

Intonation and Linguistic Theory

David Crystal

University of Reading

Background

My starting point is the assertion that prosodic features in general, and intonation in particular, have until very recently never held a prominent place in discussions of linguistic theory. Two simple reasons account for this: on the one hand, intonation analysts have on the whole not involved themselves with general questions of theory; and on the other hand, theoretical models have on the whole been orientated towards other things than prosody. Neither of these points is intended to be derogatory: they simply reflect the alternative preoccupations of earlier generations of scholars. Within the European tradition of intonation studies, for example, the priorities were dictated by demands for materials capable of being used in the context of foreign language teaching. There was an over-concentration on points of phonetic detail, and on the means of intonational transcription, and negligible attention paid to the criteria for establishing phonological categories. There were numerous ad hoc comments about the grammatical function of particular intonation patterns, but no systematic attempt to investigate the relationship between intonation and other aspects of language structure, especially syntax. The range of data which was taken into account was extremely restricted—usually the intonation of written language read aloud—and judgements about frequency of use or semantic interpretation were largely impressionistic. The pedagogical orientation for the work was not a compatible framework for theoretical enquiry, and the methodological difficulties involved in the collection and ana-

lysis of intonational data made systematic investigations impracticable, until the development of appropriate instrumentation. Whenever theoretical claims were made, therefore, they tended to be vague or oversimple — for instance, that intonation was an affective phenomenon, or that for every tone there was a basic meaning. The European tradition produced many descriptive insights into intonation, but whatever theoretical position underlied the work of such scholars as Passy, Palmer, Jones and Armstrong, it remained largely implicit. The American tradition of studies before Chomsky was quite different in its general orientation, but again, what was produced was primarily a descriptive framework and a set of methodological directives, with little by way of an explicit theory within which prosodic observations could be interrelated and predicted.

These points have been made before,¹ and bear witness to the general neglect of theoretical considerations by students of intonation. They can be paralleled by the neglect, amounting at times to total ignorance, of prosodic features by linguists working within the 'mainstream' of their discipline. Segmental phonology, morphology and syntax have each received a respectable measure of attention in post-Saussurean linguistics, but despite the pleas of Pike, Trager, Bolinger and others, the dominant attitude in the fifties was still to see intonation as 'on the edge of language', referring to it in descriptive or theoretical enquiry only when absolutely necessary, or reducing parts of it to units of the same kind as segmental phonemes.² Early work in generative grammar generally maintained the structuralist emphasis, either by ignoring intonation altogether, or by dismissing it as mere performance. And even more recently, there has been a reluctance to deal with intonation at all other than to see it as an extension of word-level stress phonology

¹A general discussion, with full bibliography, is to be found in Crystal 1969:Ch. 2 and p.253ff.

²Early criticism of segmental reductionism came from Haas (1957:159), Bazell (1954:133) and Bolinger (1949, 1951, 1958). The arguments have nowhere been answered.

(as in 'The Sound Pattern of English', discussed further below). It is this reluctance which current trends in linguistics are beginning to erode.

In the past ten years, views of intonation have radically altered. The intonation analysts have been affected by the general stimulus of the Chomskyan approach to search for general explanations of particular events, and the theoreticians have been impressed by a confluence of ideas about prosodic features from a number of distinct branches of the subject which suggests that such general explanations are both possible and relevant. Prosodic features seem nowadays to hold an importance for the explanation of linguistic behavior not previously realised. This was first recognised in the various educationally-, anthropologically- and sociologically-orientated studies in the mid-sixties, out of which developed, inter alia, concepts of communicative competence and the ethnography of communication (see Gumperz & Hymes 1972). The ethnographic approach was in principle concerned with the microscopic analysis of the data of social interaction, to establish what features in the act of speaking had the function of signalling the various social assumptions, relationships and categories found in different types of situation. Early on it became clear that conventional analysis in terms of the segmental phonology, syntax and vocabulary of sentences was leaving out a great deal of significance, and a closer attention came to be paid to prosodic phenomena. A recent paper on classroom interaction shows this emphasis, suggesting 'that the prosodic component encompassing stress, pitch and timing along with speech features usually termed paralinguistic is . . . part of an optional set of communicative strategies that can be used alternatively with syntactic, lexical or phonological variables' (Gumperz & Herasimchuk 1972:19-20). The key role of non-verbal factors (under which term is subsumed both linguistic features such as intonation and non-linguistic features such as voice-quality) is a main theme of Robinson (1972). This book takes the view that 'patterns of stress and intonation (and other non-segmental

patterns, as the author makes clear elsewhere, DC) are best treated as essential rather than peripheral features of the language' (187), and in his review of social psychological research into linguistic behaviour, it is remarkable how many quite distinct lines of research conclude that the understanding of these patterns is a key to the problem being investigated: not only the marking of emotional states, but the marking of personality, social identity, role relationships, social class, the regulation of encounters, and much more. For example, an investigation by Seligman, Tucker & Lambert (1970) indicates that the prosodic characteristics of children's voices unconsciously affect teachers' evaluations of their intelligence and capabilities. Crystal (1972) goes so far as to hypothesise that 'the distinctiveness of a spoken variety of language lies primarily in its use of prosodic and paralinguistic features' (3). The social and psychological significance of non-segmental features is also accepted and illustrated, along with visual (kinesic) and other communicative modalities, in recent work on semiotics; for example the identifying function of prosody in schizophrenia is argued by Ostwald in Sebeok, Hayes & Bateson (1964), and there is a general review of the significance of paralanguage in the paper by Mahl & Schulze in that volume. More recently, the whole field of non-verbal communication has been studied in a Royal Society study group (see Hinde 1972), and due attention to the importance of prosody and paralanguage is paid there by Lyons, Argyle, and others.

The second trend towards the recognition of the centrality of intonation and related features is in language acquisition studies. There are now a number of reports which show that prosodic contrasts are used as part of a child's production long before the development of segmental phonological contrasts, and thus before identifiable 'first words'. Prosodic features are used grammatically, to identify sentence-boundaries and sentence functions ('wanting', 'identifying', etc.), and this process appears productively from around 6-7 months, and receptively much earlier (for a review of research, especially on production,

see Crystal 1973; for a discussion largely on reception, see Kaplan 1970). The specification of early semantic categories in terms of the prosodic and paralinguistic features which expound them is the direction in which the study of early language development is tending to move, and this of course has been reinforced by current developments in linguistic theory and description.

Recent developments here have given increasing recognition for intonation. We may begin with the most recent reference grammar for English (Quirk et al 1972), which is the first to give a separate section to intonation, and which illustrates throughout the role of prosodic factors in distinguishing grammatical patterns within sentences and in the connectivity of sentences and sentence parts. A number of their descriptive points will be referred to below. It is important to note that this grammar was based upon a particularly large sample of data, within which the area of spontaneous conversation was given considerable coverage. Their observations about the role of intonation are thus a major development from the statements of the pedagogical studies referred to above. A second, and more far-reaching development was the recent attempt of some generative grammarians to take account of these phenomena within the frame of reference of the standard model of generative grammar. This debate, largely published in *Language* 47 and 48, was of value in that it set a seal of approval on intonational studies, and brought them more within the purview of linguistic theory, as then understood; but it can be argued that it failed to provide an adequate account of intonation, in that it took account of but a very restricted range of data, and ignored the way in which the assumptions of an 'Aspects'-type model of generative grammar pre-set the direction of the argument, and made it impossible to give appropriate discussion to certain significant questions.

