

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

This full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/15582>

Please be advised that this information was generated on 2014-11-11 and may be subject to change.

## Prosody and the development of comprehension\*

ANNE CUTLER

*MRC Applied Psychology Unit*

AND

DAVID A. SWINNEY

*Tufts University*

(Received 28 April 1986)

### ABSTRACT

Four studies are reported in which young children's response time to detect word targets was measured. Children under about six years of age did not show the response time advantage for accented target words which adult listeners show. When semantic focus of the target word was manipulated independently of accent, children of about five years of age showed an adult-like response time advantage for focussed targets, but children younger than five did not. It is argued that the processing advantage for accented words reflects the semantic role of accent as an expression of sentence focus. Processing advantages for accented words depend on the prior development of representations of sentence semantic structure, including the concept of focus. The previous literature on the development of prosodic competence shows an apparent anomaly in that young children's productive skills appear to outstrip their receptive skills; however, this anomaly disappears if very young children's prosody is assumed to be produced without an underlying representation of the relationship between prosody and semantics.

### INTRODUCTION

A review of the literature on the acquisition of prosodic competence reveals an intriguing paradox: in certain respects, children's prosodic productions

---

[\*] Experiments one to three were reported variously to the Midwestern Psychological Society in May 1979 and to the Fifth Annual Boston University Conference on Language Development in October 1980. These earlier experiments were supported by grants from the Science Research Council (UK) and the National Science Foundation. Experiment four was supported by a grant from the National Institute of Health.

The authors thank Judith Goodman, David Gow, and Scott Tucker for technical assistance. Particular thanks go to Penny Prather for invaluable help and useful discussion. Finally, very many thanks to Dwight Bolinger for constant encouragement and for stimulating a long-overdue written report of this work. Address for correspondence: Anne Cutler, MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, UK.

appear to be more advanced than their prosodic understanding. How can this be? Children do not spontaneously develop linguistic prosody without suitable input, any more than they spontaneously develop other aspects of language without suitable input. In general, receptive abilities far outstrip productive abilities, a fact which supports the general hypothesis that children will learn to produce exactly and only those linguistic features which they hear in the speech to which they are exposed.

Yet some aspects of prosodic competence appear to be exceptions to this rule, as will be clear from the summary of the available evidence below. First, however, it should be pointed out that studying the acquisition of prosody is a little different from studying the acquisition of other aspects of language. 'Prosody' refers to a complex of suprasegmental features, including lexical stress, phrase and compound stress, sentence accent, utterance rhythm and utterance intonation. These features are realized in the speech wave in the dimensions of segment duration, amplitude and fundamental frequency. But, of course, any spoken utterance must be realized in these dimensions. Utterances without prosody are simply impossible. Thus studying the acquisition of prosody is not a matter of recording the gradual increase in produced prosodic features (as, for instance, studying the acquisition of syntax involves recording the production of identifiable words, of multiword utterances, of negation, inflections etc.). Instead, studying the acquisition of prosody involves assessing the CORRECTNESS of the prosody which is produced with respect to the adult model.<sup>1</sup>

Studies of prosodic production in young children have shown that even at a very early age children can successfully use prosody as part of their linguistic competence. Allen & Hawkins' (1978) study of five 3-year-olds showed them to be clearly in command of aspects of utterance rhythm and intonation – they made systematic distinctions between strong and weak syllables, and marked phrase boundaries with appropriate phrase final intonation. Smith (1978) likewise found that the underlying rhythmic structure of children's utterances was analogous to that of adult utterances. Even 2-year-olds can use prosody communicatively, e.g. Furrow (1984) found children of this age consistently distinguishing utterances in which eye contact was made from other utterances by the use of different prosody. Tonkova-Yampol'skaya (1973) found similarly that infants' cries which had different underlying intentions were prosodically distinguished. The same observation is used to argue that children's communicative competence at the two-word stage is greater than

---

[1] This view ignores the fact that there are theories of prosodic phonology which treat the prosodic repertoire as a system of discrete units, so that the acquisition of prosody can be treated as the acquisition of a series of units or functions in much the same way as the acquisition of syntactic units and functions is described. However, such analyses of prosodic acquisition are theory-dependent in a way the study of syntax acquisition is not.

would appear from their syntactic competence: identical two-word strings may be uttered in completely different contexts with quite distinct prosodic patterns, and it is claimed that they express quite different propositions (Miller & Ervin 1964, Brown 1973). Klein's (1984) study of a 2-year-old's lexical stress patterns found that although this child had considerable difficulty imitating lexical stress patterns, his spontaneous productions of words familiar to him used consistent and correct stress placement.

At the utterance stress level, MacWhinney & Bates (1978) showed that children can use stress to distinguish new from given information by age three; Wieman (1976) showed similarly that stress patterns in two-word utterances clearly reflect semantic structure; and Hornby & Hass (1970) found that 4-year-olds correctly assigned contrastive stress in a picture description task.

The comprehension evidence is less clear-cut. Even very young infants show some sensitivity to prosodic aspects of the speech of adults (Morse 1972, Mehler & Bertoncini 1979). Prosodic structure seems to play an important role in many aspects of acquisition, e.g. Gleitman & Wanner (1982) have argued persuasively that crosslinguistic asymmetries in the acquisition of certain morphosyntactic features can be explained as the universal application of a strategy 'pay attention to stressed syllables'. Thus language-specific interrelations of stress patterns and morphology underlie language-specific acquisition patterns.

The prosodic perception of young children is in some cases BETTER than might be predicted; Allen (1983) showed that 4-year-old French children can correctly perceive stress contrasts not found in their language (but typical of English and other stress languages), whereas the same children at age five, more fully in command of the prosodic structures of their own language, can no longer reliably distinguish the non-native contrasts.

