

2. Prosodic development

David Crystal

1. The nature of prosody

Unlike the well-established fields of grammar and segmental phonology, the field of prosody requires a certain amount of exposition, before one can talk about its development in children. What range of data is to be subsumed under this heading? For present purposes, it will be useful to select three main themes from the prosodic literature, reviewed, for example, in Crystal (1969, 1975): (a) What is excluded from the domain of a prosodic theory? (b) What variations in prosodic form may be identified? (c) What range of prosodic functions may be established?

The first question may be answered briefly, as its purpose is solely to relate prosody to other areas with which it is often confused. Given a specification of semiotic behaviour as 'patterned human communication in all its modes' (Sebeok *et al.*, 1964), then prosodic features can be identified as one component of the auditory-vocal dimension of communication, i.e. excluding the visual ('kinesic'), tactile ('proxemic') and other communicative modes referred to globally under the general heading of 'nonverbal communication' (see Argyle, 1975). Within the auditory-vocal area, a distinction is conventionally made between *segmental* and *nonsegmental* phonology, the latter usually being defined as the 'residue' of the former – what is left *after* one has studied the vowel/consonant/syllabic system of sounds. More positively, one can define nonsegmental phonology as any linguistically contrastive sound effect which cannot be described by reference to a single segment (phoneme), but which either (i) continues over a stretch of utterance (minimally, a syllable), e.g. extra loudness; or (ii) requires reference to several segments in different parts of the utterance, e.g. the use of breathy voice on vowels. Within this field, prosody is a term which traditionally has referred only to certain aspects of this variability, namely linguistic variation in pitch, loudness, speed and rhythm (including pause) of speaking. Other aspects were either ignored, or grouped loosely together under the heading of 'paralanguage' (e.g. nasal, husky, or whispered vocal effects). In the language acquisition field, the term 'prosody' has tended to

follow this orientation, though one sometimes sees it used where it plainly has a more general meaning, namely, all aspects of nonsegmental variability. Certainly, with few exceptions, the attitude adopted to prosody is very much that of the 'residue' referred to above. It is regularly assumed that segmental phonology, grammar and lexicon can be analysed without reference to prosodic features, which may, if at all, be mentioned only in passing. The methodological fallacy here is particularly worrying with reference to grammar, at the early stages of development, but its implications cover all aspects of language structure and function. There are signs, particularly in the recent literature on the interactional analysis of early language, that the importance of prosody is becoming increasingly realized; but this literature also illustrates the difficulty of working precisely within this domain, as will be discussed below.

Turning now to the second question raised above: what prosodic variability to recognize? A traditional misconception here is to talk of a child's 'prosody', implying that this is in some sense a single, homogeneous phenomenon. The linguistic literature shows very plainly, however, that several different kinds of patterning are involved in prosodic analysis, and it is important to distinguish these, in order to specify precisely what it is that is being acquired. The present chapter will be restricted to those aspects of prosodic patterning the importance of which is generally agreed among the various approaches (the terminology and particular theoretical emphasis used is that of Crystal (1975)).

(i) The primary focus must be on the way in which the several prosodic characteristics of the speech signal are integrated to produce a totality, which expounds meaning. The linguistic use of pitch, or *intonation*, ultimately develops into the most complex of all the prosodic systems, and for the adult it can to a considerable extent be studied formally as an autonomous system. But from a semantic point of view, intonation is merely one factor in communicating a meaning – as is clear when we consider what range of vocal characteristics enter into the definition of such tones of voice as sarcastic, angry, parenthetical, etc. Particularly for the first two years of life, nonintonational features (such as variations in loudness, duration, rhythmicity) are of considerable importance in the expression of meaning. This is so, not only for attitudes, but also for grammatical patterning. For example, when one tries to decide whether a sequence of two lexical items constitutes one sentence or two, one must listen out for far more than pitch contour and pause (the two features usually referred to): they may be linked by extra loudness, longer duration, marked rhythm, or with some shared paralinguistic feature (e.g. marked tension, or nasality) – all of which requires one to place intonation in its proper perspective. Intonation nonetheless has

attracted most attention in the developmental literature, and the following discussion reflects this emphasis.

(ii) A basic distinction is postulated between pitch *direction* and pitch *range*. A pitch may fall, rise, stay level, or do some combination of these things within a given phonological unit, e.g. fall-rise on a syllable; and these directional *tones* are usually isolated as one system of intonational contrastivity. But any of these tones may be varied in terms of range, seen as a separate system of contrasts, e.g. a falling pitch used relatively high, mid or low, or being widened or narrowed in some way.