The Generative Debate

The starting point of the generative discussion was the Nuclear Stress Rule (NSR) as formulated by Chomsky & Halle

(1968:17 ff.), whereby primary stress is assigned to the rightmost primary-stressed vowel within a major constituent (NP, VP, S), weakening all others. It is a cyclic rule which applies after all the rules which determine the stress of individual lexical items, and results in the traditionally 'neutral' intonation for an English sentence, the last constituent being stressed. This rule is assumed to apply to surface structures, operating as part of the phonological cycle and applying to the various constituent bracketings, starting with the most deeply embedded constituent. Chomsky & Halle themselves recognised that there are classes of exceptions to this rule, referring to an example they found in Newman (1946)—though it had worried Sweet and Palmer long before—wherein the ambiguity of *He has plans to leave* (= he intends to leave/he has documents to leave) is claimed to be resolvable by contrastive tonic placement, and to an example of contrastive stress in cases of syntactic parallelism (e.g. *He wanted to study electrical rather than civil engineering*).³

Bresnan (1971) took up the question of these types of exception. She referred to other well-known examples of cases where the final constituent is unstressed, namely sentences with final anaphoric pronouns (e.g. *Helen teaches it*), final indefinite pronouns (*The boy brought some*), and repeated items (*John knows a woman who excels at karate, and he avoids the woman*), and then concentrated on four types of syntactically complex constructions which do not conform to the NSR: (a) *George has plans to leave*; (b) *Mary liked the pro'posal that George left* (vs. . . . *proposal that George 'leave*); (c) *John asked what Helen had 'written* (vs. *John asked what 'books Helen had written*); (d) *George found someone he'd like you to 'meet*

³What does not seem to have been noticed is that tonic shifting does NOT resolve the ambiguity of this sentence: tonic on 'plans' certainly means 'documents', but tonic on 'leave' can mean that 'plans' can be interpreted as EITHER 'intentions' OR 'documents', depending on context. It is a pity that a better example could not have been found for the discussion.

(vs. *George found some 'friends he'd like you to meet*). Her hypothesis was that these cases are all predictable from the NSR, without any special modifications, assuming that the NSR is ordered after all the syntactic transformations on each transformational cycle. This claim was counter to the basic assumption in Chomsky & Halle (1968:15) that '(prosodic) contours are determined in some manner by the surface structure of the utterance', for Bresnan was arguing that the contours are determined by their underlying sentence structure. She claimed that evidence for this was provided from the above examples, whereby the stress patterns found were argued as reflecting those of the simple sentences embedded within them in deep structure. In other words, a basic stress pattern is preserved throughout the syntactic derivation.

In her article, Bresnan placed great emphasis on the consequences her claim had for generative linguistic theory, in particular that her ordering hypothesis provided evidence for a lexicalist view of language (as it reinforced the view that all lexical insertion occurs on or before the first transformational cycle). It was unfortunate, in a way, that this emphasis was made, as it attracted a discussion of various general issues which was premature, in view of the absence of any clarification of certain rather more basic matters in her article. Lakoff (1972), for instance, took up the question of the lexicalist hypothesis and related matters, and attacked Bresnan on a mixture of observational and logical grounds; his own analysis was to see the NSR as a global rule which applied in the phonological cycle. Berman & Szamosi (1972) argued that her hypothesis could be rejected on observational grounds, that it made too many incorrect predictions and missed certain generalisations. In particular, they argued that there were many important cases where surface structure was essential for the determination of sentence stress (308-310). Their stronger argument was that the NSR as such is unworkable, because there are too many cases where primary stress assignment depends on factors other than structure—especially semantic factors.

There was a reply by Bresnan (1972), some further discussion by Stockwell (1972), and an important contribution by Bolinger (1972), which I shall discuss separately below.

My present aim is not to enter into a detailed consideration of the various arguments just referred to: I leave this to those who wish to work within the framework of assumptions used by that approach to generative grammar. My purpose is to use this debate as a taking-off point for a discussion of the merits of an alternative approach. The debate usefully focussed attention on the detail of some relatively neglected facts of prosodic distribution; but it suffered from a failure to realise that many of the difficulties it got into, and much of the dispute, was due to the set of theoretical assumptions which the analysts chose to work within. It is only natural, of course, once one has opted for a particular theoretical approach, to try to make that work at all costs. But intonation, and prosody in general, is a rather different kind of phenomenon from anything generative grammar had attempted to cope with previously, and the above debate showed very clearly the problem generative grammarians had to face—the problem of having to graft onto a well-developed syntactically-orientated model a phenomenon whose importance the model in its early processes of construction totally ignored, and whose study depends on assumptions incompatible with that model.

The most illuminating attack is to query the fundamental assumption of the generative approach, that intonational phenomena are predictable from syntax. This is the line taken by Bolinger (1972). This article attacks the basic assumption, which is taken for granted in the whole of the above debate, that the location of sentence accents can be explained by reference to syntax. He agrees that stress is explicable in this way, and goes on to point out that much of the confusion is due to a failure to distinguish word- and sentence-level phonology, stress belonging to the lexicon, and accent to the utterance. He argues that placement of accent is primarily conditioned by the speaker's view of how to distribute the semantic weight of his

sentence, for which it is necessary to take account of the entire context (linguistic and situational). 'The error of attributing to syntax what belongs to semantics comes from concentrating on the commonplace' (634). Highly predictable patterns (e.g. '*work to do*, '*clothes to wear*') will have the verbs unaccented; less predictable patterns will not (e.g. '*clothes to 'laundry*'). His article concentrates on demonstrating that sentences whose accentual pattern is said to be dependent on (predictable from) syntax can have this pattern readily altered by varying contextual factors. Tonicity is to be seen, in his view, in terms of the distribution of new information in a structure, directly reflecting the speaker's intent and essentially independent of the syntax. Talking about prosody in a contextual vacuum (641) will inevitably produce innumerable disagreements about the 'empirical facts'. and the citing of 'special rules' to handle apparent problems (as the generative debate displays, e.g. Bresnan, 1972:333 fn.). His conclusion involves a rephrasing of the NSR: instead of claiming that the main accent normally goes on the last stressable constituent, 'The intonational reality is, rather, that the speaker will put the main accent as far to the right as he dares, when assertive pressure is high; and he frequently (sc. e.g. in excitedly emphatic speech, cf. p.643, DC) dares to put it on a syllable (almost but not quite always one containing a full vowel) farther to the right than the recognized lexical stress... The distribution of sentence accents is not determined by syntactic structure but by semantic and emotional highlighting.⁴ Syntax is relevant indirectly in that some structures are more likely to be highlighted than others. But a description along these lines can only be in statistical terms. Accents should not be mashed down to the level of stresses,

⁴Cf. the reluctant conclusion of Berman & Szamosi (1972:313) who, after finding that optional stress placement must be taken into account, say: 'It seems that the grammar has to be able to assign both stress contours, and to designate either one as normal, in accordance with principles which are at least in part semantic'.

which are lexical abstractions. In their zeal to reverse Trager-Smith phonology, transformationalists have fallen into the same trap. Whether one tries to set up prosodic rules for syntax or syntactic rules for prosody, the result is the same: two domains are confused which should be kept apart' (644).

This approach seems quite correct, but it should be noted that it applies only to accent-placement (or tonicity), which was almost the entire subject-matter of the generative debate. It would be misleading to suggest that this or any other conclusion can be generalised to other aspects of intonational patterning, such as tone-unit distribution or type of nuclear tone (see below): a semantic explanation may be satisfactory for one aspect of intonation, but other explanations may be required elsewhere. The problem of intonation's functional complexity is in fact well recognised in the literature—as a signal of grammatical structure, of emotional expression, of semantic organisation, of social role. It is not necessary to go any further into a sub-classification of these functions in order to see that any theory which assumes that ONE of them is primary or basic for the whole of intonation is going to have difficulty justifying this decision, apart from appealing in advance to such vague, a priori notions as simplicity or insightfulness of generalisation. I therefore take the view that prosodic phenomena should be seen as an independent component of any model of language organisation, which interrelates with other components in various ways: some features of the prosodic component interact with syntax, some with sociolinguistic categories, some with affective meaning, and so on; but there is no claim that one interrelationship is in some sense prior to the others. And while on the one hand I accept Bolinger's conclusions about the semantic conditioning of tonicity, I shall be arguing later in this paper that tone-unit placement is in fact determined by syntax.⁵

⁵The role of both semantic and syntactic factors in accounting for nuclear tone type is discussed in Crystal, forthcoming.

