therefore associated with picture (ii). Picture (iii) provides a distracter. Participants heard each sentence three times, and on that basis made a judgment as to the most appropriate picture. In the case of Sentence (2a), 14/20 adults correctly chose Picture (i), 3/20 chose the 'wrong' picture (ii), and 3/20 chose the distracter. Of the children,

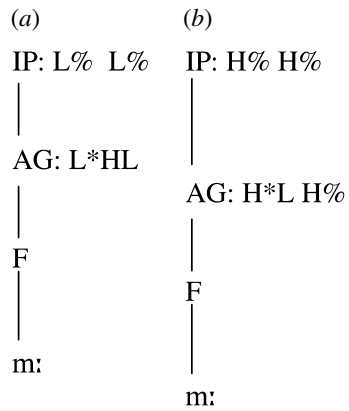

Fig. 2. Prosodic structure of *M* with: (a) rise–fall pitch to express liking, (b) fall–rise pitch to express reservation.

just 4/20 chose Picture (i), while 14/20 chose Picture (ii). By contrast, both groups performed similarly on sentence (2b): 15 adults and 17 children chose Picture (ii), the correct picture. This result suggests that adults were better than children at identifying the affective meaning of 'reservation', and that this had not been acquired by ten years of age – providing more evidence that developments of intonation comprehension are still continuing.

While the group difference is significant, there is variation among the children, four of whom succeeded on the task. Conversely, six of the adults did not display comprehension of this intonational function, indicating variation among the adults too. Such variation may be attributable either to problems with intonation comprehension or to difficulties with the procedure. The issue of variability of response is still more evident in tasks that tap in to the hypothesized relationship between intonation and other types of affect. In the same study, Cruttenden (1985) investigated a contrast between surprise, correlated with a high rising pitch as in (3a), and neutral attitude as in (3b):

(3a) she's gone a L*wayH%
(3b) she's gone a H*wayL%

The picture for (3a) showed a girl with a surprised face, and for (3b) a girl with an impassive face. While children as a group performed worse than adults, many of the adults, like the children, associated both intonational tunes with the surprised face.

Other researchers have attempted to track developmental changes in even more subtle affective meanings that have been hypothesized to have an intonational component, e.g. irony/sarcasm (Capelli, Nakagawa & Madden, 1990; Winner & Leekam, 1991). The studies failed to find differences

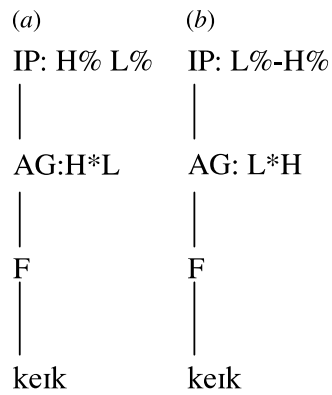

Fig. 3. Prosodic structure of *cake* with: (a) falling pitch to express affirmation, (b) rising pitch to express questioning.

between different age groups of children, though the former study did find some differences between eleven-year-old children and adults. Comparison of these two studies highlights the fact that agreement on the intonational correlates of specific emotions is hard to find. The descriptions of the 'sarcastic intonation' are different for the two experiments: Winner & Leekam's speaker used a 'flat tone' whereas Capelli *et al.* report that their speaker 'greatly exaggerated the modulation of pitch'. The results of these studies point to the difficulties involved in tapping into the supposed attitudinal meanings of intonation, and the importance of clearly specifying the hypothesized intonational correlates of the emotion to be investigated.

Interaction. In order to assess the role of prosody in interaction, we have chosen the contrast between 'yes I understand'; as opposed to 'no I didn't understand, please repeat', which can be realized through intonation on a single word, for example by a low fall compared to a high rise. This distinction between affirmation in order to confirm an understanding vs. questioning in order to check an understanding is a rather specific, well-attested instantiation of the broad (but often inaccurate) generalization that statements are realized with falling pitch, and questions with rising pitch. As in the Affect example, the opposition is at the levels of Intonational Phrase and Accent Group, as illustrated in Figure 3.

Studies of everyday talk have revealed many ways in which adult speakers deploy intonational features for interactional ends (*cf.* Couper-Kuhlen & Selting, 1996). For example, intonation can be used to mark the continuation vs. the end of a conversational turn – a skill that seems to be evident in young children in the second year of life, as they first begin to use multiword utterances, and which may serve to create the interactional space that allows them to develop more complex grammatical structures

(Branigan, 1979; Corrin, Tarplee & Wells, 2001). Researchers of early intonation development have been attracted particularly to the hypothesis that rising or high terminal pitch may be associated with communicative functions that demand a response from the child's conversational partner (often a parent), while falling or low pitch may be associated with other functions. Observations of individual children have lent some support to this idea (e.g. Halliday, 1975) but the association between intonational form and communicative function is by no means invariant, and a good deal of individual variation has been found (e.g. Flax, Lahey, Harris & Boothroyd, 1991). The upshot is that there is no clear developmental picture of this interactional use of intonation. There is a dearth of research into interactional functions of prosody in older children, perhaps because of the methodological difficulties of tapping interactional function in an experimental design. For instance, Cruttenden (1985) does not investigate this type of meaning in his study of intonation comprehension in ten-year-olds. Nevertheless, given their importance in adult talk, the development of such uses of intonation for interactional purposes merits further research.

Focus. This refers to the speaker's use of phonetic prominence to indicate which item is most important in an utterance. In English a pitch accent is located on the final stressed syllable of the constituent to be focused. This can lead to some ambiguity of interpretation for a string such as *I wanted chocolate and *honey*, where the domain of focus could be *honey*, i.e. narrow focus, or the whole phrase *chocolate and honey* (broad focus). Narrow focus can be located on non-final as well as final constituents, e.g. *I wanted *chocolate and honey*.

In the present study, as in most of those described below, the investigation concentrates on narrow focus, in the context of doing correction. The accentual structures for Non-final Narrow Focus and Final Narrow Focus respectively, can be represented as in Figure 4.

While the two structures are identical from the level of the Foot and below, there are differences at the levels of Intonation Phrase (IP) and Accent Group (AG). While each utterance consists of a single IP with a low final boundary tone, they differ in the initial boundary tone: the Non-final Focus begins with a high boundary tone, whereas the Final Focus begins with a low boundary tone. Each utterance also contains only one AG, realized by a H*L pitch accent. However, the utterances differ in the alignment of the AG. In the Non-final Focus utterance, the AG has as its head the leftmost Foot, which begins with *chocolate*, while *honey* forms an enclitic Foot, attached to the preceding Foot. In the Final Focus utterance, the AG has as its head the Foot consisting of *honey*, while *chocolate and* forms a proclitic Foot, attached to the Foot that follows it.