(iii) Features of pitch direction and range, along with features of rhythm and pause, are organized into prosodic configurations, or *tone-units* (or 'primary contours', 'sense groups', etc.), which expound meanings over and above the accompanying 'verbal' meanings. Tone-units provide the most general level of organization that can be imposed upon prosodic data, equivalent in status to the notion of 'sentence' in grammatical analysis.¹ For example, the normal tone-unit segmentation of the utterance

When he comes/ tell him I'm out/

is as indicated by the slant lines. In general, the assignment of tone-unit boundaries seems motivated by syntactic reasons, e.g. to mark the boundary between clauses (see further, Crystal, 1975: ch. 1). We might accordingly expect such a fundamental notion to be an early characteristic of prosodic development (see below).

(iv) The prosodic feature which seems to carry the next most important linguistic contrastivity is the placement of maximum prominence on a given syllable (or, occasionally, on more than one syllable). This is primarily a matter of pitch movement, but extra loudness is involved, and duration and pause may be used to heighten the contrast between what precedes and follows. The prominent, or *tonic* syllable may be seen capitalized in the utterance

Because we stayed until MIDNIGHT/ we got in TERRIBLY late/

This is the focus of most of the discussion on intonation in the context of generative grammar (e.g. Bresnan, 1971), where the aim was to demonstrate that tonicity (i.e. tonic syllable placement) had a syntactic function, being used to disambiguate sentences or signal a distinction between grammatical and ungrammatical. While this sometimes happens, the alternative view (which I share) is that of Bolinger (1972), who argues that the factors

¹ One awaits the demonstration that tone-unit sequences operate generally and systematically (cf. paragraphs), though some interesting suggestions along these lines have already been made (e.g. Fox, 1973).

governing tonic placement are primarily semantic, e.g. the signalling of new information in context.

(v) The next most noticeable prosodic characteristic is the specific direction-range of the tonic syllable, e.g. whether the *tone* of the syllable is high-falling-wide, low-rising-narrow, etc. These tones seem to signal primarily attitudinal information, though certain tonal contrasts can expound grammatical meaning, e.g. the 'asking' versus 'telling' distinction in tag-questions:

You're còMING/ĀREN'T you/
You're còMING/ĀREN'T you/

The third question raised at the beginning of this chapter concerned the range of functions which these prosodic patterns might be said to perform. Here five roles need to be distinguished.

(i) In a *grammatical* function of prosody, the prosodic feature(s) signal a contrast, the terms of which would be conventionally recognized as morphological or syntactic in a grammar, e.g. positive/negative, singular/plural, statement/question. These contrasts are common in tone languages, such as Twi, but some of these may also be found in English, where tone-units, tonic syllables and tones can all perform grammatical roles (cf. the tag-question contrast above). In a related sense, prosody may be used obligatorily to mark a grammatical distinction already overt in word order or morphology, as in co-ordinated utterances such as

I'll ask the FİRST question/ and you ask the SÈCOND one/

(ii) The *semantic* function of prosody subsumes a speaker's organization of meaning in a discourse, whereby he signals which parts of what he is saying are most important, which parenthetic, etc. This includes the emphasizing of a relatively unexpected lexical item in an utterance, as Bolinger argues (1972), as well as reflecting the presuppositions about subject matter or context when focusing on a specific item, as when, for example *There was a bus in the road* implies a context where someone had queried whether this was in fact the case.

(iii) The *attitudinal* function of prosody is usually distinguished from (ii), on similar grounds to the classical distinction between denotation and connotation. Personal emotions are signalled concerning the subject matter or context of an utterance, e.g. anger, puzzlement, surprise. It is unclear how far such emotions use prosodic features specific to a language, and how far they rely on universal characteristics of emotional expression (cf. Bolinger, 1964).

(iv) The *psychological* function of prosody is evident from the several

experiments which have shown that performance in short-term memory, recall, perception and other variables is affected by the prosodic character of the utterance, e.g. words containing tonic syllables are more readily recalled (cf. Leonard, 1973).

(v) The *social* function of prosody signals information about the sociolinguistic characteristics of the speaker, such as his sex, class, professional status, etc. (Crystal, 1975: ch. 5). The importance of this function in facilitating social interaction in dialogue is being increasingly recognized, e.g. when the intonation of a stimulus sentence prompts someone to respond, or implies that no further comment is needed. The role of prosody in expressing the illocutionary force of a speech act, such as persuading or commanding, is also now seen to be significant – though whether one might refer to this as primarily a social, attitudinal, semantic or grammatical role is very much an open question!