The most noticeable difference between the production and the comprehension evidence, however, is that in comprehension, sentence prosodic features, in particular sentence stress patterns, have repeatedly been shown to be very poorly processed by young children. Lahey (1974) found that children who were asked to act out complex sentences were not significantly worse at doing so when the sentences were presented with a monotonous list-type intonation than when all the customary sentence prosody cues were available. Bates (1976) similarly showed that children's imitation of sentences was clearly disrupted by departures from canonical agent-action-object word order, but was not disrupted by highly inappropriate focal stress assignment, and Bosshardt & Hörmann (1982) found that sentence repetition by 4- to 6-year-olds was (in contrast to repetition by adults) quite unaffected by abnormal prosody. Hornby (1971), in a parallel study to the experiment by Hornby & Hass (1970) cited above, found that first- and third-graders performed essentially at chance in interpreting stress cues to topic-comment

structure. MacWhinney & Price (1980) replicated and extended this finding. MacWhinney, Pleh & Bates (1985), in a study of sentence understanding in Hungarian, found that 6-year-olds could use stress as a cue to thematic role assignment almost as efficiently as adults could; 3- and 4-year-olds, however, failed to make use of the stress cue. Scholes, Tanis & Turner (1976) found 5-year-olds unable to use disjuncture as a syntactic disambiguating cue, a result which has also been replicated (Cowin, Mann, Schoenheimer & Berman 1984). Finally, Solan (1980) found that 5-year-olds failed to make correct use of stress cues to pronominal reference.

Perhaps the most conclusive demonstration that comprehension of sentence-level prosody is acquired relatively late is Cruttenden's (1974, 1985) finding that even 9- and 10-year-olds fall well short of adult performance in correctly interpreting the function of certain intonation contours in context.

The paradoxical advantage of production over comprehension has even been demonstrated in the same children. Hornby (1971) found that his subjects who performed at chance in using stress cues in the comprehension of topic-comment structure nevertheless produced the same cues appropriately; and Atkinson-King (1973) found that children who could reliably produce compound stress distinctions of the 'blackbird - black bird' type could not reliably perceive the same distinctions.

The performance paradox therefore seems to be chiefly associated with phrase- and sentence-level prosody, and chiefly noticeable in children of around the pre-school/first-grade age (5-7 years). Children at this stage can produce sentence prosody, particularly sentence accent patterns, which sound to adults as if they are entirely appropriate to the sentence semantics; but the same children appear not to extract semantic information from the sentence prosody, or even to process it at all (since ill-formed prosody does not disrupt their comprehension). Such an unusual advantage of production over comprehension demands explanation.

There is, however, always the possibility that the body of available evidence is misleading. Specifically, it could quite easily be the case that many studies have underestimated children's comprehension abilities. With adults, it is normal to assess comprehension performance 'on-line', i.e. to use response time measures of understanding. Such measures are usually considered to produce a purer picture of factors affecting comprehension than 'off-line' measures such as correctness of question-answering or the like, in which there are considered to be rather more intervening processes between input and response. But response time techniques are - for obvious reasons - extremely unusual in the study of children's comprehension. None of the prosodic comprehension studies cited used on-line techniques. It could therefore be the case that the failure to find certain prosodic processing effects in young

children is due simply to insensitivity of the tasks which were used: perhaps the children could process the prosodic structure, but could not carry the results of this processing through the necessary additional cognitive stages to the desired response. In other words, the paradox might disappear if comprehension were tested with an on-line task.

There is one on-line task which has recently been used successfully to study children's comprehension. This is the word-monitoring task, in which subjects listen for a specified word target, and press a response key as soon as this word occurs in the sentence or passage to which they are listening. Foss, Bias & Starkey (1978) and Tyler & Marslen-Wilson (1981) have successfully used this task with pre-school populations. The word-monitoring task therefore seems a suitable on-line measure for the study of young children's use of prosodic structure in comprehension.

With an analogous monitoring task, phoneme-monitoring (in which subjects listen for words beginning with a specified target sound), it has been established that adult listeners make very active use of the prosodic structure of speech, in particular to direct their attention towards the most important parts of incoming messages. This series of studies showed, firstly, that target-bearing words which carry sentence accent are responded to consistently faster than target-bearing words which are not accented (Cutler & Foss 1977). For instance, subjects listening for the phoneme /k/ in (1) will respond faster if they hear version (1*a*) than if they hear version (1*b*) (words bearing sentence accent are in small capitals):

- (1*a*) Does John really want to KEEP that old van?
- (1*b*) Does John really WANT to keep that old van?

Furthermore, the reaction time advantage of accented words is not due merely to acoustic factors: differences in the prosodic structure of the part of the sentence preceding the target indicate where accent will fall, and suffice to direct listeners' attention to the highly stressed words. In another experiment, sentences like (2) were recorded in two prosodic versions, with the target-bearing word accented in one version and unaccented in the other:

- (2*a*) She managed to remove the DIRT from the rug, but not the BERRY stains.
- (2*b*) She managed to remove the dirt from the RUG, but not from their CLOTHES.

The target-bearing word itself (in this example, 'dirt') was then edited out of each recording and replaced by acoustically identical copies of the same word taken from a third, relatively neutral, recording of the sentence. Thus the experimental versions of each sentence had acoustically identical target

words, but in one case the prosodic contour of the rest of the sentence was appropriate for an accented word in the target word's location, while in the other version the contour was appropriate for the occurrence of an unaccented word at that point. Under these circumstances targets in accented position still produced faster response times than targets in unaccented position (Cutler 1976). Since the target words themselves were not different, the listeners must have been making use of the differences in prosody which preceded the target.

Using prosodic cues in order to locate accented words ought to be a useful sentence comprehension strategy for the following reason: speakers usually accent the most important parts of their message. Thus, for a listener, locating accented words is equivalent to locating the semantically most central part – the focus – of the utterance. Not surprisingly, when sentence focus is varied independently of prosodic contour, it is found that listeners respond more rapidly to focussed than to non-focussed target words. In an experiment by Cutler and Fodor (1979) listeners heard sentences like (5) preceded by one of two alternative questions, of which one – in this instance (3) – focussed attention on the first part of the sentence, while the other – (4) – focussed attention on the last part of the sentence.

- (3) Which woman was it that went into the office?
- (4) Which office was it that the women went into?
- (5) The woman with the bag went into the dentist's office.