There are however other difficulties with the generative approach which ought to be pointed out before moving to the constructive part of this paper. In particular, one has to comment on their curious semi-awareness of the problem of reliable data. They often remark on the methodological problems involved in their work, but—perhaps because of their traditional antagonism towards discovery procedures—they fail to do anything about it. Much of the difficulty arises out of a disagreement as to what the basic (i.e. 'normal') data to take into account are. Bresnan et al seem to have used one of two techniques to establish their basic data: they have either used their own intuition, to decide on normal intonation patterns, or they have asked their colleagues to read sentences aloud. Neither technique is a satisfactory basis for intonational study. I am not of course disputing the relevance of intuition as a datum for certain areas of linguistic analysis, but there seems no reason why one should assume that our intuition is equally powerful a means of judgement for all parts of language. Our conscious tacit knowledge of prosodic regularities is in fact negligible. It is well-known that response to a fixed set of prosodically contrasting sentences varies widely from day to day. Some days we accept more sentences as possible than on other days; the range of semantic interpretations varies considerably (even when presented with the data in multiple choice form); we rate as acceptable more the longer we are given to think about a sentence; and so on. Berman & Szamosi note one aspect of the complexity, but do so dismissively in a footnote: they feel that 'the oddness or normalness of a given stress contour depends partly on the speaker's ability to provide a satisfactory context', and add 'It is interesting to note that those of our acquaintances who are most inventive in concocting contexts in which "semantically anomalous" sentences are acceptable found the widest diversity of possible stress patterns' (1972:314). They are therefore—they say—hesitant about marking examples as ungrammatical. But they still do asterisk a large number of sentences, and their attack on Bresnan

depends entirely on a naturalness condition of some kind, either by showing that other patterns than Bresnan's are possible and normal, or by criticising her outputs as being impossible or abnormal. The point is made also by Lakoff (1972:286), who points to the normality of other stress patterns than those Bresnan gives, but who himself makes judgements about unacceptability, norms, contrastive stress, etc., presumably on a personal intuitive basis. It would be very easy to go through most of the asterisked examples in this debate and provide contrasts which make them acceptable, without necessarily introducing a contrastive stress element, but Bolinger has already shown how elementary a matter this is.

In other words, the empirical basis of the whole argument is open to question ('insecure' is the word Chomsky uses for it, 1969:24—and see his fn. 21 for his awareness of the problem of determining normal intonation, though he does not attempt a solution). As long as the analyst relies wholly on his own intuition for the verification of his prosodic hypotheses, his approach remains suspect: one can convince oneself in minutes that a pattern is possible, that problem patterns are idiolectal or dialectal (cf. the unsupported speculations in Berman & Szamosi 1972:314), and so on. Before one can interpret claims that 'most people' do such-and-such in intonation, one needs some rather basic background information—like how many informants were asked, whether they were phonetically naive, whether they were well-educated, how the task was presented to them. At least Berman & Szamosi try, but their results are unimpressive. They give the following sentences as 'normal, non-contrastive, non-emphatic' patterns (312): (a) *We liked the proposal that George 'mentioned*; (b) *Whose church did they 'vandalise*; (c) *There are new worlds to 'win*; (d) *There are new areas to 'explore*. They say: 'Most people (sic) find that shifting the primary stress... to the head noun (or questioned noun)... results in sentences that are distinctly odd. (Of course, they can all be given contrastive interpretations, in which case there is nothing strange about them.)' (312-13).

Bolinger took up the point that, given appropriate context, the above could be given just as contrastive an interpretation as the alternatives. I simply want to add that asking informants to judge for 'oddness' in these sentences does not obtain the agreement Berman & Szamosi claim. Even if one uses their technique, and asks informants to say the sentences out loud, there is considerable variability in the response. For example, when 30 educated and phonetically naive British English speakers were asked to say the above sentences aloud, the following tonic placements were obtained: (a) tonic on *liked* (4 times), *proposal* (4), *George* (8), *mentioned* (9); compound tonic on *liked* and *mentioned* (2), *proposal* and *mentioned* (2), *George* and *mentioned* (1); (b) tonic on *whose* (1), *church* (17), *they* (1), *vandalise* (11); (c) tonic on *worlds* (10), *win* (18); compound tonic on *worlds* and *win* (2); (d) tonic on *areas* (14), *explore* (11); compound tonic on *areas* and *explore* (5).

All of this indicates the unreliability of the data used in the above papers. To the extent that the authors realised the inadequacy of their own intuitions and asked other informants, their data is useful insofar as they make their methodology explicit. At times, however, their methodology seems positively in error. It would seem that one technique that was used was to present colleagues with a written version of a sentence and ask them to read it aloud, e.g. Berman & Szamosi, 1972:314 fn. 11). The prosodic response was then taken as the basis of a judgement of normality. But this is in principle a suspect procedure. Berman & Szamosi give us no information about how they went about judging for normality, what statistical techniques they used (if any), and so on. Were their informants equally and consistently confident in their output (cf. Davy & Quirk 1969)? At the very least, one would want to present these informants with a set of alternatives which they would be asked to react to (e.g. by rank ordering)—otherwise one will reach premature conclusions (for example, assuming that the first pattern to be produced is the 'normal' one, whereas the sentence may have more than one version to which the term

'normal' may appropriately be applied, a possibility Berman & Szamosi themselves recognise, p. 313—though one which is apparently denied by Stockwell 1972:87-88). There was apparently some variability: in referring to a certain set of sentences (p. 315), Berman & Szamosi state 'sentences like 40 are almost never ambiguous, because a unique stress contour is associated with each of the possible readings'. A footnote (13) adds: 'We refer, of course, to the most natural readings of the sentences, factoring out contrastive or other such marked interpretations'. But who decided the naturalness question here? Berman and Szamosi? Using their intuitions solely? Or does 'factoring out' mean what it says? My impression is that they, as others, are using informants in a thoroughly unscientific way, and the force of their case is correspondingly much reduced. (In any case it should be pointed out that notions of naturalness, contrastivity and markedness are extremely obscure.)

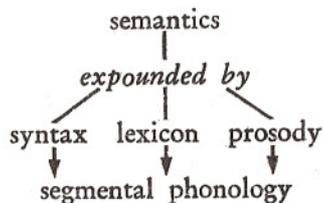
But there is a second problem with the reading aloud test: even with proper methodological controls, it works only with short, simple sentences—and these are sentences which on the whole present few problems. As soon as a sentence gets at all complex, or as soon as the context is given in full, punctuation enters in. Now what does one do about this? If one leaves the punctuation in, in presenting a text to an informant, one is begging the question by giving the informants an explicit indication of where the tone-unit boundary falls. If one leaves it out, one is presenting the informants with an unfamiliar (and probably ambiguous) representation, which they will be unable to give a 'natural' reaction to. And there may be other difficulties, to do with the written medium on the spoken form, e.g. names beginning with a capital letter might tend to attract the accent. (This is not simply hypothetical: my same thirty informants were given the following sentences, separated from each other by other tasks: (a) *I've just bought some excellent badges*, (b) *I've just bought some Liverpool badges*. For (a), all 30 put the tonic on *badges*; for (b), 18 put the tonic on *badges*, and 12 on *Liverpool*.)