This last point raises a principle of fundamental importance, which needs to be emphasized, in view of the extent to which it is neglected in child language studies – namely that there is no one-to-one correspondence between the above categories of prosodic form and prosodic function, nor between any of the individual features subsumed within these categories. A rising tone, for example, signals far more than a questioning meaning, and a grammatical question may be uttered using other tones than rising ones (for discussion and references, see Crystal, 1969: ch. 1). It is accordingly fallacious to assume that a child who uses rising tones is thereby ‘asking a question’, ‘making a questioning speech act’, or the like: everything depends on the careful analysis of the accompanying behaviour and situation before one can be justified in ascribing such an interpretation to the utterance.

2. Prosodic acquisition

Given the limited empirical study which has taken place (almost entirely within the first two years of life), talk in terms of clear stages of development in this area may well be premature. On the other hand, the evidence which is available does agree so far on several points, hence the following progression. Five stages can be distinguished, of which the last two are particularly important.

Stage I. There have been many studies of the prelinguistic antecedents of prosodic features, usually under the heading of ‘infant vocalization’. On the whole, these studies recognize a period of biologically determined vocalizations (e.g. the ‘basic cry’ pattern, underlying hunger, pain, etc. states described in Wolff (1969: p. 82)), and a period of differentiated vocalizations which permit general attitudinal interpretation only (e.g. ‘pleasure’,

'recognition'). Systematic variation in these vocalizations can be ascribed to such factors as the baby's sex or environment. There seems to be little difference in their physical characteristics and attitudinal function across languages. This stage, from birth until around 6 months, is reviewed in detail in Crystal (1975: ch. 8).

Stage II. The first sign of anything linguistic emerging is the awareness of prosodic contrasts in adult utterances directed to the child. This has long been known to be present in children from around 2 to 3 months, as the reports in Lewis (1951) testify. But this literature is rather anecdotal, and experimental studies are lacking which attempt to separate prosody from other semiotic features of the stimuli, and to identify the roles individual prosodic features might play within the adult utterance, e.g. whether pitch or loudness are discriminated first. Kaplan (1970), for example, demonstrated that a contrast between falling and rising tones could be discriminated from around 4 months, but it is difficult to be sure of the relative roles pitch and loudness had to play.

Stage III. The increasingly varied vocalizations of children around 6 months have begun to be studied in detail, using a combination of acoustic, articulatory and auditory criteria, and it is possible to isolate a wide range of nonsegmental parameters in terms of which the patterns of crying, babbling, etc., can be classified. Stark *et al.* (1975), for example, cite breath direction, pitch, loudness, and several kinds of glottal and supraglottal constriction; within pitch, they distinguish contrasts of range, direction and continuity.

Gradually, these nonsegmental features come to resemble prosodic patterns of the mother tongue – from as early as 6 months, according to most scholars (see the review in Crystal, 1975: p. 136). Initially, the resemblance is only hinted at, by the occasional use of a language-specific prosodic characteristic within a relatively long stretch of nonlinguistic vocalization. Such instances are very striking when they occur on a tape, as they stand out as something much more familiar, discrete and transcribable than the general background of utterance. Increasingly, at this time, babbling patterns become shorter and phonetically more stable: accordingly, when a babbled utterance of only one or two syllables is used in conjunction with a language-specific prosodic feature, the result is going to be very much like an attempt at a meaningful utterance. Such combinations are quickly focused on by parents, who will comment on what they think the baby is 'saying', often providing lexical glosses. It is, however, very difficult to be precise about the nature of the development at this stage. To say that a language-specific feature has been detected is to say very little: recognition of language-specificity involves both phonetic notions (e.g. the 'community voice quality' or characteristic 'twang' of a language) and phonological notions (e.g. the selection

of contrasts which produce an identifiable accent), and it is by no means clear how to distinguish these in the child's vocalization at this stage. The boundary area between the phonetic use of pitch, loudness, etc., during the first 6 months of life, and the phonological use of pitch that has emerged by around a year, is totally uncharted territory (cf. Olney and Scholnick, 1976).