There were two alternative targets – in (5) the /b/ of 'bag' or the /d/ of 'dentist's – and the sentence itself remained acoustically identical irrespective of which preceding question or which target specification a particular subject heard. Focussed targets produced consistently faster responses – thus in (5), subjects listening for /b/ responded faster if they had heard question (3) than if they had heard question (4), whereas the reverse was true for subjects listening for /d/: (4) produced faster responses than (3). Varying sentence accent and varying sentence focus thus produce analogous effects on phoneme-monitoring response time, which suggests that the same effect may be involved in both cases: the reason listeners make such good use of cues to accent is that accented words are focussed words.

In adult listeners the use of prosodic information during sentence comprehension is obviously well developed; prosodic cues to accent provide an effective pointer to the sentence's focus. The word-monitoring task is a much more sensitive procedure than any previously used to study children's prosodic processing. It is possible that this on-line technique will show that children who can produce adequate sentence accent patterns can also use these patterns in comprehension.

Our first experiment was a direct analogue of the first phoneme-monitoring study described above, in which adult listeners showed a response time

advantage for accented target words. Our child subjects heard sentences like (6) and (7):

(6*a*) The nurse brought a CLEAN towel and took away the DIRTY one.

(6*b*) The nurse brought a clean TOWEL and took away the dirty one.

(7*a*) The family is already AT the summer cabin.

(7*b*) The family is ALREADY at the summer cabin.

In these examples the word targets were 'clean' and 'at' respectively.

Because our review of the literature did not motivate precise predictions about prosodic processing abilities at specific ages, we tested a fairly wide selection of pre-school and early school-age children, from four to eight years of age.

## EXPERIMENT 1

### METHOD

#### *Materials*

Sixteen experimental sentences were constructed, of which (6) and (7) are examples; half of the sentences, like (6), contained open class (content word) targets, while in the remaining sentences, like (7), the target was a closed class (function) word. The target words, all monosyllabic and of high frequency, were: *ball, door, coat, keep, scrub, bite, old, clean; can, is, at, in, but, the, this, and my*. Each sentence had two prosodic versions, one in which the target word was the most prominent word in the sentence, i.e. received primary sentence accent, and one in which the primary accent fell elsewhere than on the target word.

A tape was recorded (by a speaker of standard American), containing both prosodic versions of each experimental sentence plus six filler sentences without occurrences of the specified targets. This was to ensure that subjects did not just wait and press the button at the end of each sentence. One version of each sentence occurred in each half of the tape, and the prosody (accented versus unaccented targets) and word class (open versus closed class targets) conditions were well mixed in the presentation order. The target word for each sentence was specified on the tape immediately prior to the sentence. The experimental set was preceded by a set of practice sentences.

#### *Subjects*

Child subjects were students at the Eliot Pearson Nursery School, affiliated with Tufts University, and at a primary school in the Weston, Massachusetts School District. After elimination of a few subjects who failed to produce an acceptable number of correct responses to the comprehension questions (see below), we were left with 21 child subjects between the ages of four and eight years. Ten undergraduates at Tufts University formed an initial adult control



group. Because the response times produced by these subjects with the simple materials of this experiment were very fast, a further eight adult controls from the same population were tested with the same materials presented under a light distorting mask of white noise (signal-to-noise ratio 16dB).

### *Procedure*

The child subjects were tested at their schools, the adults at the Tufts University psychology department. All subjects were tested individually and heard the sentences over headphones. They were instructed to understand the sentences and also to listen for the target word in the sentence and to press the response button as soon as they heard the specified word. The response button for the child subjects was especially enlarged (1 in. diameter). Subjects were given practice trials (up to 10) until it seemed they understood the task. Response time was measured from the onset of the target word. Response times were displayed on a visual display and recorded by the experimenter.

Comprehension was checked by asking simple questions after 15 of the experimental sentences. Only subjects who gave at least 12 correct answers were included in the data analysis.

### RESULTS

Preliminary inspection of the data suggested that age of subject was strongly related to the pattern of results in the child group. Accordingly we divided our child subjects into two groups, an older group (ages 6;5 to 7;11) containing 11 subjects, and a younger group (ages 4;0 to 6;1) containing 10 subjects. Mean response times for each condition were computed for each of these two groups and for the two adult control groups (ages 18;1 to 22;2 for the No Noise condition, and 18;7 to 21;3 for the Noise condition). These means are shown in Table 1.

The results for the adult control groups replicated the results of Cutler & Foss (1977). Response times for both groups were faster to accented than to unaccented words ( $F(1,9) = 15.2$ ,  $P < 0.003$  for the No Noise condition,  $F(1,7) = 14.3$ ,  $P < 0.01$  for the Noise condition), but neither the difference between open and closed class target words nor the interaction of the word class and accent variables reached significance for either group. These results indicate that for the adult subjects the word-monitoring task produces the same pattern of results as the phoneme-monitoring task used by Cutler & Foss. (Another study, investigating the word class and accent effects in the comprehension of aphasic patients, used the word-monitoring task with a group of hospitalized normal controls and again found results very similar to those found for adults in the present study – Swinney, Zurif & Cutler 1980.) The fact that the results for both control groups are essentially the same, although adding noise to the presented stimuli added some 200 msec

PROSODY AND COMPREHENSION

TABLE 1. *Mean response times (msec) for each subject group and each condition. Experiment 1*

	Adults: No noise condition ( $N = 10$ )			Adults: Noise condition ( $N = 8$ )		
	Open	Closed	$\bar{x}$	Open	Closed	$\bar{x}$
Accented	203	178	191	423	390	407
Unaccented	239	268	254	439	483	461
$\bar{x}$	221	223		431	437	
	Older children ( $N = 11$ )			Younger children ( $N = 10$ )		
Accented	317	347	332	601	748	675
Unaccented	432	506	469	579	762	671
$\bar{x}$	375	427		590	755	

to the average response time, suggests that the results for the No Noise condition are not simply so fast that differences between open and closed class target words are minimized (a 'floor effect'); there is a main effect of accent, but no main effect of word class, even when the task is made more difficult by degrading the stimuli. Thus the word-monitoring task appears to be tapping the same comprehension processes as the phoneme-monitoring task.