I can think of no way in which one can extract information about prosodic acceptability from informants without to some extent prejudging the results or introducing uncontrollable variables; and this is why I insist on the regulative function of the corpus. As far as I can tell, none of the authors being discussed—except, of course, Bolinger—have systematically analysed a corpus of spontaneously produced, recorded material. If they had, many of Bolinger's criticisms would have been anticipated. Perhaps this is the traditional suspicion of corpus-based work emerging again. But I would point out that one reason why Chomsky's strictures of corpus work on syntax were so well-received was because people had already spent years analysing the syntactic patterns of corpora, and knew what the data were like. Chomsky's remarks made a great deal of sense. But we are now talking of intonation, not syntax, and here we lack a corresponding tradition of detailed descriptive work which we can use as a basis for formulating more powerful hypotheses. Bresnan and others, it seems, were jumping the gun. They were trying to develop a generative account of prosodic patterns without having carried out the necessary descriptive spadework. And this spadework means, in effect, the transcription of quantities of spontaneously produced material. No-one would claim that the description of this material could be anything other than observationally adequate for the corpus in question, but it can provide the necessary controls on the process of extending the analysis into the realm of competence.⁶

The above debate is largely of historical interest now, in that alternative theories of generative grammar have developed within which it would be easier to incorporate prosodic

⁶This paragraph would have been unnecessary if generative authors had familiarised themselves with the wealth of descriptive detail collected in the European pedagogical tradition. It is really rather amusing to see examples of constructions attributed with thanks to recent lectures of Ross or Lakoff, when they are cited in the work of Henry Sweet or Joshua Steele!

phenomena (though the cautionary points about methodology obviously still apply). For example, there was the basic difficulty with the standard model of knowing what to do with intonational primitives once they were permitted within the base component (cf. Stockwell, 1960). Semantically-oriented models of language structure, whether generative or interpretive, will presumably find it much easier to take intonation into account, because they allow a more direct relationship between intonational form and semantics than could be presented in terms of the standard model. Given a set of semantic categories, one might postulate a prosodic component which works along with the syntactic and lexical as a means of realization, as follows:



But I have not seen a detailed attempt to incorporate intonation into an approach of this kind. Nor has the interpretive hypothesis been fully explored from this point of view. In Chomsky (1969), tonicity provided the main argument for the view that semantic interpretation has to be partially determined by surface structure; but this essay was of course only exploratory: it illustrated the relevance of intonation in relation to the concepts of focus and presupposition⁷ from only a small set of sentences, and it did not go into crucial questions of formalization. It was (and still is) quite unclear (a) how one would get from the rules of phonological interpretation, which would assign an intonation contour to surface structures, to the seman-

⁷The focus is a phrase containing the intonation center; the presupposition, an expression derived by replacing the focus by a variable' (26).

tic rules, which would interpret a phrase containing an intonational centre as a focus of utterance, and how the semantic rules actually operate; (b) how certain surface structure phrases would be marked to receive 'contrastive stress' ('by grammatical processes of a poorly-understood sort', p. 25), and in what ways these would be allowed to affect the operation of the phonological rules; and (c) how to account for gradience in naturalness of response in cases of contrastive stress, whereby 'naturalness declines far more sharply as longer and longer phrases containing the intonation center are considered as a possible focus' (25). In the absence of any suggestions as to how these problems should be tackled, the merits of an alternative approach, without taking sides on the generative vs. interpretive issue, could usefully be explored. The main features of this approach are the splitting of the phonological component into two, the SEGMENTAL and the NON-SEGMENTAL, and of the latter into a set of functionally distinct sub-components.

The Present Approach

Intonation cannot be slotted into one's model as a single category: it is not a unitary, homogeneous phenomenon. It is often talked about as if it were—when one hears references to 'the intonation of a sentence' or to 'the learning of intonation', for example. I have argued in earlier papers that intonation is not a single system of contours or levels, but the product of the interaction of features from different prosodic systems—TONE, PITCH-RANGE, LOUDNESS, RHYTHMICITY and TEMPO in particular. For example, a particular tone (e.g. FALLING) can be seen to vary in terms of its relative height (e.g. HIGH vs. LOW) and width (e.g. WIDE vs. NARROW); a stretch of utterance can be articulated as 'parenthetic' if it is given LOW pitch range with optional PIANO loudness and ALLEGRO tempo. The various features are given a partly hierarchic organisation, such that the basic unit of prosodic organisation, the TONE-UNIT, is seen to consist minimally of a TONIC SYLLABLE, expounded by one of a set of NUCLEAR TONES (falling, rising, etc.), and optionally

preceded and followed by other syllables involving differing degrees of pitch- and loudness-prominence. The full system is given in Crystal 1969: see especially Chapters 4 and 5. In the rest of this paper, I want to look at the main structural characteristics of this model—and in particular at the concept of the tone-unit—from the point of view of how they might be integrated within a broader model of language structure. What is immediately clear from this approach, regardless of its merits, is that there is far more to the study of 'intonation' than tonic placement, or TONICITY. But the generative debate for some reason has arbitrarily restricted the subject-matter of intonation to tonicity. This is a restriction which goes back to Chomsky & Halle, and it needs to be removed. In fact it does not stop the writers from referring to other prosodic matters when they find it necessary to do so, but these references are always introduced in an ad hoc way, or not systematically distinguished from tonicity. In particular, there is a failure to give appropriate recognition to the independent roles of tone-unit boundary features and nuclear tone type, and to the systematic basis of other prosodic contrasts (such as pause, rhythm). The only exception, as far as I can see, is Stockwell (1972:103 ff.), who refers to two unpublished dissertations on the topic of phrase boundaries and intonation within the generative framework, and in commenting on Pope's (1971) view that intonation assignment depends on prior stress assignment, argues for tonicity and tone to be seen as autonomous systems (96-7): 'It is quite likely that the contour, and its center, are altogether independent phenomena'.

The present view is based on an analysis of the prosodic patterns used in a corpus of some eight hours of informal, spontaneous conversation, in which the participants were unaware that they were being recorded. This analysis is intended to justify the particular theoretical approach used, which, for the sake of clarity, will be given some exposition first. Figure 1 is a first attempt at a model in which the main constructs required to handle my data are interrelated. The first four

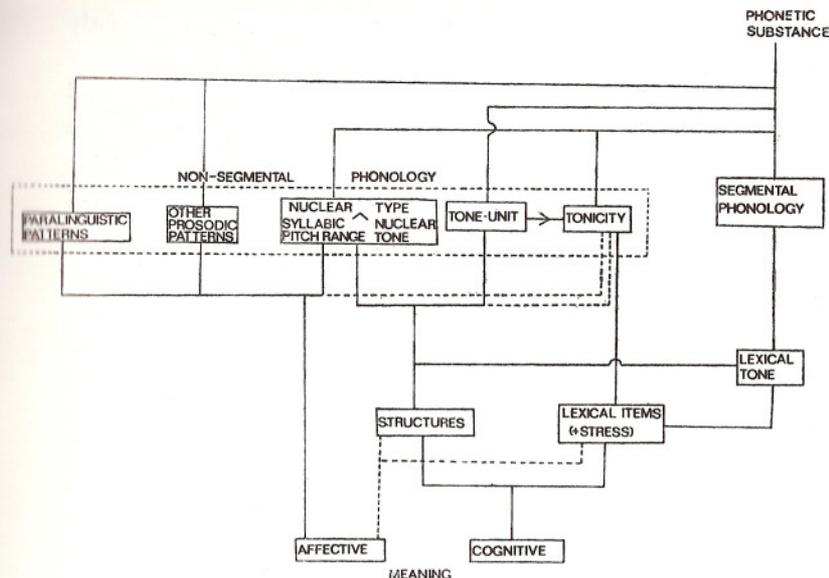


FIGURE 1. Functional model: non-segmental phonology and other components.

points which follow are preliminary assumptions, which I wish to take for granted for the purposes of this paper.

1. The model may be interpreted as a model of either production or reception; but for the purposes of illustration, I will outline its operation in terms of a process of production.

2. MEANING and PHONETIC SUBSTANCE are taken as given: the question is to determine what components are necessary in order to interrelate them.

3. A distinction between syntax and lexicon is here accepted, and represented using the labels STRUCTURES (*sc.* syntactic and morphological structures and categories, as given in some grammatical handbook) and LEXICAL ITEMS (as listed in some dictionary). It is assumed that the internal stress pattern of lexical items is given as part of this description, and no further reference will be made to stress in this paper. (This view of

stress as a word-level concept is also Bolinger's, as already mentioned.)