Stage IV. However the transition to phonology takes place, it is evident that learned patterns of prosodic behaviour are characteristic of the output of the child during the second half of the first year. These patterns can be studied both formally and functionally. From the formal viewpoint, the increasingly determinate and systematic character of these patterns is readily statable: a configuration of features is involved, using primarily pitch, rhythm and pause. This configuration has been variously labelled: a prosodic 'envelope' or 'matrix' (Bruner, 1975: p. 10) or 'frame' (Dore, 1975). Weir (1962) had previously talked about the splitting up of utterances into 'sentence-like chunks', at this stage. Lenneberg (1967: p. 279) describes the process thus:

'The first feature of natural language to be discernible in a child's babbling is contour of intonation. Short sound sequences are produced that may have neither any determinable meaning nor definable phoneme structure, but they can be proffered with recognizable intonation such as occurs in questions, exclamations or affirmations. The linguistic development of utterance does not seem to begin with a composition of individual, independently movable items *but as a whole tonal pattern*. With further development, this whole becomes differentiated into component parts . . .' (my emphasis)

The important point is that these primitive units have both a segmental and a prosodic dimension, but it is the latter which is the more stable, and the more readily elicited. In one child studied at Reading, aged 1;2, the phrase *all-gone*, regularly said by the parent after each meal, was actually rehearsed by using the prosodic component only: the child hummed the intonation of the phrase first, viz.  , only then attempting the whole, producing an accurate intonation but only approximate segments ([ʌ?dʌ]). The phrase could be easily elicited after any meal, but it was not until a month had gone by that the child's segmental output became as stable as his prosodic. Menn's Jacob (1976a: pp. 195ff.) also produced 'proto-words' with a distinctive prosodic shape – the ones reported being used at 1;4 for a peekaboo game, an item with demonstrative function (*ésa*), and, later, a name-elicitor (*zi*). Dore (1975) refers to the formally isolable, repeated, and situationally specific patterns observed at this stage as 'phonetically consis-

tent forms', whose 'protophonemic' segmental character is complemented by a distinctive prosody, which is the more stable.

From a functional point of view, these prosodically delimited units can be interpreted in several ways – semantic, syntactic and social 'explanations' have all been mooted. The latter view is perhaps the most widely held: here, prosody is seen as a means of signalling joint participation in an action sequence shared by parent and child. This view, emphasized particularly by Bruner (1975), is part of a developmental theory wherein vocalization is seen as one component in a communication activity alongside such non-vocal behaviour as reaching and eye-contact. In a peekaboo game, for instance, both the utterance and the activity of hiding-and-reappearing are obligatory, interdependent components (as the absurdity of attempting to play the game without either indicates). And when adults play these games, the lexical character of the utterance regularly varies ('Peep-bo', 'See you', etc.) whereas the prosodic features display much less variation. Another example is in action sequences such as nuzzling the child or jumping him up and down, where there are parallel prosodic patterns. The development of 'turn-taking', either between parent and child (Snow, 1977a) or between children (Keenan, 1974) also involves prosodic delimitation and interdependence. One Keenan twin, for example, would regularly take the prosodic character of the other's utterance and 'play' with it. Another child, studied at Reading, marked the end of a jargon sequence with a distinctive two-syllable pitch movement (· ·), which was openly described by his parents as 'their cue to speak'.

Several attempts have been made to describe the social or 'pragmatic' functions of such utterances, especially using the metalanguage of speech act analysis. Dore, for example (1975: pp. 31ff.), argues that prosodic features provide crucial evidence for the development of speech acts. Primitive speech acts are said to contain a 'rudimentary referring expression' (lexical items) and a 'primitive force indicating device' ('typically an intonation pattern', p. 31), as in labelling, requesting and calling. The distinction between referent and intention is pivotal: 'whereas the child's one word communicates the notion he has in mind, his prosodic pattern indicates his intention with regard to that notion' (p. 32). Likewise, Menn says about Jacob (1976a: pp. 26–7), 'he . . . consistently used certain intonation patterns in conjunction with actions that communicated particular intentions, so we can ascribe meaning to his use of those contours'. The difficulty with all such approaches, of course, is empirical verification of the notion of 'intention'. As has been argued in other areas of child language, the fact that parents interpret their children's prosody systematically is no evidence for ascribing their belief patterns to the child's intuition. At best, one can argue,

as does Menn (1976a: p. 192) that 'consideration of adult interpretation of intonation contour on vocalizations does give us information about what the child *conveys*, if not what he/she *intends*'. It is difficult to go beyond this, and know that a child at this stage intends a distinction between, say, 'calling' and 'greeting' (two of Dore's categories). Searching for one-to-one correlations between prosody and other aspects of the child's behaviour is unlikely to be successful, because the situations in which the language is used are often indeterminate, and the gestures and other kinesic features accompanying are usually ambiguous. There are also fewer pitch patterns available in a language than there are situations to be differentiated. It is possible that more detailed behavioural analyses will give grounds for optimism, but these ought to begin with the most concrete, determinate and replicable of situations (e.g. the daily, ritualized settings referred to in Bruner (1975) and R. Clark (1977)). Attempting to establish developmental speech act theories using as sole data a sample of unstructured, spontaneous play interaction is liable to produce a set of unfalsifiable interpretations about what went on. The intuitive plausibility of interactional approach to the study of utterances at this stage is thereby much reduced.