The results for the child groups are different. The older children showed a significant effect of accent ( $F(1,10) = 46.96, P < 0.001$ ), but also a significant effect of word class ( $F(1,10) = 5.46, P < 0.05$ ). (The interaction between the two factors was not significant.)

The younger children, on the other hand, showed a significant effect of word class ( $F(1,9) = 12.32, P < 0.01$ ), but no effect of accent ( $F < 1$ ) (and again no interaction).

DISCUSSION

Both the child groups showed a non-adult pattern of results. It appears that closed class words do not present an easy processing task for children's comprehension. Certainly the difficulty of processing them is not essentially equivalent to the difficulty presented by open class words, as is the case for adults' comprehension. We must conclude from this aspect of our results that the specialized word recognition processes which adults use for closed class words have a developmental history, and in fact do not fully develop till quite late – until after age seven, at least. This is in fact in line with a good deal of other evidence in the language acquisition literature, particularly concerning the most widely-studied category of closed class words, pronouns.

Solan (1980), for instance, found that pronominal reference is very difficult for 5-year-olds to process. Karmiloff-Smith (1979) concluded that full grasp of anaphoric pronouns is not achieved until about age nine.

A more surprising aspect of our results is the failure of the youngest group to show a response time effect of sentence accent. It might be argued that this null result is simply a 'ceiling effect' – the response times are longer for this group, and it might be suggested that the variability is also higher, to such an extent that much greater mean differences would be required to reach the required level of statistical significance. But this cannot be so, since on the one hand the word class effect is clearly significant, and on the other hand there is not the slightest trace of an accent effect.

The reason that this lack of an effect is particularly surprising is that the acoustic differences between accented and unaccented target words alone should give some response time advantage for accented targets. Stressed words are typically longer and louder and express more pitch movement than unstressed words. For these reasons they are acoustically clearer. Although in the case of open class words the unaccented targets may still have retained rhythmic stress, so that adding accent could only improve perceptibility by a negligible amount, this was not the case with the closed class words; in unaccented position they were definitely unstressed, so that their simple perceptibility should have been much greater in accented position. (It will be noted that although the interaction of word class and accent did not reach significance, both adult control groups showed a greater accent effect for closed class than for open class target words. This was also the case for the older child group. It was also, as it happens, the case for both experiments conducted by Cutler & Foss (1977), and for the normal control group of Swinney *et al.* (1980). Simple acoustic differences are responsible for this highly consistent pattern.)

In order to test whether our subject population was insensitive to stress effects as a whole or simply to accent effects in sentences, we performed a simple control experiment: we presented subjects from the same population with the same monitoring task, using the same materials reordered such that they formed lists, not sentences.

## EXPERIMENT 2

### METHOD

#### *Materials*

The 38 sentences of Experiment 1 were each separately scrambled so that they formed 38 separate syntactically ill-formed lists. Thus sentence (6) above became 'The took and a clean brought the nurse one dirty away towel'. The targets were as in Experiment 1. Each target occurred in the same position in the list as it had in the original sentence, and the lists occurred in the same

order as the sentences had. The lists were read at a normal rate but without sentence prosody; each list occurred in two versions, in one of which the target word was 'stressed' (i.e. spoken with longer duration, higher pitch and greater intensity than in the 'unstressed' production).

### *Subjects & Procedure*

Ten subjects, aged from 5;0 to 7;1, from the same children's school as in Experiment 1 took part. The procedure was as in Experiment 1 except that no comprehension questions were asked.

TABLE 2. *Mean response time (msec) per condition. Experiment 2*

	Open	Closed	$\bar{x}$
'Accented'	468	531	500
'Unaccented'	497	577	537
$\bar{x}$	483	554	

### RESULTS

Mean response times for each condition are shown in Table 2. The main effects for word class ( $F(1,9) = 12.6$ ,  $P < 0.01$ ) and for stress level ( $F(1,9) = 9.3$ ,  $P < 0.05$ ) were both significant; the interaction of the two variables was not significant ( $F = 1.3$ ).

### DISCUSSION

The results of Experiment 2 suggest that it is unlikely that our younger subjects in Experiment 1 were insensitive to stress effects *per se*. Simple perceptibility effects do show up when the materials are not presented in sentence form. Why then do they not show the same effects when they are processing sentences? (Recall that we do know that they ARE processing the sentences – only subjects who passed our comprehension test were included in the data analysis.)

The answer is provided by the body of evidence we cited at the beginning of this paper. The more sensitive on-line measure attests to the conclusion derived from off-line measures: children of this age group are very inefficient at processing prosodic cues of any kind in sentences. In fact, they are rather inefficient at processing all aspects of language which pertain to the sentence – syntax and pragmatics are very imperfectly exploited. As Tyler & Marslen-Wilson (1978) have argued, for children of this age group semantics over-rides all else in comprehension – however, they have not yet learned that processing

syntax and prosody can be part of processing semantics. By default, lexical semantics must be the strongest factor in their comprehension. Thus it is not surprising that they show a strong word class effect – open class words are ‘weightier’ with respect to lexical semantics than are closed class words. The effect of varying accent position, however, is strictly a manipulation of sentence semantics – lexical semantics will remain unaffected by such variations. Therefore such effects are of little importance in young children’s comprehension.

But this claim in its strongest form is circular: it would imply that young children process sentences as if they were lists of unrelated words. If this were the case, we would have expected precisely the same results for this age group in Experiment 1, in which the materials were sentences, and Experiment 2, in which the materials actually were lists of unrelated words. In order to obtain a clearer picture of the processing of sentence semantics in the age group four to six, our next experiment explicitly manipulated the variable of sentence focus. As in the adult phoneme-monitoring study (Cutler & Fodor 1979) described in the introduction, focus was manipulated by varying a question which preceded the sentence in which the target word occurred.