Developmentally, the ability to manipulate the location of the main accent has been attributed to young children as soon as they begin to produce

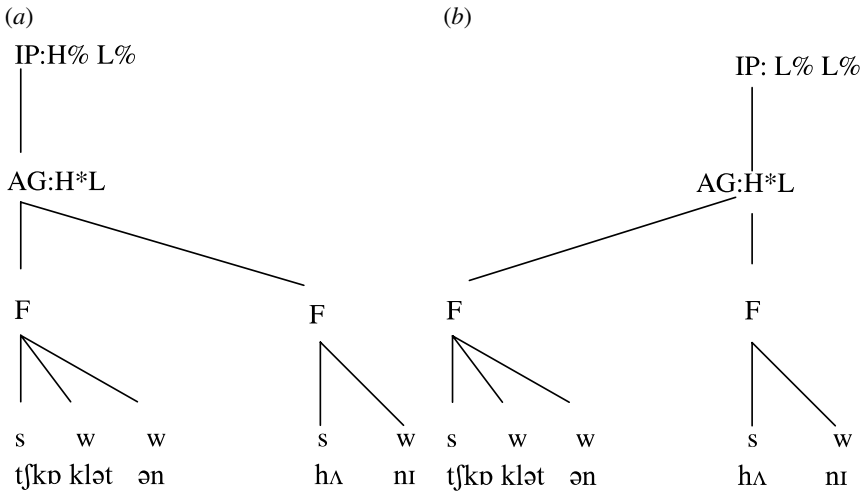

Fig. 4. Prosodic structure of *chocolate and honey* with: (a) non-final narrow focus, (b) final narrow focus.

utterances of more than a single word (Crystal 1987, p. 73), and in an influential study Wieman (1976) claimed that the location of the accent was not random, but influenced by information structure: 'Children operate with an appreciation of what is new in their utterance, and apply stress accordingly'. Thus, whereas in general an utterance like *Blue man* would have the stress on the noun rather than the attribute, this would not be the case when the noun had already been mentioned, such as *Man. Blue man.* In the latter case, the child would use non-final narrow focus, on *blue.* Wells & Local (1993) present some evidence that casts doubt on the generality of that claim, suggesting that some children may start off with a pattern of accenting the final stressed syllable of the utterance, irrespective of focus considerations, and only later learn to manipulate accent placement for focus purposes. This is further evidence, albeit from very young children, of variability across children in intonation development.

Experimental studies of production have demonstrated that children aged 3;0 to 5;11 can use accent placement to achieve narrow focus, in order to do corrections (Hornby & Hass, 1970; Hornby, 1971; Macwhinney & Bates, 1978), suggesting that this ability is well developed in the preschool period. On the other hand, Cruttenden (1985) found that his ten-year-old subjects, while performing above chance level, were significantly worse than adults at assigning to the correct picture sentences which differed in focus/accent structure: *John's got *four oranges* vs. *John's got four *oranges.* This result echoes the finding of Cutler & Swinney (1987) that younger children (aged

4;0–6;1, $n=10$), while sensitive to stress *per se*, differed from children aged 6;5–7;11 ($n=11$) and adults in being unable to use the location of sentence accent to aid sentence processing. In conclusion, while there is evidence that young children are able to manipulate accent for the purposes of narrow focus in their own speech, studies of comprehension suggest that some aspects of the interaction between accent and focus may not be mastered by ten-year-olds.

The research reviewed above suggests that within each of the four communicative areas, certain aspects of children's intonation comprehension and production continue to develop during the school years. At the same time, they suggest that there may be considerable variation in intonation performance among children of the same age. Thus a major challenge for research into intonation development is to determine to what extent variation in intonation performance can be related to age, and thus to developmental factors; and to what extent the variation is attributable to other factors. It has been noted in passing that variation is found among adults too. This was explored systematically by Peppé, Maxim & Wells (2000), who noted that variation may in theory be due to a range of factors, including dialect/accent, phonological or discourse context, gender, age, or education level. Controlling for dialect/accent and context, they found that some variation was attributable to education level when combined with age, but not to gender. However, a large amount of variation remained which could not be attributed to any of the factors investigated, suggesting that even among adults there are considerable individual differences.

Each of the child intonation studies reviewed has addressed a relatively circumscribed area of intonation development, with a restricted age-range of children, and widely different methods have been used. Consequently, there are still important gaps in our knowledge of children's later intonation development. Furthermore, these studies do not address the issue of whether intonation development is related to other aspects of linguistic development. In order to identify the variation in children's intonation performance that is attributable to age, it is necessary to control as far as possible for other factors that may lead to differences among children. The present study investigates groups of children aged between 5;0 and 14;3, matched for gender, who have been drawn from a homogeneous population in terms of accent and educational background. Materials are used that control for contextual effects as far as possible. In order to explore the relationship between intonation and grammatical development, tests of grammatical comprehension and production are also administered. In this way, the study aims to contribute to a more comprehensive picture of intonation ability in the school years.

In order to carry out the present study, a set of tasks was drawn from a prosodic test battery, PEPS-C (Profiling Elements of Prosodic

Systems – Child version), recently developed as an assessment tool that could be used by professionals working with children with communication difficulties (Wells & Peppé, 2001, 2003; Peppé & McCann, 2003). The battery was based on an earlier procedure for testing adults: Profiling Elements of Prosodic Systems – PEPS (Peppé, 1998; Peppé *et al.*, 2000). Each of the four communicative areas described above (Chunking, Affect, Interaction and Focus) is tested for both comprehension and production. The broad question addressed by this study is: how does functional prosodic performance change after the age of 5;0?

METHOD

Participants

120 children were recruited from state-maintained schools in North London. Participants met the following criteria: English was their first language and the language spoken in the home; they had no identified speech and language problems; they had no general educational problems; they had been resident in the south east of England for at least three years. Participants were selected by age to form groups (30 per group: 15 male, 15 female), separated by approximately three years; because of administrative constraints, precise intervals were difficult to achieve, and the average ages of the groups, in years, were as follows: 5;6 (range 5;0–5;11, *S.D.* 0·226); 8;7 (range 7;11–8;11, *S.D.* 0·23); 10;10 (10;5–11;2, *S.D.* 0·193); 13;9 (range 13;5–14;4, *S.D.* 0·254). A supplementary group of 73 children (approximately 18 in each of the above age groups) was selected to do the two Focus tasks, which were developed and refined after data collection for tasks in the other three areas had been completed.

Materials

Each of the four communicative areas in PEPS-C is tested for both Input and Output, giving a total of eight tasks. Each Input task has sixteen items, and each Output task twelve items. Care was taken in devising the tasks to ensure that in their responses the child is obliged to draw on prosodic resources rather than other linguistic or non-linguistic means. This was achieved principally by controlling the lexical and grammatical content of test items. In addition, various measures were taken to ensure that the tester decided the intended meaning on the basis of only the prosodic features of the child's response: a screen hid the facial expressions of both tester and child, and confirmation of his/her intended meaning supplied only after the tester had made a judgement. The tasks are now described, with a summary in Table 1.