As an alternative to a social approach, it is possible to see these prosodic frames – or primitive tone-units, to use the terminology above – primarily as having a formal or grammatical role. Bruner, for instance (1975), at one point describes the function of these frames as 'place-holders': a mode of communication (such as a demand, or a question) is established using prosody, and primitive lexical items are then added. In a stretch of jargon, from around 12 months, it is often the case that one will recognize a word within the otherwise unintelligible utterance (cf. also several of the utterances in Keenan (1974)). And the transitional stage between one- and two-element sentences also contains uninterpretable phonetic forms which may perhaps be interpreted as remnants of a primitive prosodic frame (cf. Dore *et al.*, 1976; Bloom, 1973). Dore *et al.* (1976: p. 26) in fact suggest seven transitional stages at this point:

- (i) prosodically un-isolable, nonphonemic units ('prelinguistic babbling')
- (ii) prosodically isolable, nonphonemic patterns ('prelinguistic jargon')
- (iii) prosodically isolable, nonphonemic units ('phonetically consistent forms')
- (iv) conventional phonemic units ('words')
- (v) word plus 'empty' phonetic forms in single prosodic pattern ('presyntactic devices')
- (vi) chained conventional phonemic units forming separate intonation patterns ('successive single-word utterances')
- (vii) prosodically complex patterns ('patterned speech')

The phonological and phonetic details of the development of these frames into determinate tone-units with a definable internal structure are, however, not at all clear, so little empirical work having been done. In particular, it is unclear whether tonicity or tonal contrastivity develops first, or whether they emerge simultaneously. The suggestion that the development is simultaneous is based on the observation that tonicity contrasts are early evidenced in jargon sequences (in which sequences of rhythms are built up which resemble the intonational norms of connected speech), whereas tone contrasts are early heard in the use of lexical items as single-word sentences. Menn, for instance, finds her child's semantic control of certain tones on 'babble carriers' and their contrastive use on words to be almost simultaneous (1976a: p. 186). If one ignores jargon, however, as being both less central to communicative development and less systematic in its patterning, then it would seem that tone develops before tonicity. Polysyllabic lexical items at this stage tend to have fixed tonic placement (Atkinson-King, 1973), though they may vary in terms of pitch direction and range, e.g. *dàda* (said as daddy enters the room), *dáda* (said when a noise was heard outside at the time when daddy was expected). Of the two, range seems to become contrastive before direction, especially high versus low, but also wide versus narrow. Most of the contrasts noted by Halliday (1975), for example, involve range rather than direction – mid versus low first, later high – e.g. the distinction between seeking and finding a person, signalled in his child by high versus mid-low range, from around 1;3. Eight pitch range variations are in fact used by Halliday in his transcription (very high, high, mid high, mid, mid low, low, wide, narrow), as well as four directions (level, fall, rise, rise-fall). The notion of high versus low register is discussed further in Konopczynski (1975), and a great deal of early child data can be interpreted in this way, e.g. in Keenan (1974) and Menn (1976a).

Based on Menn (1976a), Halliday (1975), and my own study, a tentative analysis of early tonal development – the contrasts involving both direction and range – emerged as follows:

(i) Initially, the child uses only *falling* patterns. Menn states that – except for imitations of adult rises – her child used rises on words only after these words were first used with falls (1976a: p. 195). Halliday's range contrasts are all on falling tones (1975: p. 148).

(ii) The first contrast is *falling* versus *level* tones (high level in Halliday (1975: pp. 150–1)), the level tone often being accompanied by other prosodic features, e.g. falsetto, length, loudness variations.

(iii) This is followed by *falling* versus *high rising* tones, the latter being used in a variety of contexts. Menn's special study of rising tones brought to light a large number of contexts between 1;1 and 1;4, including offering,

requesting, attention-getting, and several 'curiosity' noises (e.g. when peering). Several of these notions, moreover, are complex—e.g. 'request' includes requests for help, recognition, permission, to obtain an object, etc., all of which are distinguishable in the situation (1976a: pp. 186ff., 198–9). The 'natural' distinction between fall and rise is characterized as 'demanding' versus 'requesting/offering' (p. 193). Halliday's high rises are first used in association with falls, as compound tones (1975: p. 151).