### EXPERIMENT 3

#### METHOD

##### *Materials*

In order that the target bearing sentence and its preceding focussing question should not be considered as separate entities, the materials of this experiment consisted of brief stories. An example story is given in Appendix I. Six such stories were constructed. The word targets for which subjects listened were the names of characters in the story. Each story had two potential target names, and each name occurred five times in its particular story – once in the opening sentence (counted as a practice item) and thereafter twice in focussed position, twice in non-focussed position. Focus was varied by means of questions which preceded the sentences containing occurrences of the targets. Each story had two versions, and each individual target occurrence was focussed in one version and non-focussed in the other. Thus for the example story in Appendix I, one version contained the upper alternative questions, while the other version contained the lower alternatives. The two versions of each story were created in the following manner: one master tape was recorded of each story, in which the reader (the same male speaker of standard American) read out both alternative questions before the sentence containing a target. For the example story, for instance, he read ‘...with big windows. Which one lived in the house with chimneys? Which one lived in the house with big windows? The house that Patti lived in had big windows, and Jenni’s house had many chimneys...’ and so on. Two copies of this master tape were

then made, and the experimental versions of the stories created by splicing the copies to remove the unwanted alternative questions. Thus each version contained the same rendition of each target sentence irrespective of the preceding question.

### *Subjects*

Subjects were 44 children varying in age from 3;0 to 7;10. (Four of these subjects were over the upper limit of our age range four to six, and their data were not included in the analysis.<sup>2</sup> Data from 12 further subjects were discarded because they failed to reach criterion performance either on comprehension or on target detection.) Ten subjects were students at the Eliot Pearson Nursery School at Tufts University. The remaining subjects were drawn from the neighbourhoods around Tufts University. In addition, 12 Tufts University undergraduates served as an adult control group.

### *Procedure*

Child subjects were tested at their schools or play groups, adult subjects in the Tufts University psychology department. Subjects were tested individually. At the beginning of testing the children were told that they would be playing a listening game. They were instructed to place their hands on the response key, and to press it as quickly as they could whenever they heard certain names. Each subject heard only one of the four versions of each of the six stories. Before each story subjects were told which name they were supposed to listen for. If they successfully responded to the (dummy) target in the first sentence they heard the whole story without interruptions. If they failed to respond to this first target, they were given the instructions again, and then heard the first sentence a second time. If they failed on second attempts in two stories they were judged unable to perform the word-monitoring task. Seven child subjects were rejected from the experiment for this season.

Comprehension was tested by questions at the end of each passage. There were four groups of subjects since for each story there were four possible combinations of story version (first or second) with name target (A or B).

### RESULTS

No subject in this experiment failed to produce acceptable answers to the comprehension questions. However, five child subjects produced an unacceptably high error rate, missing nine or more experimental target occurrences

---

[2] Each of these older children showed a significant response time advantage for focussed targets.

TABLE 3. Mean response time (msec) for each subject group and each condition. Experiment 3

Adults	Focussed	397
	Non-focussed	469
Children	Focussed	1674
	Non-focussed	1708

out of a total of 24. The data for these subjects were discarded, leaving a total of 28 child subjects and 12 adult controls. Mean response times for the adult and child groups as a whole are presented in Table 3.

It can immediately be seen that the overall mean response time for each group is considerably longer than in Experiment 1. We suggest that this is simply an effect of the increased difficulty of monitoring for a word target in connected prose as opposed to isolated sentences. (A similar effect has been found with phoneme-monitoring by Rudnicky 1980.) Comprehension of a story is rather more interesting, and requires more higher-level integrative processing, than comprehension of a sequence of unrelated sentences; this necessarily distracts some attention from the monitoring task, even for adult subjects.

The focus effect was significant, as predicted, for the adult group ( $F(1, 11) = 13.9, P < 0.01$ ). It was not significant for the child group as a whole. However, inspection of the individual results suggested that there was a clear tendency for a focus effect to appear with increasing age. Accordingly we divided the 28 children by age, into three roughly equal groups of about the same size as we had used in Experiment 1: 3;0-4;6, 4;7-4;11 and 5;0-5;8. The means broken down in this manner are shown in Table 4. Analysis of the focus effect showed that it was not significant for either of the two younger groups ( $t(9) = -0.10, P > 0.9$  for the youngest group,<sup>3</sup>  $t(8) = -0.12, P > 0.9$  for the middle group), but was significant for the five-year-old children ( $t(8) = 2.47, P < 0.04$ ).

#### DISCUSSION

It would appear from the results of this experiment that the processing advantage for focussed words is not fully developed in pre-school children; like the processing advantage for accented words, it develops some time between the age of four and six.

[3] This group contained nine children in the age range 4;2-4;6, and one bright younger child of 3;0. Removing the 3-year-old's data makes no difference to the levels of significance, and leaves the two condition means for this group still only 3 msec apart.

PROSODY AND COMPREHENSION

TABLE 4. *Mean response times (msec) per condition for three child age ranges. Experiment 3*

Age 3;0-4;6 ( <i>N</i> = 10)	
Focussed	2114
Non-focussed	2105
Age 4;7-4;11 ( <i>N</i> = 9)	
Focussed	1546
Non-focussed	1532
Age 5;0-5;8 ( <i>N</i> = 9)	
Focussed	1363
Non-focussed	1488

As we have argued on the basis of the adult studies, the reason that accented words enjoy a processing advantage is precisely because they represent the semantic focus of the sentence; detecting the semantic focus as rapidly as possible is a useful strategy in sentence comprehension. Thus it is not surprising that where there is no processing advantage for focus, there is no processing advantage for accent.

The subjects we used in our focus study (Experiment 3), however, were all from the age range which in our earlier study (Experiment 1) failed to show an accent effect, i.e. the under 6-year-olds. Yet the 5-year-olds in this group showed a significant processing advantage for focus. Although our samples have been small – 21 and 28 children respectively – we might nevertheless advance a tentative conclusion that the focus effect appears before the accent effect.

This is of course precisely what one would expect on the basis of our account of the accent effect in the first place. If the accent effect is indeed an indirect effect of focus – attention is directed towards accented words precisely because they tend to be focussed, i.e. important information – then one would expect its genesis to be parasitic on the earlier appearance of a focus effect. In other words, listeners must develop the strategy of searching for focussed information first, because only the existence of this strategy will prompt the development of a prosodic processing strategy as an efficient way of realizing this search.