TABLE 1. *Brief description of prosodic tasks*

Task name	Description
Chunking Input	Identification: recorded voice names 2 foods (e.g. FRUIT-*SAL*AD AND *MILK*) or 3 foods (e.g. *FRUIT*, *SAL*AD AND *MILK*).
Chunking Output	Naming: picture-strip shows either 2 foods (e.g. FRUIT-SALAD, MILK) or 3 foods (e.g. FRUIT, SALAD, MILK).
Affect Input	Identification. Single food item on picture. Recorded voice likes it ([m] with rise–fall) or is not keen ([m] with fall–rise).
Affect Output	Child hears food-item (e.g. BANANAS) and, with [m] only, expresses liking or not keen.
Interaction Input	Identification. Child names picture (e.g. CUP) which tester repeats either fall with low onset (affirming, i.e. 'go on') or rise with high onset (questioning, i.e. 'repeat'). Child decides whether the tester wants child to go on to the next item or to repeat.
Interaction Output	Recorded voice speaks a non-word (e.g. PARGLE) or a real word (e.g. CARROT). Child repeats word, to sound as if questioning in order to check understanding (non-word) or affirming, to confirm understanding (real word).
Focus Input	Identification. Recorded stimuli, e.g. 'I wanted *CHOC*OLATE and honey'/'I wanted CHOCOLATE AND *HONE*Y'. Child decides which food the speaker had not received.
Focus Output	Tester offers child a picture saying e.g. 'How about a green bike?' Child has to respond so as to get the picture s/he actually needs e.g. 'I WANT A *WHITE* BIKE'.

Input. The Input tasks are designed as identification tasks: the child hears a spoken stimulus, and has to assign it to one of two meaning categories. In the Chunking Input task, the child (having first been shown pictures of the food-items involved, to check vocabulary) hears a single utterance such as the following: *fruit-*salad% and *milk%*; or alternatively **fruit%, *salad% and *milk%*. The difference is in the number of small intonation phrases and consequently, accent groups. In the second utterance, there are three small intonational phrases, each with its own accent, whereas in the first utterance there is no separate small intonational phrase and accent for *salad* (*cf.* Figure 1). The child is asked to say whether the utterance sounds like two items of food or three.

In the Affect Input task, the child has two pictures: a smiley face, representing 'liking' and a doubtful face, to represent 'reservation'. The child has to indicate the picture corresponding to the utterance that is played on the tape. Tones are carried on the single syllable M. Liking is conveyed by a rise–fall pitch, reservation by a fall–rise. This opposition is reflected in the boundary tones of the Intonational Phrase, and in the accent type at the level of Accent Group, as illustrated earlier in Figure 2.

In the Interaction Input task, the intonational distinction is between a low fall, with affirmative meaning, 'yes I understood'; or a high rise, with questioning meaning, 'no I didn't understand, please repeat'. The opposition is again at the levels of Intonational Phrase and Accent Group (*cf.* Figure 3). The child is given a set of pictures, bound together, each depicting a single object, e.g. *cup*, *key*, etc. The child says what is in each picture, and the tester repeats the word, with either high rising or low falling intonation. If the intonation is rising, the child is expected to repeat the word; if falling, to go on to the next picture.

In the Focus Input task, the child hears a single utterance, e.g. **chocolate and honey %* or alternatively *chocolate and *honey %*. The child has to identify which of two items in the utterance is highlighted by the speaker, and indicate this by pointing to the appropriate picture. Intonational prominence (pitch step-up to the start of the main syllable of the focal item, then falling pitch-movement) serves to focus on one item of food. The phonological opposition is at the levels of Intonation Phrase (difference in initial boundary tone) and Accent Group (location) (*cf.* Figure 4).

Output. In the Chunking task, the child has a pile of picture-strips, each of which depicts either two items of food (e.g. *chocolate-biscuits*, *honey*) or three items (e.g. *chocolate*, *biscuits*, *honey*). The child picks up one picture-strip, unseen by the tester, and tells the tester what is on it. The tester notes down whether the child sounded as though s/he was talking about two items of food or three, and then checks by looking at the picture strip. When scoring, the tester compares what the response sounded like with the contents of the picture-strip itself; thus the child is assessed on whether or not s/he can realize his/her own communicative intention by signalling the correct number of small intonation phrases/accent groups, with their boundaries aligned appropriately to the lexicogrammatical structure.

In the Affect Output task, the child has two cards: a smiley face and a doubtful face. The tester explains that she wants to know what food the child likes and what he is not too keen on. The tester names an item of food, e.g. *bananas*. If the child likes it, he should say [m:] with an appropriately enthusiastic intonation. This intonation could be a rise–fall, as in the Input stimuli (*cf.* Figure 2a), but other intonation patterns may also be scored as correct if deemed by the tester to convey 'liking', e.g. fall starting high with wide pitch range. If the child is not too keen, he should pronounce [m:] with an appropriately unenthusiastic intonation such as the fall–rise used in the Input tasks (*cf.* Figure 2b), or a small F0 variation low in the pitch range. While the child is responding his face is hidden from the tester by a screen, so that only his phonetic production can signal his inclinations. The phonological opposition can involve intonation phrase and accent group. The tester has access to the child's intention because after uttering each response the child has to point to either a smiley face or a doubtful face. In this way

the child's ability to realize his communicative intention phonetically can be assessed.

In the Interaction Output task, the child has one card with a tick on it and another with a question-mark. S/he hears a list of words and is required to repeat the word with an appropriate intonation. The word may be familiar, e.g. *carrot*, in which case the child repeats the word in such a way as to affirm that it has been understood, e.g. with a falling pitch contour. Alternatively, the word may be unfamiliar, e.g. *pargle*, in which case the child should query it – for example by using a rising intonation. The phonological opposition can involve intonation phrase and accent group (*cf.* Figure 3). As in the Affect task, the child's facial expression is not seen by the tester. The child indicates his/her communicative intention after responding, this time by pointing to the tick or the question-mark.

The Focus Output task taps into the child's ability to use accent placement in order to focus on a specific item in the utterance, for the purposes of repair. It takes the form of a lotto game, in which the child is offered a picture that does not match the ones he has; the child asks for a different picture, emphasizing the thing that differentiates the picture the child wants from the one that had been offered. Exchanges such as the following occur: Tester: 'How about a green bike?' Child: 'I want a *white bike%'; or Tester: 'How about a black boat?' Child: 'I want a black *bus%.' The child's response is scored as correct if he conveys narrow focus on the item of new information, by aligning the Accent Group with it (*cf.* Figure 4). The tester presents items with broad focus, using an intonation contour that does not highlight either the colour or the vehicle. Typically this is a downdrift contour with a low fall on the final word.