(iv) The next contrast is between *falling* and *high falling* tones, the latter especially in contexts of surprise, recognition, insistence, greetings. Halliday reports a high falling contrast between 1;1 and 1;3, and further distinguishes a mid fall.

(v) A contrast between *rising* and *high rising* tones follows: the Reading study suggested a particular incidence of high rises especially in playful, anticipatory contexts. Menn notes the latter mainly in 'intensification' contexts: the child gets no response to an utterance with a low rise, and repeats the utterance with a wider contour—the extra height, according to Menn, is the 'essential information-carrying feature' (1976a: pp. 193–4). Halliday's mid versus high rise emerges at 1;3 to 1;4.

(vi) The next contrast is between *falling* and *high rising-falling* tones, the latter being used in emphatic contexts, e.g. of achievement (e.g. *thèrē*, as an extra brick is placed on a pile) or impressiveness (e.g. *bùs* vs. *bûs*, the former being used by one child studied to refer to 'any' vehicle, the latter to a *real* bus). Menn reports a mid-high-low contour at 1;4; Halliday has a similar contrast from as early as 1;1, but regularly from 1;3.

(vii) Next appears a contrast between *rising* and *falling-rising* tones, the latter especially in warning contexts, presumably reflecting the *be careful* pattern common in adults; cf. Halliday (1975: p. 154), between 1;4 and 1;6.

(viii) Among later contrasts to appear is that between high and low rising-falling tones, especially in play contexts.

These features appear on isolated lexical items to begin with, and for a while cannot be distinguished from a prosodic idiom (i.e. an invariant prosodic pattern accompanying a fixed lexicogrammatical utterance, as in a nursery-rhyme line). Only later, when the same lexical item is used with different prosodic characteristics, can we talk with confidence about the patterns being systemic and productive. At this point, too, the tones come to be used in juxtaposition, producing the 'contrastive syntagmas' and prosodic 'substitution games' reported by Weir (1962), Carlson and Anisfeld (1969: p. 118), Keenan (1974: pp. 172, 178) and others.

We may compare, at this point, the conclusions of some recent work on the acquisition of tone languages: Hyman and Schuh (1974), Li and Thompson (1977) on Mandarin, and Tse (1978) on Cantonese. These studies agree

that the tone system is learned in advance of the segmental, and that the period between 1;0 and 1;6 is especially important. There is also agreement that the earliest tonal contrast is one of range: high level versus low in the case of the Cantonese child, high level versus falling in the case of Mandarin. After this, rising and compound tones appear, with low rises appearing to be more difficult than high. Finally, at a much later stage, tone sandhi rules are acquired. In short, despite the very different linguistic status of lexical tone and intonation, there is a close parallel in the acquisition process. Some theoretical explanation for this, in terms of perceptual and production constraints, is provided in the above references.

Particularly with the intonational studies, one must remember that the situational interpretations used cannot be taken at face value. In much the same way as has been argued for syntax and segmental phonology (Howe, 1976; Lenneberg, 1976), it is necessary to free the mind from the constraints of adult language analyses, where situational notions such as 'question', 'request', 'permission', etc., are normal. As already agreed, it is insufficient to show that adults can differentiate these patterns and give them consistent interpretations, as several studies have succeeded in doing (e.g. Menyuk and Bernholtz, 1969): as Bloom points out (1973: p. 19), this is no evidence of contrastivity for the children. Detailed analysis of both the phonetic form and the accompanying context of utterance, moreover, readily brings to light instances of contrastivity which have no counterpart in the adult language. Halliday's child, for example, for a while used rising tones for all 'pragmatic' utterances (those requiring a response, in his terms), and falling tones for all 'mathetic' utterances (those not requiring a response (1975: pp. 29, 52)). Menn's child between 1;0 and 1;8 used a class of nonadult rising tones, e.g. between 1;1 and 1;3 he used a low rising tone (peak 450 Hz) to 'institute or maintain social interaction' (the 'adult-as-social-partner' function) and a high rising tone (peak 550 Hz) for 'instrumental use of the adult' ('obtaining an object or service') (1976a: p. 184). In the case of a child studied at Reading, the falling-rising tone was initially used only in smiling-face contexts, with a generally 'playful' meaning, and never to express doubt or opposition, as it frequently does, with frowning or neutral face, in adults.

