THE LINGUISTIC STATUS OF PROSODIC AND PARALINGUISTIC FEATURES

By D. CRYSTAL

In 1964, D. L. Bolinger published a useful little article in the Harvard Educational Review, and gave it the partial title "Around the Edge of Language"; but his gloss for this, 'Intonation', implied a state of affairs which (in academic circles at any rate) was a few years behind the times. Those days are long gone, one trusts, when phenomena such as intonation and stress were deemed 'additions' to language proper—'super' fixes or 'supra' segmentals of marginal status and interest. Most linguists these days see intonational information, its form and function having been better defined, as relevant for the main part of any linguistic statement: it is no longer on the edge of language.[1] However, intonation has also been pushed further into the linguistic fold by default, as it were, for there has been a better contender for the title of 'linguistic phenomenon of most peripheral interest'. Over recent years, linguists and others have been studying a range of vocal effects (using largely linguistic techniques of analysis and description) whose uncertain status and definition are such as to make intonation appear quite safe as a clearly intra-linguistic phenomenon. These effects, generally called 'paralinguistic phenomena' since Trager's first stimulating description of some of them,[2] are now the focus of attention for a number of scholars, and their place within or without the discipline of linguistics is clearly something which merits detailed discussion. This paper, then, attempts to review some of the issues which are currently being discussed, and to ask some questions which seem relevant to determining the linguistic status of these phenomena.

The literature in the field is now substantial, reflecting the shift in interest from "*Are* there any vocal phenomena at all like that which we traditionally call language that the linguist should be interested in?" through "Just how much systematic vocal effect is there?" to "How many of the communicative systems accompanying these vocal effects should the linguist be interested in also?" The wide range of interest and subject-matter is well reflected in a recent book, *Approaches to Semiotics* (the Hague, 1964), the Transactions of the Indiana University Conference on Paralinguistics and Kinesics, held in 1962. This basically comprises five extremely well documented 'state of the art' papers, each given by a specialist in the principal disciplines represented—psychiatry, psychology, education, linguistics and cultural anthropology—and includes the discussion which followed each paper. 'Semiotics' was chosen as a cover term for 'patterned communications in all modalities' (5), to imply "the identification of a single body of subject-matter", to act as "a persuasive device to advance unified research", and "to stress the interactional and communicational context of the human use of signs and the way in which these are organized in transactional systems involving sight, hearing, touch, smell, taste".(5). For the linguist, the central question is how much of this patterned human communication can be legitimately called linguistic, and rigorously distinguished from all other communicative systems. While well aware of the relevance of the general concept of patterned communication to his work, he is also concerned that linguistics should not be identified with communication in all its modes. Where should the boundary-line be drawn (assuming, for the moment, that there should be one)? It is unfortunate, after so much debate, that this question still has to be asked: but up to now, answers to it have been either prejudged by too limiting a definition of what the proper subject-matter of linguistics is, or obscured by the tendency to use linguistic terminology, methods of analysis and description in other modalities without sufficient attention being paid to their originally fairly narrow usage. The argument of this paper is that there is more to linguistics than has hitherto been accepted in this field, but that this extension should not be carried too far.[3]

As the field is still little-known, outside of a fairly small inner ring of specialists, some general introduction to its subject-matter and terminology may be useful. Two main lines characterise the orientation of most recent work: first, there is an attempt to put language as studied by linguistics thoroughly in a communicational perspective; second, linguistic and non-linguistic in vocal effect are usually distinguished using a Trager-Smith approach and framework of analysis and description.[4] The former is really a bird's-eye-view approach to linguistics, relating language to everything else that may take place during an act of communication, and seeing language as only one component among many, all of which function together to produce the form and meaningfulness of the communication. Trager's statement is instructive, that "language, the principal mode of communication for human beings . . . . is always accompanied by other communication systems, that all culture is an interacting set of communications, and that communication as such results from and is a composite of all the specific communication systems as they occur in the total cultural complex". In other words, one begins with a concept of meaningful communication, and isolates the components, one of which may usefully be labelled linguistic. Throughout the literature, there is the implication that as all human behaviour is potentially relevant to understanding human communication, and hence language, therefore it is to be studied by the language-student. All the components may help to communicate a meaning; and it cannot be assumed that if any of these are ignored and attention restricted to microlinguistic form (cf. below) that one can study the totality of language's meaning successfully. In this connection, the assumptions which A. S. Hayes formulates in relation to kinesic study are generalisable to paralanguage, and are worth quoting *in extenso* as a theoretical perspective:[5]

> "1. Like other events in nature, no body movement or expression is without meaning in the context in which it appears.
>
> 2. Like other aspects of human behavior, body posture, movement, and facial expression are patterned and, thus, subject to systematic analysis.
>
> 3. While recognizing the possible limitations imposed by particular biological substrata, until otherwise demonstrated, the systematic body motion of the members of a community is considered a function of the social system to which the group belongs.
>
> 4. Visible body activity like audible acoustic activity systematically influences the behavior of other members of any particular group.
>
> 5. *Until otherwise demonstrated* (my italics) such behavior will be considered to have an investigable communicational function.
>
> 6. The meanings derived therefrom are functions both of the behavior and of the operations by which it is investigated.
>
> 7. The particular biological system and the special life experience of any individual will contribute idiosyncratic elements to his kinesic system, but the individual or symptomatic quality of these elements can only be assessed following the analysis of the larger system of which his is a part."