Procedure

The procedure consisted of individual tape-recorded interviews, each session lasting generally no longer than 30 minutes, with a maximum of three sessions per participant. Stimuli for the input tasks had been pre-recorded on digital audio tape (DAT) in a recording studio. Stimuli were presented to participants via tape recorder in free field, and responses were recorded on DAT. The first session was preceded by a vocabulary-checking phase, in which it was ascertained the child was familiar with the words illustrated in the test material. In addition to the prosodic battery, each participant was also tested on independent measures of language ability. These are standardized tests, which were administered in order to ascertain that each child's language development was within normal limits, and to find out how prosodic skills correlated with other language skills. Language production was measured on an expressive language subtest of the Clinical Evaluation of Language Fundamentals – Revised (CELF-R) (Semel, Wiig & Secord,

1987). In this 'formulated sentences' subtest, the child has to make up a sentence using a given word, and is scored on the lexical appropriateness and grammatical coherence of the sentences produced. Comprehension was measured on the Test for the Reception of Grammar (TROG) (Bishop, 1989). In this test, children hear a sentence and have to match it to one of four pictures; the other three pictures show scenes and objects that might lead children to select them if they had misunderstood the grammar of the sentence.

RESULTS

On the Input tasks, each of which comprises 16 items, each child has only two choices for each item – the response is either right or wrong. According to the binomial distribution scores equal to or above 12 indicate that responses are significantly above chance. On the Output tasks, each of which comprises 12 items, the scorer marks the child's production of each item as right (2 points), wrong (0 points) or ambiguous (1 point), giving a possible maximum of 24. In order to interpret the results, it is useful to have a pass mark, above which we can be reasonably confident that the child is in command of the relevant aspect of intonation. For Output tasks, this pass mark was set at 18 (75%), since to obtain a score of 18, the child would have to make an unambiguously correct response for at least six items (50%) and make no outright errors.

Children's performance is measured using terms such as 'error' and 'ambiguous response'. These are useful categories for assessing responses in relation to the model of intonation described in the Introduction, and for providing a quantitative indicator of age-related differences in performance. However, it cannot be assumed that the intonation patterns that are counted here as 'error' responses do not occur in the adult population. While the model of intonation presented in the Introduction reflects a consensus view as to likely realizations of these intonational functions by the majority of adult speakers of Southern British English, in reality there is considerable variation in the adult population in this respect (Peppé *et al.*, 2000). This being the case, it is likely that some of the variation in children's performance reported below is not due to developmental factors, but rather reflects variation in the population at large. In the present study, the aim is not to compare children's performance against an adult 'ideal' performance, or indeed adult performance derived from an empirical study. Rather, the aim is to identify differences in performance across groups of children of different ages.

Descriptive statistics for each communicative area are presented in Tables 2 and 3. Number of participants is followed by mean scores, standard deviations, and the range. All scores are presented as percentages.

TABLE 2. *Results by age group for Input prosodic tasks*

Age group	Chunking				Affect				Interaction				Focus			
	N=	Mean	*S.D.*	Range	*N*=	Mean	*S.D.*	Range	*N*=	Mean	*S.D.*	Range	*N*=	Mean	*S.D.*	Range
5	30	75·6	14·3	43·8–93·8	30	85·8	18·2	37·5–100	29	69·0	22·9	31·3–100	18	49·7	4·0	31·3–93·8
8	30	82·5	13·2	50–100	30	94·6	8·0	62·5–100	30	90·2	17·3	31·3–100	17	67·6	19·3	37·5–100
10	30	87·1	12·7	56·3–100	30	95·6	9·2	56·3–100	30	97·3	7·3	68·8–100	18	71·5	20·0	25–93·8
13	30	93·3	7·9	75–100	30	96·5	5·1	87·5–100	30	96·7	7·3	75–100	20	91·9	13·0	43·8–100

TABLE 3. *Results by age group for Output prosodic tasks*

Age group	Chunking				Affect				Interaction				Focus			
	N=	Mean	*S.D.*	Range	*N*=	Mean	*S.D.*	Range	*N*=	Mean	*S.D.*	Range	*N*=	Mean	*S.D.*	Range
5	29	82·2	14·0	45·8–100	29	71·4	25·2	16·7–100	29	67·2	19·4	25–100	18	85·4	9·3	62·5–100
8	30	78·1	13·4	54·2–100	30	84·7	18·1	50–100	30	85·8	17·9	29·2–100	17	86·0	9·4	70·8–100
10	30	84·2	13·2	54·2–100	30	89·3	11·8	58·3–100	30	91·0	8·9	66·7–100	18	87·5	11·3	62·5–100
13	30	88·6	13·2	54·2–100	30	88·1	15·0	50–100	30	82·9	15·1	50–100	20	89·4	9·0	70·8–100

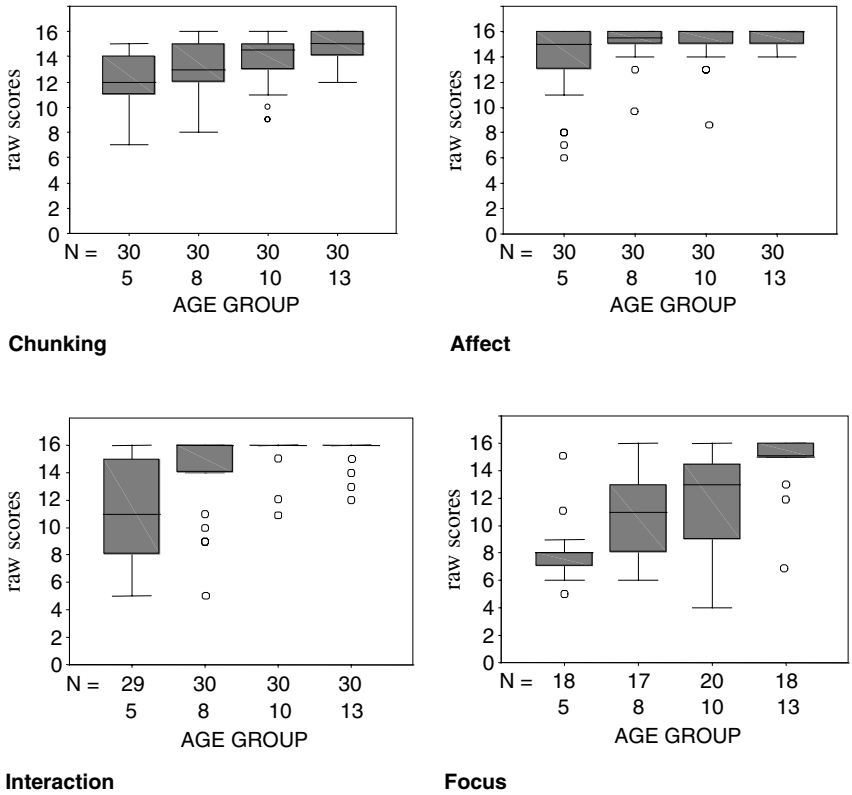

Fig. 5. Box plot for results of Input prosodic tasks.