These are all examples of a relatively familiar form conveying an unfamiliar function. The converse also applies. Throughout this stage of development, the range of phonetic exponents of the prosodic frame increases markedly, to include contrasts in loudness, duration, muscular tension and rhythmicality, not all of which are used in the adult language. At around 1 year, contrasts have been noted between loud and soft, drawled and short, tense and lax, and rhythmic and arhythmic utterances. Halliday (1975) noted, in addition to the pitch direction and range contrasts already

described, several other prosodic and paralinguistic features: slow, long, short, loud; sung, squeak, frictional and glottalized. Contrastivity involving two or more prosodic parameters emerges, e.g. the use of a low, tense, soft, husky, spasmodic voice (a 'dirty snigger'). Carlson and Anisfeld (1969) distinguish loud and soft, and staccato and drawled articulations, amongst others. Other examples are the use of marked labialization, falsetto voice for whole utterances, and spasmodic articulations (lip trills, 'raspberries', etc.). It is regrettable that a more comprehensive phonetic description of this stage of development does not exist.

Stage V. Tonic contrastivity (or 'contrastive stress', as this area is often, misleadingly called) appears as sentences get more complex syntagmatically, with the appearance of two-word utterances at around 1;6 (Bloom, 1973; Clark *et al.*, 1974: p. 49).² The general developmental process seems clear. Lexical items which have appeared independently as single-element utterances, marked thus by pitch and pause, are brought into relationship (whether syntactic, semantic or collocational need not concern this chapter). At first, the lexical items retain their prosodic autonomy, with the pause between them becoming reduced, e.g. *daddy/ . garden/*. Often, long sequences of these items appear, especially repetitively, e.g. *dàddy/gàrdén/séé/dáddy/dáddy/gàrdén/dàddy/gàrdén/séé/*. (Such sequences of course defy analysis in terms of the usually cited grammatical–semantic relations.) The next step is the prosodic integration of sequences of items, usually two, into a single tone-unit. How general a process this is, is unclear, but in several English combinations studied, it was the case that one item became more prominent than the other; it was louder, and had an identifiable pitch movement. There was a rhythmic relationship between the items (anticipating isochrony), and intervening pauses became less likely in repeated versions of lexical sequences. This step is considered to be of central theoretical importance, because it is claimed to be the main means employed by the child for formally expressing grammatical–semantic relations within a sentence – 'the simple concatenation under one utterance contour of the words which interact to create a compositional meaning that is different from the meanings of the two words in sequence' (Brown, 1973: p. 182). Unfortunately, the process of concatenation is not so 'simple' as Brown suggests. All the following sequences have been observed (. = short pause, — = pauses of increasing length).

DÀDDY/ — ÈAT/	DÁDDY/ — ÈAT/
DÀDDY/ — ÈAT/	DÁDDY/ — ÈAT/
DÀDDY/ . ÈAT/	DÁDDY/ . ÈAT/

² A single-word polysyllable in principle allows for a contrast, e.g. *dàddy* versus *daddì*. There is no evidence of such forms at this stage (Atkinson-King, 1973).

DÀDDY/ÈAT/	DÁDDY/ÈAT/
daddy ÈAT	DÁDDY eat/
DÀDDY eat	DÁDDY ÈAT/ ³

It is accordingly often difficult to decide whether we are dealing with one sentence or two – especially if the context is unclear, e.g. the child is looking at a picture. In the above example, the subject–verb relation, so ‘obvious’ to the adult observer, may motivate one set of decisions. However, in the following examples (each of which may be found with any of the above twelve patterns), the ‘compositional meanings’ are by no means so clear:

DÁDDY/CÀR/ (child is looking at daddy in a car)

DÁDDY/MÙMMY/ (child is looking at a photograph of both)

DÁDDY/NÒ/ (daddy has left the room)

DÁDDY/DÀDDY (said while being held by daddy)

Prosody, it seems, cannot be used by the analyst as a primitive discovery procedure for semantics or grammar – just as it cannot be in the adult language. It is one factor, and only one, in the simultaneity of language, behaviour, situation and adult interpretation which constitutes our analytic datum. In certain settings, prosody will be a primary determinant of meaning; in other settings, it will be discounted. The way in which these factors operate upon each other in these various settings is however by no means clear (see further, Eilers, 1975).