In an attempt to test this developmental order claim, we carried out a small follow-up study in which we tried to assess the accent and focus effects simultaneously in a group of children of the relevant age range.

This study allowed us to control one further factor. Experiments 1 and 3 differed in one important procedural aspect: Experiment 1 used isolated sentences, Experiment 3 continuous prose. Although the response time distributions clearly indicate that monitoring for word targets in stories is on



the whole more difficult than in sentences, it could be the case that effects of the kind we are looking at are somewhat more likely to show up in continuous prose comprehension. In Experiment 4 we assessed effects of accent by using stories of exactly the kind used in Experiment 3, and compared these effects with focus effects in the same subjects.

## EXPERIMENT 4

### METHOD

#### *Materials*

The materials for this experiment consisted of three of the stories used in Experiment 3, along with three new stories constructed for the present experiment. An example of these new 'accent stories' is given in Appendix II. The new stories were very similar to the focus stories in that each story had two versions, and for each story there were two potential name targets, each name occurring five times in its story, once in the first sentence as a dummy target, twice in accented position and twice in unaccented position. Accent was counterbalanced between the two versions of each story.

A single tape was constructed containing all six stories, with accent and focus stories alternating.

#### *Subjects*

Eight children from the same university nursery school took part in the experiment; their ages were: 4;5, 4;6, 4;6, 4;8, 4;11, 5;2, 5;6, 5;11.

#### *Procedure*

The procedure was as in Experiment 3. No subjects were rejected from the experiment.

### RESULTS AND DISCUSSION

The means for each story type are shown in Table 5. Both types of story produced significant effects: accented targets were responded to significantly faster than unaccented targets ( $t(7) = 3.37$ ,  $P < 0.02$ ) and focussed targets were responded to significantly faster than non-focussed targets ( $t(7) = 5.45$ ,  $P < 0.001$ ).

Thus this small study unfortunately fails to give a conclusive answer to the question of development order. Our hypothesis would predict that the focus effect would be rather stronger in a group of this age than the accent effect, simply because we hypothesize the focus effect to be logically and hence developmentally prior to the accent effect. The actual  $t$  ratios shown above certainly indicate that the variability is rather higher with the accent stories, and inspection of the individual means gives support to this suggestion: focussed targets were responded to faster than non-focussed targets by all

TABLE 5. *Mean response times (msec) for each condition. Experiment 4*

Accent stories	
Accented	628
Unaccented	648
Focus stories	
Focussed	932
Non-focussed	956

subjects, but not all subjects showed an accent effect; and five subjects showed a considerably larger focus effect than accent effect, for two subjects the size of the two effects was almost identical, while only one subject – the second oldest – showed a considerably larger accent effect than focus effect. We feel certain that the greater strength of the focus effect in this age group would be apparent in a larger-scale study.

#### GENERAL DISCUSSION

The first conclusion to be drawn from this series of experiments is that our summary of the prior evidence is confirmed: children below the age of about five or six are poor at exploiting prosodic information in language comprehension. In contrast to earlier studies, we used an on-line measure in order to assess children's performance of the actual process of comprehension. Even this much more sensitive task failed to show evidence of adult-like prosodic processing, and thus provided supporting evidence for the previous studies which had failed to find prosodic processing in comprehension using various 'off-line' tasks.

Secondly, we have shown that even the processing of utterance semantic structure (considered separately from prosodic cues to it) is poorly achieved by younger children. Within the age groups we tested, it was only children over the age of about four-and-a-half who reliably showed evidence of directing processing attention preferentially to focussed rather than non-focussed parts of utterances, in the way that adults do.

Furthermore, our data are also consistent with our third conclusion, that speeded processing of accented words is a strategy which is developed only after speeded processing of focussed words has been incorporated into the comprehension repertoire. We showed that across different groups of children the age at which the 'focus effect' appeared was somewhat earlier than the age at which the 'accent effect' appeared, and a small within-group study suggested that the accent effect was rather weaker than the focus effect. This conclusion is perhaps not so vital: what matters is that we have produced further evidence for the fact that children in the age range four to six are

clearly engaged on the task of developing efficient routines for comprehending the semantic structure of utterances. They are also clearly engaged on the task of developing efficient routines for exploiting prosodic information in comprehension. It is extremely likely that these two developmental steps are related.

Other studies have also show that children's pragmatic/semantic comprehension is not fully developed in this age range. The already cited experiments by Hornby (1971) and by MacWhinney & Price (1980) found that children's comprehension of non-prosodic cues to topic-comment structure (e.g. syntactic devices such as clefting) was, like their prosodic processing, also far from approximating adult performance. Tyler & Marslen-Wilson (1981), using the word-monitoring task, found that 5-year-olds showed a specific performance deficit in comparison to older children and adults which they attributed to inefficient discourse mapping processes. Paul (1985) found that even third- and fifth-grade children showed imperfect performance on a (relatively difficult) test of the assignment of thematic roles (given versus new) in sentences.

In the development of this type of semantic/pragmatic performance, production and comprehension abilities seem to be more or less parallel. Thus there are numerous demonstrations that young children do not always mark sentence semantic structure in the way adults do. Karmiloff-Smith (1979) found that 5-year-olds frequently use ambiguous and non-specific pronouns; Bates (1974) found that children under the age of six do not mark thematic roles (given versus new) by word order changes, a finding replicated by MacWhinney & Bates (1978), who also found that pronominalization was not sensitive to thematic structure in this age group. Thus the results of the present study corroborate the general conclusion that semantic/pragmatic abilities, both productive and receptive, are still undergoing development in 4- to 6-year-olds.

Consideration of this state of affairs makes the prosodic paradox which we discussed in the introduction yet more anomalous. In general, children's semantic/pragmatic abilities follow the general rule of linguistic performance: production is at best only as good as comprehension, it never outstrips it. Only prosodic performance seems to be an exception: at least the three studies of Hornby & Hass (1970), Wieman (1976) and MacWhinney & Bates (1978) found that children produced appropriate sentence accent cues to semantic/pragmatic roles; these children were in just that age group which other research has shown to be (a) unable to produce or comprehend semantic/pragmatic structure appropriately; and (b) unable to use prosodic structure appropriately in comprehension.