Ceiling effects were found in all tasks although they were more prevalent in the older groups. This is illustrated in Figures 5 and 6, which respectively present box plots for input and output tasks. The dark line represents the median, the interquartile range (i.e. the middle 50% of the distribution) falls in the box and the whiskers extending from the box to the highest and lowest values depict the remaining 25% at the top and the bottom of the distribution. The outliers are also clearly visible. They demonstrate that even when the majority of the children are successful at a task there are frequently some who are performing poorly. Outliers are present in 7/8 distributions and are almost invariably at the bottom of the distribution.

As ceiling effects and heterogeneity of variance were present, nonparametric tests were carried out on all the tasks. The results are reported below. In order to compare effects of age on the Input and Output tasks, *post hoc* tests were undertaken whenever the omnibus analysis was significant.

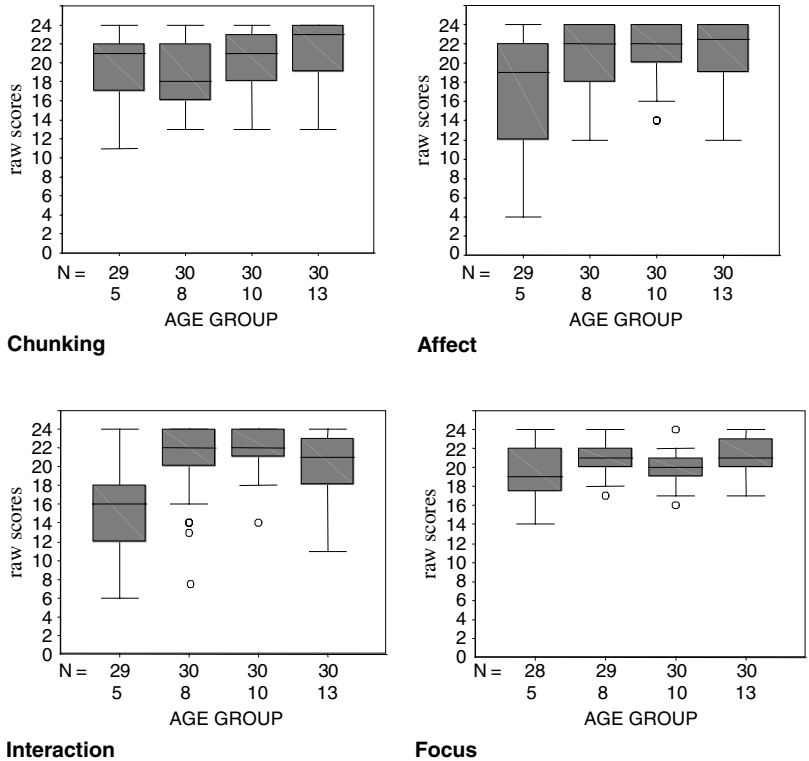

Fig. 6. Box plot for results of Output prosodic tasks.

In order to control for family-wise error a stringent probability level of $p < 0{\cdot}001$ was adopted throughout.

Reliability

As a measure of intra-rater reliability, 10% of the output results were re-scored by the same person six months after scoring had been completed. The difference in scores was 2·3%. As a measure of inter-rater reliability, 10% of the items in the original study were scored by two raters; in most tasks the discrepancy between their scores was at 1·4% or less. In the Focus Output task, however, the discrepancy was 14·9%. It seems likely that the discrepancy was caused by the particularly complex scoring procedure for this task. The 'test–retest' technique for assessing reliability was not applied, being invalid for children whose language skills may have developed in the interval.

TABLE 4. *Statistically significant age related changes on prosodic tasks*

	Input				Output			
Age	Chunking	Affect	Interaction	Focus	Chunking	Affect	Interaction	Focus
5–8			0·000				0·000	
8–10								
10–13				0·001				
5–10	0·001		0·000	0·001			0·000	
8–13	0·000			0·000				
5–13	0·000		0·000	0·000				

Gender

Mann-Whitney tests (Siegel & Castellan, 1988; Howell, 2002) were performed to examine differences in gender. There were no significant effects of gender.

Order of presentation

In order to investigate the possibility that Input tasks, if administered first, might affect performance on Output tasks, or *vice versa*, half the participants were presented with Input tasks first and half with Output tasks first. Mann-Whitney tests showed no significant effects of order of presentation.

Age

Inspection of the data in Table 2 reveals that five-year olds scored below the pass-mark of 75% on the Input Interaction and Focus tasks, but above 75% on Input Chunking and Affect tasks. The means of the three older groups were significantly above 75% on three of the Input tasks, but on the Focus Input task this was true only of the thirteen-year-olds. On the Output tasks, all the age groups attained the pass-mark on Chunking and Focus tasks, but the five-year-olds were below this criterion on Affect and Interaction tasks.

Nonparametric Kruskal Wallis One-Way Analysis of Variance (Siegel & Castellan, 1988; Howell, 2002) was used to determine differences between the four age groups. When significant at the $p < 0{\cdot}001$ level, *post hoc* Mann-Whitney tests were performed to investigate age-related differences in performance. Results are presented in Table 4.

The bottom row of Table 4 shows that there was significant improvement in scores between the youngest and oldest age groups on 3/4 Input tasks. This confirms that there are some age-related changes in intonation processing performance between the ages of 5;0 and 14;3.

TABLE 5. *Statistically significant correlations of prosodic tasks with CELF-R subtest and TROG*

	Input				Output			
	Chunking	Affect	Interaction	Focus	Chunking	Affect	Interaction	Focus
CELF	0·359		0·569	0·588		0·413	0·463	
	0·000		0·000	0·000		0·000	0·000	
TROG	0·433	0·418	0·553	0·524				
	0·000	0·000	0·000	0·000				

Correlations

Spearman correlation coefficients were calculated to explore developmental associations between performance on each subtest and (a) the formulated sentences subtest of CELF-R; and (b) TROG. These are presented in Table 5. Coefficients are given on the top line, with probability values beneath.

There were significant positive correlations of 5/8 subtests with the CELF-R subtest – three of these being Input subtests – and of 4/8 subtests with TROG, all four being Input subtests. This suggests that the improvements in intonation performance, particularly in comprehension of intonation, may be related to developments in expressive and receptive language skills.

Qualitative analysis

A more detailed analysis of responses on the four Output tasks was carried out in order to see whether there were age-related changes in the distribution of error responses and ambiguous responses. The responses of all 120 children were examined for the distribution of error responses, (score 0) and ambiguous responses, which were considered by the scorer as potentially having either meaning (score 1).

In the Chunking Output task the child is presented with a strip of three pictures depicting (for example) chocolate, biscuits and honey (=a '3-list'); or with a strip of two pictures depicting chocolate biscuits and honey (a '2-list'). The child has to describe what s/he sees.