4. A distinction between AFFECTIVE (or attitudinal) and COGNITIVE meaning is accepted as necessary for any analysis of intonational function. It is not assumed that these are the only two 'types' of meaning which need to be recognised; nor is it assumed that a clear aprioristic distinction can be made between them. (The dotted line between AFFECTIVE and STRUCTURES indicates the possibility of word-order variation, inter alia, expressing particular attitudes: see Charleston 1960.)

The points in Figure 1 which I consider controversial for the purposes of this paper are as follows.

5. A fundamental distinction is made between SEGMENTAL and NON-SEGMENTAL phonology. Segmental phonology is seen as an interpretive component of this model which (in the analysis of production) has as input the syntactic structures and categories and lexical items, integrated in some previously specified way. The various processes of non-segmental phonology are seen as operating AFTER the segmental rules⁸ have applied, as the lines emerging in the top right of the figure indicate. There are three reasons for this. Firstly, it is more convenient to specify the set of assimilations and elisions which become operative as an utterance increases in speed AFTER a normal-speed segmental analysis has been made. It is not possible to decide which of many possible degrees of reduction (e.g. *have* becoming [həv, əv, v, f]) could be established as a base-form, and the full form of the word remains the obvious choice. Likewise, junctural features at a tone-unit boundary which operate on word-final segments are best seen as operating after word-level phonology, for obviously most words in the language

⁸This is the most widely-used term, and I find it useful; but I do not presuppose that any such rules will necessarily look like those currently formulated in a generative context.

may occur in non-final as well as final position within the tone-unit. Secondly, vowels in tonic monosyllables increase in length: short vowels become long, and long vowels become longer. This fact is best taken into account after a set of vowels has been established, rather than being allowed to complicate statements of length within the segmental component. Thirdly there is a psycholinguistic point, namely that in the process of reading, the segmental graphemic properties are in a 1-1 relationship with phonemes, and since the model ought to be able to subsume both speech and writing, it is more satisfactory to see the non-segmental phonology as 'added to' segmental utterance in the spoken medium, and only partially added in the written.⁹

6. A minor point is that LEXICAL TONE—that is, the use of pitch (or some other feature) as the direct exponent of a syntactic or morphological category, or as part of the phonological identity of a lexical item—is assumed to operate before the rules of segmental phonology.

7. Within the non-segmental phonological component, this model recognises five functionally independent categories: tone-unit, tonicity, nuclear type, other prosodic patterns, and paralinguistic patterns. In a more sophisticated model, it will certainly be necessary to present the two latter categories in a more adequate way (e.g. recognising such distinctions as rhythmicality and pause), but this has not been attempted here, as unless this model works for the main non-segmental features of utterance, there is no point in trying to extend it to other areas.

8. The main empirical claims made by this model are reflected in the solid lines connecting the various components. They

⁹It should be pointed out, however, that the metaphor of 'adding' applies to a view of language solely in terms of production. When comprehension is taken into account, the primary role of non-segmental features as an initial datum for semantic interpretation has often been recognised ('It's not what he said, but the way that he said it', etc.).

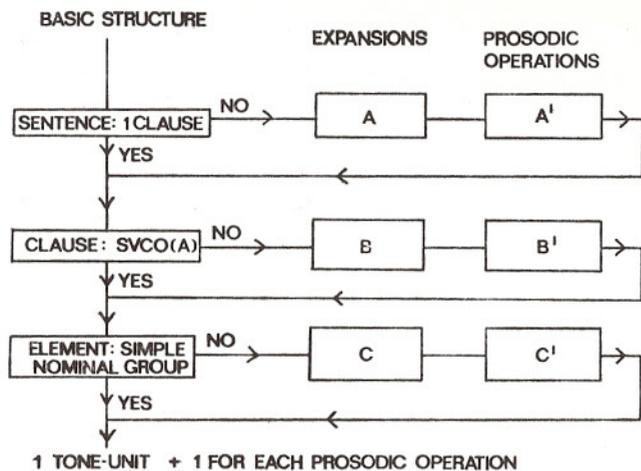
are four. Firstly, it is claimed that the placement of tone-unit boundaries is determined by syntactic structure. Secondly, it is claimed that tonicity is primarily determined by lexical or semantic factors, sometimes by specific structures, and sometimes by affective information (as the broken lines indicate). Indirectly, of course, tonicity is dependent on syntax, in that tonicity requires the prior establishment of a tone-unit to define its domain, and tone-units are determined syntactically. The arrow in the figure indicates this indirectness. Thirdly, it is claimed that nuclear types are determined both by structural and affective meaning. Fourthly, it is claimed that other prosodic and paralinguistic patterns are determined by affective meaning, and are unaffected by syntax or lexis. This is very much a simplification, but it will suffice as a first approximation.

9. Two other points should be noted. (a) As suggested in the earlier discussion, tone-unit boundaries are considered to have an independent organisational role from that of the tonic syllable expounded by a nuclear type. The motivation for this distinction is to take account of the facts of language acquisition mentioned above, where the first sign of language-particular patterning is the organisation of what Weir calls 'sentence-like chunks' (1962) out of a largely undifferentiated babbling hitherto; these chunks are primitive tone-units, which operate some time before the appearance of definable nuclear tones or tonicity contrasts (the latter, of course, not appearing until the emergence of syllable sequences, much later). Some evidence for the distinction would come from the study of tongue-slips and related matters, where it is argued that the tone-unit is the fundamental unit of neural encoding, with slips rarely crossing tone-unit boundaries or affecting tonic placement (see Boomer & Laver 1968:8-9, Laver 1970:69 ff.). (b) Other prosodic and paralinguistic patterns outrank tone-unit and other intonational organisation, as the lines at the top right of the figure indicate, in that in output, variations in speed, rhythm, loudness, etc., often reduce, subordinate or eliminate

many of the intonational contrasts. The most important variable here is speed, whereby increased speed of utterance reduces the number of tone-units in an utterance, and vice versa. (This recalls Bierwisch's notion of 'optional phrasing', 1966.) Conversely, in input, one might expect an initial processing of speed, loudness, etc. to take place as a preliminary to more detailed processes (as in one's initial reaction to voice stereotypes).

Tone-Unit Boundaries

I must now go on to consider the empirical content of the claims made in 8 above, using the corpus of data referred to earlier. In this paper, only the first claim will be tested. (Discussion of the other claims takes place in Crystal, forthcoming.) What syntactic factors determine tone-unit boundaries? At the moment, it is an open question whether the tone-unit is best described with reference to the unit sentence or to some other unit, e.g. clause (as Halliday (1967) maintains, for instance) or element of clause structure (as Crystal (1969: Ch.6) argues). Bearing in mind the need to integrate one's treatment with generally accepted models of grammar (a point not fully appreciated in my earlier work), the obvious starting-place would seem to be the sentence. Figure 2 indicates the operations required to assign tone-unit structures to sentences. (The grammatical apparatus used here, and in the following pages, is that of Quirk et al 1972.) Given an input of a sentence, then (following the left-hand side of the Figure downwards) if this sentence consists of one clause; and if this clause consists maximally of the elements Subject + Verb + Complement + Object, with one optional Adverb, in this order; and if each of the elements S, C, O, or A is expounded by a simple nominal group: then the sentence will have a single tone-unit. This is considered to be the basic pattern. Examples are: /The big boy kicked the ball yesterday/, /We gave him a lift in the car/, /Go away/, /I asked him/, /He spoke/.



OUTPUT

FIGURE 2. Operations for assigning tone-unit structure to sentences.