The paradox, seen in this light, resolves itself to a specific question of anomalous production abilities: children under six apparently cannot process semantic/pragmatic structure (e.g. given versus new, topic versus comment)

in either production or comprehension, yet their productions show appropriate accentuation patterns – new information, focussed information, commented-upon information is accented, at least sufficiently to satisfy adult listeners. From the general body of research on the discourse processing capacities of children of this age, however, one would have to conclude that they do not have the underlying representations of discourse structure which are surely the prerequisite for producing a prosodic structure marking of thematic structure. What, then, can underlie this paradoxical achievement?

We believe that the only satisfactory account applicable to this body of research is that which can be drawn from the work of Bolinger (e.g. 1983). Across the world's languages, accent (prosodic obtrusion, usually upwards from the overall contour) is used to signal focus or semantic prominence (Bolinger 1978). This, he argues, is evidence that accentual focus is a true prosodic universal; its roots, he claims further, lie in primitive physiological mechanisms. As tension in the organism (i.e. the speaker) rises, pitch rises; as tension falls, pitch levels of speech fall. Thus the basic mechanism underlying accent is that a greater level of speaker excitation is associated with certain parts of an utterance than with others, and those parts associated with greater excitation will tend to be spoken with prosodic prominence, i.e. accented. It is natural to suppose that the most semantically central parts of an utterance (i.e. the most 'interesting' parts) will be associated with greater excitation; therefore the most semantically central words will be accented (Bolinger 1983).

This mechanism, it should be noted, requires no underlying representation of utterance semantic structure for accent to be produced. It is truly innate – 'you come to the word that is most interesting and exciting, and you go up. No linguistic intention need be involved' (Bolinger 1982:19). Thus Bolinger's theory of intonation allows a simple escape from what seemed to be an inexplicable paradox. Accenting of new and focussed information is a universal of speaker physiology; therefore it is not surprising to find children accomplishing it – as our literature review showed – as early as age three, i.e. virtually as soon as they are capable of lengthy utterances in which prosodic contrasts can be realized. However, it is not necessary to assume that their linguistic competence at this age extends to the relationship between accent and sentence semantic structure. On the contrary, the balance of the evidence on children's pragmatic processing suggests most strongly that it does not. It is only later that representations of discourse structure are developed, and the prosodic production system can incorporate accent placement routines based on thematic role assignment. Only once this has occurred can the prosodic production system approximate the adult system, in which the underlying physiological basis has become 'socialized' (Bolinger 1983), so that, for instance, inappropriate accent patterns can be deliberately produced and counterfeit interest can be signalled prosodically. Similarly, the ability

to use prosodic information in comprehension as a cue to semantic structure MUST await the development of internal representations of semantic structure.

The paradox is therefore no more. Children's prosodic productions at, say, age three to four on the one hand and age five to six on the other hand, though apparently similar, especially in respect of the accenting of new information, are in fact not similar at all but qualitatively different. At the earlier stage the accenting is essentially a physiological reflex which is not symptomatic of underlying prosodic competence. At the later stage the same accent patterns may be produced via a prosodic production system referring inter alia to discourse-level factors, on the adult model.

It is to be hoped that sufficiently subtle experimentation may in the future identify ways in which accentuation processes at these two stages can be shown to differ. Meanwhile, we claim that Bolinger's theory of the underlying basis of prosody offers the only way out of what appeared to be an extraordinary anomaly in language acquisition research.

## REFERENCES

- Allen, G. D. (1983) Linguistic experience modifies lexical stress perception. *Journal of Child Language* **10**. 535-49.
- Allen, G. D. & Hawkins, S. (1978) The development of phonological rhythm. In A. Bell and J. B. Hooper (eds), *Syllables and segments*. Amsterdam: North Holland.
- Atkinson-King, K. (1973) Children's acquisition of phonological stress contrasts. *UCLA Working Papers in Phonetics* **25**.
- Bates, E. (1974) The acquisition of pragmatic competence. *Journal of Child Language* **1**, 277-81.
- (1976) *Language and context: the acquisition of pragmatics*. New York: Academic Press.
- Bolinger, D. L. (1978) Intonation across languages. In J. H. Greenberg (ed.), *Universals of human language*. Vol. 2: *Phonology*. Stanford: Stanford University Press.
- (1982) Nondeclaratives from an intonational standpoint. In *Papers from the parasession on non-declarative sentences of the eighteenth regional meeting*. Chicago: Chicago Linguistic Society.
- (1983) Intonation and gesture. *American Speech* **58**. 156-74.
- Bosshardt, H. G. & Hörmann, H. (1982) Der Einfluss suprasegmentaler Information auf die Sprachwahrnehmung bei 4 - bis 6 jährigen Kindern. *Archiv für Psychologie* **134**. 81-104.
- Brown, R. (1973) *A first language: the early stages*. Cambridge, MA: Harvard University Press.
- Cowin, E., Mann, V., Schoenheimer, J. & Berman, J. (1984) Age-related changes in the syntactic interpretation of prosodic cues. Paper presented to the 108th Meeting of the Acoustical Society of America.
- Cruttenden, A. (1974) An experiment involving comprehension of intonation in children from 7 to 10. *Journal of Child Language* **1**. 221-31.
- (1985) Intonation comprehension in ten-year-olds. *Journal of Child Language* **12**. 643-61.
- Cutler, A. (1976) Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception and Psychophysics* **20**. 55-60.
- Cutler, A. & Fodor, J. A. (1979) Semantic focus and sentence comprehension. *Cognition* **7**. 49-59.
- Cutler, A. & Foss, D. J. (1977) On the role of sentence stress in sentence processing. *Language and Speech* **20**. 1-10.
- Foss, D. J., Bias, R. & Starkey, P. (1978) Sentence comprehension processes in the pre-schooler. In R. Campbell and P. T. Smith (eds), *Recent advances in the psychology of language*. New York: Plenum.