While there was little difference in the error rates of the different age groups on the 3-lists, or in the rates of ambiguous responses for either 3 lists or 2-lists, there were more differences in errors on the 2-lists: the three younger groups performed less well than the thirteen-year-olds (Table 6). The three younger groups were more likely to make 2-lists sound like 3-lists, by segmenting the first noun as if it had been a picture on its own. For example, target *chocolate biscuits and honey* (a 2-list) should have the structure **chocolate biscuits% and *honey%* (*cf.* Figure 1a). The younger

TABLE 6. *Errors and ambiguous responses in Chunking Output task (%), by age group*

	5	8	10	13
Errors in target 3-lists	4·9	11·7	7·1	5·8
Errors in target 2-lists	15·7	13·3	8·5	1·1
Ambiguous in target 3-lists	13·2	15·6	12·5	16·3
Ambiguous in target 2-lists	19·0	19·4	18·8	16·0

TABLE 7. *Errors and ambiguous responses in Affect Output task (%), by age group*

	5	8	10	13
Errors in target 'like'	15·3	3·4	3·7	3·6
Errors in target 'not keen'	18·6	3·9	1·5	5·1
Ambiguous in target 'like'	28·7	28·3	22·0	19·4
Ambiguous in target 'not keen'	20·0	12·6	12·8	15·4

children, when making an error, would produce it in a way that was interpreted by the listener as having the structure: **chocolate% *biscuits% and *honey%* (*cf.* Figure 1b). The children were thus failing to subordinate *biscuits* (anyway a separate foot) to *chocolate* as part of the same small intonational phrase/accent group. In the oldest group, there are fewer errors overall, and the situation has reversed. The children in this group make proportionately more errors by producing 3-lists like 2-lists. Thus, *chocolate, biscuits and honey* (*cf.* Figure 1b) has the structure of *chocolate-biscuits and honey* (*cf.* Figure 1a). It appears that sometimes, for some thirteen-year-olds, the demands of fluency override the requirements of accurate delimitation of intonational phrases.

The distribution of errors and ambiguous responses on the Affect Output task was analysed to see whether there was any developmental change in children's ability to express affective meaning (Table 7).

In the case of both 'like' and 'not keen' responses, five-year-olds made more errors than the other three age groups, who made few errors. Thus five-year-olds appear to have difficulty in expressing both options. In the case of the 'not keen' option: when indicating that they were doubtful about a food item, their intonation did not convey this. Similarly, in the case of the 'like' option, they could not use intonation consistently to convey this feeling. For ambiguous responses, there was a more gradual reduction with age. As was illustrated in the Introduction in Figures 2(a) and 2(b), a common way of realising both options in this variety of English is by using complex tones. One possibility is that the five-year-olds have less control than older children over the deployment of complex tones.

TABLE 8. *Errors and ambiguous responses in Interaction Output task (%), by age group*

	5	8	10	13
Errors in affirming responses	8·5	7·1	1·8	1·8
Errors in questioning responses	41·5	7·8	6·7	10·8
Ambiguous affirming responses	13·4	15·6	7·3	12·7
Ambiguous questioning responses	18·2	9·2	11·8	27·2

TABLE 9. *Errors and ambiguous responses in Focus Output task (%), by age group*

	5	8	10	13
Errors in target non-final focus	6·5	1·3	0·6	2·0
Errors in target final focus	0·8	2·5	3·1	2·4
Ambiguous target non-final focus	10·1	1·9	5·4	0·7
Ambiguous target final focus	37·3	37·6	50·9	36·3

The distribution of errors and ambiguous responses on the Interaction Output task was analysed, to see whether there was any developmental change in children's ability to express interactional meaning (Table 8).

On the affirming responses, there were relatively few errors in all groups, and they declined gradually with age. On the questioning responses, errors were very common (41·5%) among the five-year-old group: there was a strong tendency for the 'questioning' response to sound affirming. This suggests that the younger children have more difficulties in producing a rising tone (as illustrated in Figure 3b) in order to convey a particular communicative need. The thirteen-year-olds produced the highest number of ambiguous questioning responses. This unanticipated result suggests that the thirteen-year-olds are less likely to use intonation to check understanding, than the two younger age groups.

The distribution of errors and ambiguous responses on the Focus Output task was examined (Table 9), with a view to discovering whether there were any developmental patterns in the ability to communicate information focus that were not evident from the quantitative measures presented in Table 3.

The error rate for the Focus task was small for all groups, though the five-year-olds made more errors than the other groups on the Non Final Focus responses, where the colour was to be emphasized. An error in a 'colour' response (target non-final focus) meant that it sounded as though the object was being emphasized, not the colour, e.g. *How about a white car? – I want a green *car*. In terms of the structural descriptions presented earlier, the tendency is for young children to prefer a prosodic structure where the AG (Accent Group) dominates the final foot (Figure 4b) over a structure where the AG dominates a non-final foot (Figure 4a). The

opposite error rarely occurred in any group, e.g. *How about a white car? – I want a *white bike*. The five-year-olds also made more ambiguous responses than the older groups for these target non-final (colour) responses.

However, the most striking finding is the number of ambiguous responses that were consistently obtained for the 'object'-responses in the Focus task. This is not a developmental trend. This result shows that there was a strong tendency for words in the final position of an utterance not to be phonetically emphasized even when there was clear motivation for producing narrow focus. The implications of this finding are explored in the next section.

DISCUSSION

The main findings of the study are now considered.

(a) *The ability to use intonation functionally is largely established by the age of 5;0.*

Most Output skills tested here (Chunking, Affect, Focus) are established for most of the five-year-old group. This tallies with the widely-held view that prosodic resources are already used effectively by very young children to convey communicative intent, e.g. focus/emphasis (Wieman, 1976) and utterance delimitation (Branigan, 1979). Consequently, not all prosodic skills show a clear pattern of development through the age range covered in the present study. On the Chunking Output task, which tests utterance delimitation, the high mean scores (82·2% for the five-year-olds, where the pass-mark is 75%) indicate that by that age, many children have acquired the skill of phrasing an utterance in order to convey the desired meaning. This conflicts with the finding of Katz *et al.* (1996), reported in the Introduction, that seven-year-olds were unable to mark prosodic phrase boundaries in order to disambiguate otherwise identical strings. The reason for the discrepancy may be that in the present study, the children had to signal a lexical distinction (compound noun vs. string of two nouns), whereas in the earlier study, the distinction was syntactic, marking different coordinate structures.