The above needs two clarifications, concerning the notions of clause structure and simple nominal group. As regards the former, it should be clear that different combinations of elements of structure will occur, depending primarily on the type of verb, e.g. SV, SVC, SVO, SVOO. I assume the classification of Quirk et al (1972:§2.10). Minor (elliptical) sentences, e.g. S, O, are also subsumed under the above statement. Concerning the second point, Quirk et al distinguish 'simple' nominal groups (§13.76), in which a noun is modified only by closed-system items (*the, that...*) or realised solely by a pronoun or proper name, from 'complex' nominal groups, which involve heavier modification. Within complex groups, they distinguish (i) modification consisting of one adjectival premodifier and/or one prepositional phrase postmodifier, from (ii) all other cases of 'multiple' modification. They discuss these distinctions from the point of view of the stylistic differentiation of texts. From the intonational point of view, however, it makes better sense to class together simple and class (i) complex groups, as they

are identical as far as intonational distribution is concerned. 'Multiple' for me therefore includes all nominal groups containing minimally either two premodifying elements (determiners excluded) or two postmodifying elements (cf. C3 below).

Returning now to the top left-hand corner of the figure, we find that various possible syntactic EXPANSIONS occur at each of the grammatical levels recognised in the model. Each expansion carries with it a PROSODIC OPERATION which involves the placing of a tone-unit boundary (or boundaries) at a point in the construction which demarcates the point of expansion of the basic structure, and which does not already have a boundary present (due to some previous prosodic operation). Thus at clause level, the sentence may contain an indefinite number of clauses: each clause is then subjected to a prosodic operation which gives it an intonational identity. The output from the clause level is then fed into the next level, and EACH clause in the sentence is analysed for its element structure: if it contains other elements than those recognised in the basic pattern above, or a marked order of elements, each of the expansions (under which term I subsume reordering) is given an intonational identity. The output, as a sequence of elements of clause structure, is then fed into the next level, and each nominal group (or phrase containing a nominal group) is analysed for its complexity, in the above sense, any expansions being intonationally identified. No further levels of grammatical structure affect tone-unit boundary placement. The output of tone-units is therefore ready for analysis in terms of placement of the tonic syllable (see Crystal, forthcoming).

Now follows the set of expansions which need to be recognised in order to account for the corpus. This is at present little more than an inventory. Doubtless it will prove possible to interrelate these structures more elegantly, but in the first instance I was anxious to present the facts with a minimum of theoretical overlay. The symbol : separates a syntactic expansion from its prosodic operation. / = tone-unit boundary. A, A', B, B', C, C' refer to Figure 2.

Expansions A: Prosodic Operations A'

1. *Structural parallelism*: *put / after each component*, e.g. /he came at three/ left at four/ was home by five/ and in bed by six/.

2. *Coordinate clauses*: *put / after each component*, e.g. /John picked up the phone/ and a voice asked him the time/; /either you've got the answer/ or I have/; /he asked me what I'd said/ and whether I was interested in more/. There are a number of categories of exception to this rule, which have not been fully explored. The most important case is Ellipsis. No boundary is used if the Subject of a coordinated clause is elided, and a sequence of coordinated verbs produced, e.g. /Susan will sing and dance/. On the other hand, the ellipsis of Object or Complement or part of the Verb from the first clause in a coordination requires the prosodic operation: *put / at the point of ellipsis, and optionally at the parallel point in the coordinated clause*, e.g. /Gerald likes/ but Peter hates/ Mary/ (cf. Quirk et al § 8.91); /it wasn't Jim/ but John who asked her/ (cf. Quirk et al § 14.18).

3. Subordinate clauses.

(i) *Adverbial initial*: *put / after clause*, e.g. /when he comes/ tell him I'm out/.

(ii) *Adverbial medial*: *put / on either side*, e.g. /the man in the corner/ if you must know/ is my cousin/.

(iii) *Adverbial final*: *put / before clause*. The conditions operating here are not entirely clear. / is obligatory if the adverbial status of the construction needs to be made clear, e.g. /tell me/ to save time/ (= in order to save time), as opposed to /tell me to save time/, which is ambiguous between adverbial and object. It is also obligatory when the preceding sentence structure is long or complex, as in /you're all going out to the museum on Saturday/ if the weather stays fine/. (The question of whether phonological length of the preced-

ing structure as such is a determining factor needs to be investigated.)

(iv) *Nominal as subject*: *put / after subject* (cf. Quirk et al, § 11.16), e.g. /what I said/ was of the utmost importance/; /how the book will sell/ is impossible to say/.

(v) *Medial non-restrictive*: *put / on either side*, e.g. /my brother/ who's abroad/ sent me a letter/; /the man/ dressed in a raincoat/ came towards me/. The boundary between non-restrictive clauses and phrases is by no means clear-cut (cf. Quirk et al § 13.31 and C4 below).

(vi) *Final non-restrictive*: *put / before*. This includes those with sentential antecedent (cf. Quirk et al § 13.15), e.g. /he likes linguistics/ which surprises me/.

(vii) *Appositive*: *put / after* (cf. Quirk et al § 13.16), e.g. /the fact is/ that he doesn't love her/.

4. *Medial parenthetical clause*: *put / on either side*, e.g. /each one of the children/ I insist/ will have to go/; /Michael Smith/ he's the one in the black suit/ has just got married again/. There is no clear boundary between clause and phrase here, once again, as Quirk et al point out; consequently one might include here (rather than under C below) examples of the following type: /Michael Smith/ the man in the black suit/ has just got married again/. Exclamatory asides, because of their parenthetical properties, ought also to be taken at this point, and not with apposition under C, e.g. /John/ the butcher/ he's ruined all my plants/ — which might be contrasted with /John the butcher/. . . , where John deals in meat.

5. *Direct speech*: *put / after the reporting verb*, e.g. /then they said/ who's coming/ (= . . . said, 'Who's coming'.); /tell me/ where's your brother/; /naturally he said/ I'm very interested/ (= 'Naturally', he said, 'I'm very interested'.)

6. Comment Clauses (cf. Quirk et al § 11.65)

(i) *Initial*: *put / after*, e.g. /you know/ I think it's going to rain/ - cf. /you know I think. . . /

- (ii) *Medial, emphatic: put / on either side*, e.g. /and Mrs Jones/ you know/ is going to be there/ ('you know' = 'you do know her, don't you')
- (iii) *Final, emphatic: put / before*, e.g. /and everything came out quite well/ you know/.

7. *Tag utterances: put / before*, e.g. questions, /you're staying/ aren't you/; statements, /that was a lark/ that was/; /he likes a drink/ Jim does/. Tag statement clauses with elided verb might also be included at this point, e.g. /he's coming/ John/; cf. Quirk et al § 14.50.

Expansions B: Prosodic Operations B'

1. *Initial vocative: put / after*, e.g. /John/ are you coming out with us/.

2. *Adverbials* (conjuncts/disjuncts, in the sense of Quirk et al Ch. 8)

- (i) *Initial: put / after*, e.g. /therefore/ I think he's right/; /normally/ he doesn't take medicine/ (cf. /he doesn't take medicine normally/ = 'in the usual manner'); /every six weeks/ he visits us/. I also include interjections, etc. at this point, e.g. /well/ I think he will/.
- (ii) *Medial; put / on either side*, e.g. /he did say/ on the other hand/ that he'd like to/; there were /with respect/ twelve people there/.
- (iii) *Final: put / before*, e.g. /I didn't ask him/ anyway/; /he didn't read the book/ stupidly/ (i.e. 'unfortunately'). Only a subset of conjuncts and disjuncts operate in this way (cf. Quirk et al § 8.91). I have not investigated how this subset might be defined.

3. *Adverbial sequence: put / after each element*, e.g. /I spoke to him quietly/ without fuss/.

Expansions C: Prosodic Operations C'

1. *Structural parallelism: put / after each component*, e.g. /I went to see the girls from class 1/ the boys from class 2/ and all prefects/; /he's gone to buy gin/ whiskey/ eggs/ and tea/.