- Furrow, D. (1984) Young children's use of prosody. *Journal of Child Language* **11**. 203-13.
- Gleitman, L. R. & Wanner, E. (1982) Language acquisition: the state of the state of the art. In E. Wanner & L. R. Gleitman (eds), *Language acquisition: the state of the art*. Cambridge: C.U.P.
- Hornby, P. A. (1971) Surface structure and the topic-comment distinction: a developmental study. *Child Development* **42**, 1975-88.
- Hornby, P. A. & Hass, W. A. (1970) Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research* **13**. 395-9.
- Karmiloff-Smith, A. (1979) *A functional approach to child language*. Cambridge: C.U.P.
- Klein, H. B. (1984) Learning to stress: a case study. *Journal of Child Language* **11**. 375-90.
- Lahey, M. (1974) Use of prosody and syntactic markers in children's comprehension of spoken sentences. *Journal of Speech and Hearing Research* **17**. 656-68.
- 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.
- MacWhinney, B., Pleh, C. & Bates, E. L. (1985) The development of sentence interpretation in Hungarian. *Cognitive Psychology* **17**. 178-209.
- MacWhinney, B. & Price, D. (1980). The development of the comprehension of topic-comment marking. In D. Ingram, F. Peng & P. Dale (eds), *Proceedings of the first international congress for the study of child language*. Lanham MD: University Press of America.
- Mehler, J. & Bertoncini, J. (1979) Infants' perception of speech and other acoustic stimuli. In J. Morton & J. Marshall (eds), *Structures and processes (Psycholinguistics series 2)*. London: Elek.
- Miller, W. & Ervin, S. (1964) The development of grammar in child language. In U. Bellugi and R. Brown (eds), *The acquisition of language. Monographs of the Society for Research in Child Development* **29**. No. 92.
- Morse, P. A. (1972) The discrimination of speech and nonspeech stimuli in early infancy. *Journal of Experimental Child Psychology* **14**. 477-92.
- Paul, R. (1985) The emergence of pragmatic comprehension: a study of children's understanding of sentence-structure cues to given/new information. *Journal of Child Language* **12**. 161-79.
- Rudnicky, A. (1980) Structure and familiarity in the organisation of speech perception. Unpublished PhD Dissertation, Carnegie-Mellon University.
- Scholes, R. J., Tanis, D. C. & Turner, A. (1976) Syntactic and strategic aspects of the comprehension of indirect and direct object constructions by children. *Language and Speech* **19**. 212-23.
- Smith, B. L. (1978) Temporal aspects of English speech production: a developmental perspective. *Journal of Phonetics* **6**. 37-67.
- Solan, L. (1980) Contrastive stress and children's interpretation of pronouns. *Journal of Speech and Hearing Research* **23**. 688-98.
- Swinney, D. A., Zurif, E. B. & Cutler, A. (1980) Effects of sentential stress and word class upon comprehension in Broca's aphasics. *Brain and Language* **10**. 132-44.
- Tonkova-Yampol'skaya, R. (1973) Development of speech intonation in infants during the first two years of life. In C. A. Ferguson and D. I. Slobin (ed), *Studies of child language development*. New York: Holt, Rinehart & Winston.
- Tyler, L. K. & Marslen-Wilson, W. D. (1978) Some developmental aspects of sentence processing and memory. *Journal of Child Language* **5**. 113-29.
- & — (1981) Children's processing of spoken language. *Journal of Verbal Learning and Verbal Behaviour* **20**. 400-16.
- Wieman, L. (1976) Stress patterns in early child language. *Journal of Child Language* **3**. 283-6.

## APPENDIX I

Target A: Jenny

Target B: Patti

Once there were two friends called Jenny and Patti. They lived just across the road from one another in a small town in Massachusetts. One lived in a tall house with chimneys, and the other in a wide house with big windows. Which one lived in the house with <sup>chimneys</sup><sub>big windows</sub>? The house that Patti lived in had big windows, and Jenny's house had many chimneys. One house had a swimming pool, and the other had an orchard. Whose house had a <sup>swimming pool</sup><sub>orchard</sub>? The house that Jenny lived in had a swimming pool, and Patti's house had the orchard. So the two friends always had good places to play. Their fathers were also friends. One was a doctor and one was a teacher. Whose father was the <sup>teacher</sup><sub>doctor</sub>? The teacher was Patti's father, and Jenny's father was a doctor. Often the two families went on vacation together, but one year they went separately, one family to Florida and the other family to Hawaii. Whose family went to <sup>Florida</sup><sub>Hawaii</sub>? Florida was where Jenny's family went, and Patti's family went to Hawaii. Because they went to different places, the two friends sent each other lots of postcards.

[Example story used in Experiment 3. Version 1 of the story contained all upper alternative questions, version 2 contained all lower alternatives. Each subject heard only one version and listened for only one name target. The four experimental sentences which contained targets were counter-balanced such that each target was focussed twice in each version; order of occurrence in the sentence was also controlled.]

## APPENDIX II

Target A: Dick

Target B: Cathy

Version 1

This is about a brother and sister named Dick and Cathy. They lived with their parents in a great big house. The parents liked their son to help them with whatever he could, and often Dick helped CATHY too. The brother and sister were very close, and although SHE sometimes had a hard time from some of the other kids at school, it never happened that DICK was nasty to Cathy. And just the same way, although she was often mad at one or another kid at school, Cathy was never mad at DICK. In fact, whenever the parents were helping their son with his homework, CATHY would try to help Dick, too. They always enjoyed being together.

## Version 2

This is about a brother and sister named Dick and Cathy. They lived with their parents in a great big house. The parents liked to help their daughter with her homework, and often DICK helped Cathy too. The brother and sister were very close, and although HE sometimes gave some of the other kids at school a pretty hard time, it never happened that Dick was nasty to CATHY. And just the same way, although many of the other kids were often mad at him, CATHY was never mad at Dick. In fact, whenever she was helping her parents with one thing or another, Cathy would try to help DICK, too. They always enjoyed being together.

[Example 'accent' story used in Experiment 4. As in the focus stories, each target occurs in four experimental sentences, twice accented and twice unaccented in each version; order of occurrence in the sentence is also controlled.]