(b) *Some functional prosodic contrasts are hard for younger children to produce, but these are for the most part mastered by eight-year-olds.*

Despite the general finding summarized under (a) above, there is evidence that some five-year-olds still have difficulties with certain aspects of the intonation system. These include:

(i) *The ability to incorporate two words with potential lexical stresses into a single intonational phrase*. This marks the string as a compound, as in 'coffee-cake'. On the Chunking Output task, error analysis showed a tendency among the youngest group to prefer more small intonation phrases/accent groups in the intonation phrase, e.g. one accent group per foot, rather than

to subordinate a foot into an adjacent accent group. It is possible that this error pattern among some of the youngest children reflects immature prosodic competence – an interpretation lent some support by the results of the Input task, presumed to tap competence in this aspect of intonation, where significant improvement was found between five- and ten-year-old groups (Table 4). An alternative possibility is that the tendency of the youngest children to produce more intonational phrases is a by-product of a generally slower speech rate characteristic of younger children, although speech rate was not controlled for in the current study.

(ii) *The ability to use on-syllable prosodic features, to convey meaning.* On the Affect Output task, error analysis (Table 7) indicated that the five-year-olds in particular would confuse the expression of 'liking' vs. 'not keen'. This could reflect immature prosodic competence, either phonetically or at the (functional) phonological level. As the five-year-olds as a group exceeded the pass mark on the corresponding Input task (Table 2), the difficulty is more likely to be phonetic. However, it should be noted that in this Output Function task, the child is not constrained to use a specific contour to realize the required meaning, e.g. a rise–fall for 'like' and fall–rise for 'not keen' (the contours used for the Input task), and it is possible that the poorer performance of the youngest group could be due to non-phonetic factors, e.g. pragmatic aspects of the task. Even at the phonetic level, it is possible for information about 'liking' etc. to be conveyed through articulatory and voice-quality features, as well as intonation e.g. 'smile voice' for positive and groaning for negative affect. The failure of the youngest group to use high rising pitch (in particular) to convey interactional meanings, such as requesting clarification, is another case where on-syllable pitch movement is apparently not used appropriately. Error analysis of Interaction Output task responses indicated that the problems encountered by the younger children were with the expression of questioning, suggesting that for at least some five-year-old children, rising pitch accents may not yet be fully incorporated into the functional intonation system. This parallels the finding of Snow (1998) that children have more trouble with final rising tones than final falling tones. In this case, the results from the corresponding Input task (Table 2) show that the five-year-old group had not yet reached the pass-mark, so their difficulties with the Output task may in part be due to immature phonological (functional) competence.

(iii) *The ability to realize prefinal focus by placing the accent in non-final position.* In the case of the Focus Output task, qualitative analysis revealed a bias among five-year-olds that is not present in older age groups, towards utterance final position for focus accent placement, e.g. 'I want a *green car' realized as 'I want a green *car' (Table 9). This pattern is in line with results from previous experimental and observational studies suggesting that where children make errors with focus accent placement, it is by

shifting the accent to the last word in the utterance, and that this is particularly likely in younger children (Wells & Local, 1993).

(c) *Aspects of intonation continue to develop after age five.*

Within the age span covered by our study, it is apparent that some developments occur in the period between 5;0 and 11;2. Notably, there are developments through the age range in prosodic comprehension (Input), supporting earlier findings that aspects of intonation comprehension are not adultlike by age ten (Cruttenden, 1985). The Chunking Input task shows developmental improvement over a wide age range, and correlates significantly with both language measures (TROG and CELF subtest). Both Interaction tasks show significant improvement between 5;0 and 11;2 (Table 4). This suggests that the skills measured by the Interaction tasks – which involve the contrastive use of simple pitch movements to confirm or check an understanding – are acquired in the early school-age period, and uniformly; while other functions are acquired later and gradually. An example is the child's understanding of the interlocutor's use of accent/focus to highlight the key part of the utterance, as tested in the Focus Input task (Table 2). This lags behind children's ability to use the phonetic features functionally in their own speech (Focus Output, Table 3). This points to the complexities of form/function mapping, and lends some support to the conclusions of Cutler & Swinney (1987), that children may be able to use accent to realize focus in their own speech, before they can make use of accentuation to interpret other speakers' focus.

More generally, it was found that the children's performance on the Input tasks correlated strongly with measures of receptive and expressive language development (Table 5). This suggests that during the school years, intonation and prosodic competence, as measured by the Input tasks, develops in line with other aspects of grammatical comprehension and production, as measured by TROG and the sentence formulation subtest of the CELF.

(d) *There is variation among children in all age groups.*

Although there was a pattern of improvement with age on the majority of tasks, ceiling effects were found even among the five-year-olds, indicating that some children showed early mastery of functional intonation across different communicative areas. On the Chunking Input task (Figure 5), although even the five-year old children attain 75% accuracy, the range of scores for each of the four age groups shows that the task is sensitive to individual variability: while some children are at ceiling, there are some ten-year-olds still responding at chance level. On the Affect tasks there are ceiling scores of 100%, even in the five-year-old group, but also some of the

lowest individual scores in the entire battery, particularly on the Output task (Figure 6).

Similarly on the Chunking Output task, in all age groups some children are scoring at ceiling while others are scoring around half marks (Figure 6). Dankovičová *et al*. (2004) analysed acoustically the responses of ten children from the eight-year-old group from the present study, selected at random. 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-list or 3-list). 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 (simple nouns: 3-list target), as opposed to the utterances without a prosodic boundary (compound noun: 2-list target). 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-list and 3-list 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. Thus there is variability both across children in the same age band and within the individual child.

(e) *Some intonation systems may never be acquired by some individuals*.

It was not the case that all the children in the oldest group performed at ceiling on all tasks. This may indicate that some aspects of intonation are acquired later than the age range covered in the present study. A further possibility is that some aspects of the intonation system as described in the classic studies of British English intonation are never acquired, or at least are not used consistently even by adults. On Chunking Output, although there was no developmental progression, there were 31 out of 119 children who scored below 75%: over a quarter of the children tested, distributed across the age range do not consistently use the expected pattern (Figure 6).

One area in particular that showed no developmental change in the present study was in the number of 'errors' in marking final narrow focus by means of accentuation. The most striking finding of the qualitative analysis of Focus Output was the high incidence of ambiguous responses, across all age groups, for items targeting final narrow (contrastive) focus (Table 9). For example, the unambiguous response to *How about a green bike?* would be *Oh I want a green *car*, with contrastive (narrow) focus on *car*. An ambiguous response is where there is (for instance) a final fall on

car, but this is not accompanied by a step up in pitch, or increased loudness or duration. This is a classic way of indicating broad focus over the whole utterance, rather than narrow focus on the final word. Alternatively, ambiguity may arise because more than one item is accented strongly (*I want a *green *car*). Although not predicted by theoretical accounts of English intonation, phonetic ambiguity in speakers' expression of final narrow focus has been reported to be quite common in the speech of adult speakers of Southern British English. Peppé *et al*. (2000) used an earlier version of the PEPS procedure with 90 adults from the same geographical area as the children in this study. They reported that a significant minority of adults made this kind of 'error' on the Focus task. There is thus a range of variation in the adult population, even from a single dialect area, which needs to be taken into account when considering children's acquisition.