However it is arrived at, it is plain that around 1;6 in most children, two-element sentences within single prosodic contours are used, and tonic prominence is not random. In the adult language, the prominence in a sentence consisting of one tone-unit is in 90 per cent of cases on the last lexical item (Chomsky and Halle, 1968: pp. 17ff.; Crystal, 1969, 1975: ch. 1). Bringing the prominence forward within the tone-unit is possible, for both grammatical and semantic reasons. In the former case, one may be constrained by rules of cross-reference within the sentence (e.g. *Jack saw JÍM/ and HÉ said . . .*); in the latter case, one may be making a (referential or personal) contrast between lexical items (e.g. *the RÈD dress/ not the GRÉEN dress/*). The presuppositions and attitudes of the speaker also promote marked tonicity (e.g. *I WÀNT a red dress/, he ISN'T coming/*). At the two- to three-word stage in children, there will obviously be little to note in relation to the prosodic marking of grammatical or lexical relations – such contrasts are likely to be more apparent when clause sequences appear. The most fruitful way of analysing variations in tonicity at the two-element stage, as a result, is therefore to establish a relationship between the changes in the

³ A compound-tone unit, such as this, is in fact singled out by Du Preez (1974) as an important transitional stage.

child's environment and his prosody. Wieman (1976), for example, attempts to show that new information in a sentence affects tonic placement, whereas old information does not (cf. also Gruber, 1967). If a child is given a marble, he might say /got MÀRBLE/, but thereafter he is likely to say /SÈE marble/, because *marble* is old information the second time. It is not difficult to hear examples which do not confirm this hypothesis, however; and plainly, there are difficulties in working with notions of 'information', 'presupposition' and the like with young children. How does one know that what is 'new' to the observer, interpreting the situation in terms of adult expectancies, is also going to be new to the child? How does one establish the emergence of personal, attitudinal contrastivity, equivalent to the emphasis an adult might give, quite out of the blue, to the following sentence: *I LÌKE Bartok* (I've decided!) (Where it is not necessarily the case that this had previously been in doubt - 'I didn't know you DÌDN'T like him', one might respond).

3. Future research

Once grammatical patterns and lexical sets develop, then the tracing of prosodic patterns becomes a much more straightforward task. What is important here is for researchers to remember the important role prosody has in the adult language in relation to the delimitation and integration of such structures as relative clauses, co-ordination, adverbial positioning, direct/indirect-object marking, etc. (cf. Quirk *et al.*, 1972). Very little research seems to have been done on the later development of such patterns, but it is probable that this kind of learning continues until puberty (and, in terms of the development of one's stylistic control over prosody, e.g. in dramatic speaking, into adult life). Four research tasks suggest themselves:

- (i) We need to follow the development of specific prosodic features in given grammatical lexical structures; questions have been quite well studied, but little else (cf. Menyuk, 1969; Wode, 1975).
- (ii) Analogous studies are needed of the comprehension of these features. See, for example, Cruttenden's (1974) study of certain aspects of intonation in a restricted class of co-ordinated utterances (football results), which showed that awareness of the rules involved was in the process of development between 7 and 10 years; also the developing awareness of co-referential pronouns in certain co-ordination contexts, as in Chomsky (1969), Maratsos (1973).
- (iii) Several sociolinguistic studies have now been made which show the importance of prosody in the variety of language adults use in talking to children (e.g. Blount and Padgug, 1977; Ferguson, 1977); there has however been much less discovered about whether children's prosody shows

comparable 'exaggerated' characteristics, though this is a common impression. This is certainly the case, according to Weeks (1971), who reports the emergence of a wide range of nonsegmental characteristics with a variety of functions, in children between 1 and 5 years: exaggerated intonation, loudness, high pitch and 'clarification' (slower, carefully enunciated speech) characteristics were of particular importance. Likewise Sachs and Devin (1976) report the use of higher pitch and wider intonation patterns when 3 to 5 year old children talk to a baby or doll, or role-play a baby, but the relevant information here is introduced informally. In general, it is unclear how widespread a phenomenon this is.

(iv) Lastly, it is particularly important to focus on the specific role prosodic variables (especially tonicity) play in psycholinguistic experimentation, e.g. in experiments involving recall, paraphrase, comprehension, and imitation. Varying the prosodic input does influence response patterns, as has been shown both for normal children (e.g. Du Preez, 1974) and adults (e.g. Cutler, 1976), and in the context of disability, for example by Goodglass *et al.* (1967) and Stark *et al.* (1967). Du Preez, for instance, shows very plainly that children have a predilection to imitate tonic syllables: words occurring finally in a tone-unit are imitated first, and those earlier in the tone-unit are imitated only when they are given marked tonicity, the tonic apparently 'act [ing] as a signal to notice what to follow' (1974: p. 71). On the other hand, the common strategy of 'leaving the prosody out', by presenting a series of stimuli in identical tones of voice, involves assumptions which themselves require investigation (such homogeneous sequences are abnormal in parent-child interaction; how far might such unfamiliar stimuli affect responses?). Here, as in the other research areas listed, the problems are urgently in need of systematic investigation.