2. *Multiple heads*

- (i) *Separate premodification: put / after first head*, e.g. /in that shop you'll find some very nice chairs/ and tables/ (i.e. the tables are not necessarily very nice; cf. /... very nice chairs and tables/).
- (ii) *Separate postmodification: put / after first head*, e.g. /the man/ and the woman dressed in black/ came to see us/ (i.e. only the woman is dressed in black; cf. /the man and the woman dressed in black/ ... , where both are).
- (iii) *Non-restrictive apposition* (cf. Quirk et al § 9.133): *put / on either side of the apposed phrase*, e.g. /Mr Jones/ the architect/ is over there/ - cf. /Mr Jones the architect/ ... , which is restrictive, and implies an opposition with 'Mr Jones the butcher' or 'Mr Smith the architect'. This also includes the use of apposition markers (cf. Quirk et al § 9.138), e.g. /the plane/ or rather the elephant/ will be there by two/.
- (iv) *Noun phrase tags: put / before*, e.g. /they're all the same/ these unions/.

3. *Multiple modification*

- (i) *Premodification, general adjectives: put / after each component except the last*, e.g. /I was talking to that very tall/ pretty/ rather awkward girl/. This also includes coordinated premodification, e.g. /an equally serious/ but more interesting situation/ is... In Subject position, / is put after the completed Subject element, e.g. /the very tall/ pretty/ but rather awkward girl/ was talking.../. The term 'general' excludes certain adjectival categories which display order-restrictions (e.g. colour,

nationality) and which would not be split intonationally, e.g. /the big red house/: see Quirk et al § 13.65. However, if C1 is applied, and the adjectives are introduced as if in a list, even order-restricted adjectives are affected, e.g. /it was a big/ red/ Gothic kind of building/.

- (ii) *Postmodification, in subject: put / after subject* (cf. Quirk et al § 13.39), e.g. /the man in the raincoat standing near the bus-stop/ is my brother/. An additional / is introduced where ambiguity has to be avoided, cf. /the man in the car/ waiting near the bus-stop/ ... and /the man in the car waiting near the bus-stop/ ...
- (iii) *Postmodification, in passive agent: put / before agent*, e.g. /the butler had been fiercely attacked/ by a man wearing a raincoat and holding a gun/.
- (iv) *Postmodification, in non-final object: put / after object*, e.g. /I gave the book you'd been waiting for/ to the man from upstairs/. This seems obligatory only when the postmodification is clausal, but there is a strong tendency for / to be used even with phrases, e.g. /I gave everyone in the room/ a big wave/.

4. *Medial non-restrictive phrases: put / on either side*, e.g. /the man/ in a raincoat/ came towards me/.

An illustration of the application of the above rules to the corpus is as follows. Here is a sentence as it was first transcribed (though certain features of the transcription not relevant to this paper are omitted). Short and long pauses are represented as . and — respectively; words containing the tonic syllable are in small capitals, with an indication of the pitch movement over the stressed syllable (' rising, ` falling).

the second deplorable thing ABÓUT it/—is the FÁCT/ . that .
 THÍŠ CHÁP/ . a NEURÔTIC/—his mother calls him a NEURÓTIC/
 from the age of TWÓ/ . who can't stick in the army for twenty-
 four HÓURS/ is the kind of PÉRSÓN/ who is made a public idol
 of the DÁY/ —

Given a grammatical specification of this sentence in terms of Quirk et al, the following operations would apply:

A: *clause level: / after fact* (under A3 vii), the first instance of *neurotic* and *two* (under A4), *hours* (under A3v), *person* (under A3 vi), and *day* (sentence-final). The sentence now looks like this:

the second deplorable thing about it is the fact/ that this chap a neurotic/ his mother calls him a neurotic from the age of two/ who can't stick in the army for twenty-four hours/ is the kind of person/ who is made a public idol of the day/

B: *element level: no operations applicable.*

C: *group level: / after it* (under C3 i), *chap* (under C2 iii) *neurotic* (under C3 iv). The sentence is now in its proper tone-unit form, and goes forward for tonicity insertion, obligatory pause insertion (e.g. pauses required around the appositional group, *a neurotic*), etc. (see Crystal, forthcoming).

Out of the 12,000 tone-units examined, about 100 were unable to be predicted from the above rules.¹⁰ How are these to be dealt with? An example of such a sentence was /we gave him/ a lift/ on a Tuesday/, where the two internal boundaries are not predictable from the above rules. There are only four possible explanations. Firstly, the usage is sociolinguistically explicable (or 'stylistically' explicable, in the sense of Crystal & Davy 1969): that is, it is predictable for the entire output of a particular group, e.g. sports commentator or political speech-maker. This hypothesis would be easy to check statistically. Secondly, the usage is individually explicable: a predictable

¹⁰To avoid the charge that the transcription was biased by an awareness of the grammatical constraints outlined in this paper, I should make it clear that the prosodic transcription of the data used here was carried out between 1963 and 1970, the transcription being checked by at least two analysts, the sole criterion that they were given to work with being auditory agreement as to the prosodic variables involved (see Crystal 1969:Ch.1). The present paper was prepared for a conference in 1973. It is of course possible that we were all of us being unconsciously influenced in our transcription by some innate knowledge of these rules!

idiosyncrasy, which may be analysed as part of voice quality or idiolect. There would presumably be no difficulty in verifying this hypothesis either. Thirdly, the usage is (socio- or psycho-)linguistically unpredictable because of performance interference, e.g. a physiological reflex (such as a cough, or a pause for breath) interrupts speech flow, or the speaker arbitrarily chooses to restart a construction. In all such cases, there will be some other indication than the non-segmental anomaly that an aberrant utterance is in progress. Lastly, the usage is attitudinally predictable: tone-unit boundaries are put in or omitted because of the attitude of the speaker. This is the problem area, because of the difficulty of specifying the range of possible attitudes precisely. But in principle it should be clear that, given an alphabet of affective category-labels applicable to the whole of a sentence (and not just the non-segmental component), it would be possible, given enough data, to establish the attitudinal conditions which affect the placement of tone-unit boundaries. For example, one might say: if the utterance is interpreted as X (excited, tired, irritated. . .), then Y prosodic operations apply. Impressionistically, this is not difficult to do. The above sentence is explicable in this way, for instance: here a husband's story had been interrupted by his wife, who had queried whether the hitch-hiker had been picked up on a Tuesday; the husband, plainly irritated, replied with the above. One might then hypothesise that in the attitudinal context IRRITATION, optional tone-unit boundaries are introduced starting at clause level and continuing downwards, depending on the degree of irritability present, up to and including the morphemic level, e.g. /we gave him/ a lift/ on a Tuesday/; /we/ gave him/ a lift/ on/ a Tuesday/; /we/ gave/ him/ a/ lift/ on/ a/ Tuesday/—/in/dis/putably/ (he might have added!) The matter obviously requires further study. Given the above norms, under what circumstances will data present (a) more and (b) less tone-units? Other prosodic variables will have to be considered (especially TEMPO and PAUSE), as will the question of the type of grammatical unit

with which tone-units can best be correlated. There is some evidence in the data, for instance, to suggest that clauses and sentences are the units of organisation for relatively informal, fluent discourse, but elements of clause structure for more formal or less fluent situations. It is hoped that, using the cautionary perspective and normative framework of the present paper, such questions—along with the descriptive and methodological issues raised earlier—will now begin to be systematically investigated.