CONCLUSIONS

The present study suggests that while five-year-old children have acquired many functional prosodic skills, there are further developments in prosodic comprehension between 5;0 and 8;7; and that some aspects of intonation continue to develop after that. Furthermore, functional prosodic comprehension correlates significantly with the development of other aspects of language. Despite these generalities, however, the picture of prosodic development presented here highlights variability. It appears that the age of acquisition of a specific prosodic ability may vary; levels of ability in a specific skill vary across children; and competence in different modes (comprehension and expression of prosody) may become evident at different ages. Such findings go some way towards explaining why the picture of intonational and prosodic development has been so unclear hitherto. Above all, our study suggests that it is unrealistic to examine one aspect of prosody (such as focus/accent) and assume that ability in this area is representative of all aspects of prosody. To gauge a particular child's stage of prosodic development it is necessary to establish what aspect of prosody is in question and to look at peer performance. From a practical perspective, such information is becoming increasingly important. In the UK, for example, there is growing emphasis on spoken language skills ('oracy') in the school years, with teachers being required to teach and assess pupils in this area. A description of what children might reasonably be expected to know at different ages, and of the degree of variability that might be expected within a demographically homogeneous group, thus forms useful background knowledge for education professionals.

Previous developmental research has tended to favour an experimental approach for investigating the comprehension of prosodic features, while production has been studied through analysis of more or less spontaneous

speech samples, as well as through experimental elicitation. In the present study, an experimental, test-based approach was used to investigate output as well as input. This has the advantage of facilitating comparisons between them. However, in this as in other areas of language development, questions remain about the relationship between children's ability as demonstrated by performance on tests and their ability as demonstrated by their competence in naturally occurring interactions. It may be that the analysis of children's production of intonation and orientation to others' use of intonation in spontaneous interaction will reveal a somewhat different picture. By combining the two methodologies we should arrive at a fuller understanding of this neglected but communicatively important aspect of children's language development.

REFERENCES

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**, 1148–60.

Bishop, D. (1989). *Test for Reception of Grammar* (*TROG*). Second edition. Manchester: Department of Psychology, University of Manchester.

Branigan, G. (1979). Some reasons why successive single word utterances are not. *Journal of Child Language* **6**, 411–21.

Capelli, C., Nakagawa, N. & Madden, C. (1990). How children understand sarcasm: the role of context and intonation. *Child Development* **61**, 1824–41.

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

Couper-Kuhlen, E. & Selting, M. (1996). *Prosody in conversation: interactional studies*. Cambridge: CUP.

Cruttenden, A. (1985). Intonation comprehension in ten-year-olds. *Journal of Child Language* **12**, 643–61.

Crystal, D. (1986). Prosodic development. In P. Fletcher & M. Garman (eds), *Language acquisition: studies in first language development*. Second edition. Cambridge: CUP.

Crystal, D. (1987). *Clinical linguistics*. Second edition. London: Edward Arnold.

Cutler, A. & Swinney, D. (1987). Prosody and the development of comprehension. *Journal of Child Language* **14**, 145–67.

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.

Ferrand, C. & Bloom, R. (1996). Gender differences in children's intonational patterns. *Journal of Voice* **10**, 284–91.

Flax, J., Lahey, M., Harris, K. & Boothroyd, A. (1991). Relations between prosodic variables and communicative functions. *Journal of Child Language* **18**, 3–19.

Ford, C. E. & Thompson, S. A. (1996). Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E. A. Schegloff & S. A. Thompson (eds), *Interaction and grammar*. Cambridge: CUP.

Halliday, M. A. K. (1975). *Learning how to mean: explorations in the development of language*. London: Edward Arnold.

Hornby, P. (1971). Surface structure and the topic–comment distinction: a developmental study. *Child Development* **42**, 1975–88.

Hornby, P. & Hass, W. (1970). Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research* **13**, 395–99.

House, J., Dankovičová, J. & Huckvale, M. (1999). Intonation modelling in ProSynth. *Speech, Hearing and Language: Work in Progress*, Department of Phonetics and Linguistics, University College London 11, 51–61.
Howell, D. C. (2002). *Statistical methods for psychology*. Fifth edition. Pacific Grove, CA: Duxbury Press.
Katz, W., Beach, C., 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**, 3179–91.
Ladd, D. R. (1996). *Intonational phonology*. Cambridge: CUP.
Lloyd, P., Mann, S. & Peers, I. (1998). The growth of speaker and listener skills from five to eleven years. *First Language* **18**, 81–103.
Local, J. (1980). Modelling intonational variability in children's speech. In S. Romaine (ed.), *Sociolinguistic variation in speech communities*. London: Edward Arnold.
Macwhinney, B. & Bates, E. (1978). Sentential devices for conveying givenness and newness: a cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behaviour* **17**, 539–58.
Nathan, L., Wells, B. & Donlan, C. (1998). Children's comprehension of unfamiliar regional accents: a preliminary investigation. *Journal of Child Language* **25**, 343–65.
Peppé, S. (1998). *Investigating linguistic prosodic ability in adult speakers of English*. Unpublished PhD thesis, University College London.
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 & Phonetics* **17**, 345–54.
Peppé, S., Maxim, J. & Wells, B. (2000). Prosodic variation in Southern British English. *Language and Speech* **43**, 309–34.
Semel, E., Wiig, E. & Secord, W. (1987). *Clinical evaluation of language fundamentals – revised*. London: The Psychological Corporation.
Siegel, S. & Castellan, N. J. (1988). *Nonparametric statistics for the behavioural sciences*. Second edition. New York: McGraw-Hill.
Snow, D. (1998). Children's imitations of intonation contours: are rising tones more difficult than falling tones? *Journal of Speech, Language and Hearing Research* **41**, 576–87.
Van der Meulen, S., Janssen, P. & den Os, E. (1997). Prosodic abilities in children with specific language impairment. *Journal of Communication Disorders* **30**, 155–70.
Wells, B. & Local, J. (1993). The sense of an ending: a case of prosodic delay. *Clinical Linguistics & Phonetics* **7.1**, 59–73.
Wells, B. & Peppé, S. (2001). Intonation within a psycholinguistic framework. In J. Stackhouse & B. Wells (eds), *Children's speech and literacy difficulties 2: identification and intervention*. London: Whurr Publishers.
Wells, B. & Peppé, S. (2003). Intonation abilities of children with speech and language impairments. *Journal of Speech Language & Hearing Research* **46**, 5–20.
Whiteside, S. & Hodgson, C. (2000). Some acoustic characteristics in the voices of 6- to 10-year-old children and adults: a comparative sex and developmental perspective. *Logopedics Phoniatrics Vocology* **25**, 122–32.
Wieman, L. A. (1976). Stress patterns of early child language. *Journal of Child Language* **3**, 283–86.
Winner, E. & Leekam, S. (1991). Distinguishing irony from deception: understanding the speaker's second-order intention. *British Journal of Developmental Psychology* **9**, 257–70.