BIBLIOGRAPHY

- Bazell, C.E. Choice of Criteria in Structural Linguistics. *Word* 10.126-135. 1954.
- Berman, A. and M. Szamosi. Observations on Sentential Stress. *Language* 48.304-325. 1972.
- Bierwisch, M. Two Critical Problems in Accent Rules. *Journal of Linguistics* 4.173-178. 1968.
- Bolinger, D.L. Intonation and Analysis. *Word* 5.248-254. 1949.
- . Intonation - Levels v. Configurations. *Word* 7.199-210. 1951.
- . A Theory of Pitch Accent in English. *Word* 14.109-149. 1958.
- . Accent is Predictable (if You're a Mind-Reader). *Language* 48.633-644. 1972.
- Boomer, D.S. and J. Laver. Slips of the Tongue. *British Journal of Disorders of Communication* 3.2-12. 1968.
- Bresnan, J.W. Sentence Stress and Syntactic Transformations. *Language* 47.257-281. 1971.
- . Stress and Syntax: a Reply. *Language* 48.326-342. 1972.
- Charleston, B.M. Studies on the Emotional and Affective Means of Expression in Modern English. *Swiss Studies in English* 46. Berne 1960.
- Chomsky, A.N. Deep Structure, Surface Structure, and Semantic Interpretation. Indiana Linguistics Club 1969. Reprinted in R. Jakobson & S. Kawamoto (eds.), *Studies in General and Oriental Linguistics*. Tokyo.
- and M. Halle. *The Sound Pattern of English*. New York 1968.
- Crystal, D. *Prosodic Systems and Intonation in English*. London 1969.
- . Non-Segmental Phonology in Religious Modalities. Paper Read to Georgetown Round Table Section on Language and Religion 1972.
- . Non-Segmental Phonology in Language Acquisition: a Review of the Issues. *Lingua* 1973.
- . The Analysis of Nuclear Tone. Forthcoming, *Festschrift for D.L. Bolinger*.

- Crystal, D., and D. Davy. *Investigating English Style*. London 1969.
- Davy, D., and R. Quirk. An Acceptability Experiment with Spoken Output. *Journal of Linguistics* 5.109-120. 1969.
- Gumperz, J., and E. Herasimchuk. *The Conversational Analysis of Social Meaning: A Study of Classroom Interaction*, Mimeo, University of California 1971.
- and D. Hymes (eds). *Directions in Sociolinguistics: The Ethnography of Communication*. New York 1972.
- Haas, W. The Identification and Description of Phonetic Elements. *Transactions of the Philological Society* 118-159. 1957.
- Halliday, M.A.K. *Intonation and Grammar in British English*. The Hague 1967.
- Hinde, R.A. (ed.). *Non-Verbal Communication*. London 1972.
- Kaplan, E.L. *Intonation and Language Acquisition. Papers and Reports on Child Language Development* 1.1-21. 1970.
- Lakoff, G. The Global Nature of the Nuclear Stress Rule. *Language* 48.285-303. 1972.
- Laver, J. The Production of Speech. In J. Lyons (ed.), *New Horizons in Linguistics*, 53-75. Harmondsworth 1970.
- Newman, S.S. On the Stress System of English. *Word* 2.171-187. 1946.
- Pope, E. Answers to Yes-No Questions. *Linguistic Inquiry* 2.69-82. 1971.
- Quirk, R., S. Greenbaum, G.N. Leech, and J. Svartvik. *A Grammar of Contemporary English*. London 1972.
- Robinson, W.P. *Language and Social Behavior*. Harmondsworth 1972.
- Sebeok, T.A., A. Hayes and M.C. Bateson (eds.). *Approaches to Semiotics*. The Hague 1964.
- Seligman, C.R., G.R. Tucker and W.E. Lambert. *The Effects of Speech Style on Teachers' Attitudes Toward Pupils*. McGill University mimeo 1970.
- Stockwell, R.P. The Place of Intonation in a Generative Grammar of English. *Language* 36.360-367. 1960.
- . *The Role of Intonation: Reconsiderations and Other Considerations*. D.L. Bolinger (ed.), *Intonation*, 87-109. Harmondsworth 1972.
- Weir, R. *Language in the Crib*. The Hague 1962.

DISCUSSION

Robert Austerlitz: This approach makes me, at least feel less intimidated about undertaking prosodic analysis than I ever felt before.

Do you take for granted a complete (or incomplete, or no) grammatical analysis before you assign your prosodic values?

A speculative question: have you ever begun the analysis by assigning prosodic values first and only then grafting informational units upon them? Is there any value in this?

Claes-Christian Elert: One way in which prosody of the Scandinavian languages differs from that of British English seems to be that the relevant units are shorter. This may be the background for the contrast between the long-range intonation changes in your reading of some examples and the undulating intonation in a comparable Swedish example: *Jag tror att han har nå't att komma med där*.

Göran Hammarström: In what I call PROSODEMES (and also in what I call CONTOUREMES) I usually believe that four components should be distinguished: (a) pitch course, (b) loudness course, (c) quality course, (d) length course. In his talk Prof. Crystal dealt with something like component (a) and he related linguistic functions to 'tone units'. I would prefer, however, to provide each prosodeme as a whole (be it a syllabeme, lexeme or syntagmeme prosodeme) with a function. (See my book *Linguistische Einheiten*, 1966:7-13, 33-37, and my paper 'On linguistic terminology', *Actes du X^e congrès international des linguistes*, I:321-325. Bucarest 1969.)

Stig Eliasson: In presenting your prosodic framework, you did not state how you conceive of the structural dependencies that hold, in many languages, between prosodic and non-prosodic facts. Would you consider so-called determined prosodic factors (such as—for the most part—stress, tonality, and quantity in Swedish) as DERIVED on the basis of non-prosodic factors WITHOUT there being in the grammar ANY PROVISION FOR RECOVERING the conditioning elements (whether in part or in full)? If so, this raises a very fundamental question about the prosody of Swedish and similar languages, namely, why should these languages have a conventional prosody at all? In a unidirectional phonological framework like the generative one, which regards much of prosody as completely predictable,

there should to this same extent be no structural need whatever for it; indeed a paradoxical situation.

David Crystal: To Austerlitz: The prosodic and grammatical analyses have always been carried out independently. The prosodic values are assigned on the basis of auditory agreement that some change (of pitch, loudness, etc.) has taken place. The assumption is that change is significant unless it can be shown to be otherwise, and so in our complete prosodic description, some very narrow sound features are transcribed. We do not however assume that all changes are equally significant: one of the main principles underlying this approach is that there are varying kinds and degrees of linguistic contrastivity, and this is one motivation for seeing non-segmental phonology as a set of prosodic systems. Only in a very small number of cases, where there is lack of auditory agreement, do we need to consciously refer to our grammatical knowledge in order to make a prosodic decision (e.g. as to where a tone-unit boundary should go), and these cases have been ignored in the above presentation. The grammatical analysis of the data was carried out separately (I used the Quirk grammar, as this is the one I know best, but any model might have been used). The aim was thus to determine whether the correlations between prosody and grammatical analysis were sufficiently systematic to motivate a theory as to the relationship between the two. The investigation shows that they are highly systematic; but of course it is still an open question as to the best way of incorporating these facts into a linguistic theory.

To Elert: This may be so. But it is worth pointing out that some kind of shorter unit needs to be specified in order to handle the Swedish examples too, otherwise you will end up with an unmanageably large range of intonation-cum-rhythm patterns in your description. By aiming to keep the basic units of intonation relatively short, a more economical description can be obtained. Concatenations of tone units, then, when pronounced in distinctively rhythmic ways, can be described with reference to a system of specifically rhythmic variations

(which is what I do for English).

To Hammarström: I do not think there is a very great difference between us, in respect of the issue you raise, despite the terminology. I too work in terms of these four variables. My notion of tone-unit is defined primarily, but not solely, with reference to pitch; and I do not state that all linguistic functions are ascribable to tone-units. The tone-unit has one range of functions, tonicity has another, pitch range another, and so on. Where we differ is that I do not think there is anything to be gained by insisting that all these kinds of linguistic contrastivity are 'emic', in any single, clear sense.

To Eliasson: It seems to me that part of the paradox is due to the use of a grammatical model in which specific notions of derivation and recovery are required that may be incompatible with the idea of prosodic-grammatical relationship. There are after all many types of relationship: for example, a grammatical category may be wholly or partially realised by the prosody; its morphological exponence may co-occur with an obligatory prosodic pattern, or there may be no more than a tendency towards co-occurrence. The English data illustrates all these types, and thus it is quite unclear what the best way might be of stating all these relationships within a theoretical grammar. It may be that some kind of polysystemic statement will be more helpful.

The following members of the Conference also contributed to the discussion: Eva Gårding, Siiri Karlsson, William Labov.