At least the following stages must be (and have been) distinguished in analysis:

I. The total personal and environmental physical setting in which an act of communication takes place. From the linguistic point of view, the interesting aspect of this is the idiosyncratic vocal background characteristic which has been called 'voice quality' or 'voice set', and which we intuitively discount after we have recognised it for what it is.[6] It has a mainly person-identifying function.

II. Against this background, acts of communication take place. These are 'bundles' of interacting events or non-events from different communicational sub-systems (or 'modalities') simultaneously transmitted and received. This communicative activity has been variously called a 'signal syndrome', a 'total configuration', and a 'communication network'. It is distinct from the physical background in that the latter is primarily physiologically-determined, idiosyncratic activity (in normal utterance) whereas the specific act of communication is always, in the present sense of the term, culturally-determined and conventional.[7] Questions of psychosomatically-motivated activity, semi-habitual responses, environmentally-controlled responses, and so on (in short, whether communication is intended or not), are obviously crucial in deciding the validity of this distinction, but these problems are too large to allow discussion here[8], nor are they relevant to the present topic, which concerns the status of vocal effects that are usually incontrovertably controllable and conventional.[9]

The act of communication, of course, cannot be studied without individual reference being made to its constitutive elements, so that the different functions and interrelationships of these elements may be better assessed.[10]. It is also expected that paralinguistic and kinesic phenomena, for example, are significant in themselves, apart from any combined function they may have as part of a 'communication network'. These components, all of which are variables, at the most general level are related to the five senses—vocal/auditory, visual, tactile, olfactory and taste. In theory, all these components are relevant for determining the full meaning of an act of communication; but for the linguist (and even, more generally, for the communication analyst—or 'semiotician'?), attention has been primarily centred on the vocal/auditory and visual components. These two are of central importance in discussing patterned communication of most kinds, because they involve a large number of physical variables which can be used for making contrasts, and are relatively easily controllable, transmittable and perceivable (as opposed to the others). Only these are frequently used in normal (or abnormal) communication—and consequently were the main subject-matter of the Indiana Conference. The main question is, "How far should one allow all this meaningful but non-vocal communicative activity to become part of the province of linguistics? The linguist, faced with this mass of new and potentially relevant data, and bearing in mind the newly-discovered complexity of more central linguistic areas currently being investigated, will naturally take a conservative position (cf. the reluctance of some linguists to allow even writing a place in linguistics). This is largely justifiable. To allow the whole of non-vocal communicative activity into linguistics (with the exception of writing, which may be distinguished from visual communication in general on such grounds as that it systematically reflects and does not accompany speech) would overburden any (already complex) linguistic theory; and there seems to be no internal structure in body motion, etc., comparable to that of language.[11] In short there is just too little in common. Again, non-vocal fields other than kinesics have received very little study, nor is knowledge about them likely to increase substantially in the near future—data is difficult to come by, and even in kinesics such questions as the boundary between conventional and idiosyncratic are only beginning to be studied. To require that the linguist treat these components along with language would be a retrogade step at present, only hindering his progress in understanding the latter. There is also the point that the consensus of opinion at the moment is against even prosodic and paralinguistic features being allowed into linguistics proper (other than features of pitch, stress and juncture when given a phonemic status)—hence a policy of caution would seem wise, despite a certain amount of theoretical motivation. Consequently there is no argument in this paper for the inclusion of non-vocal communicative activity under the heading of linguistics—only for more of the vocal.

III. Turning now to the vocal/auditory component, this may be broken down into the following categories:[12]

1. 'Segmental'. This is the traditional centre of linguistic attention, which in its widest definition would cover segmental phonetics and phonology, morphophonology, morphology, syntax, lexis, and, these days, semantics.

2.1. Vocalisations, *e.g.*, 'mhm,' 'shhh', which are really lexical items, and which may function grammatically in certain types of structure. They are quite distinct from, and should not be taken along with 3 and 4 below.

2.2 Hesitation. There is an unclear formal boundary between vocalisations and voiced hesitation ('hmmm' - - 'ə:m', etc.)[13], and a distributional and semantic overlap with voiceless hesitation.

3. 'Non-segmental'. These are aspects of the phonic continuum which have an essentially variable relationship to the words selected, as opposed to those features (*e.g.*, the phonetic distinctive features, the lexical meaning, etc., cf. 1 above) which have a direct and identifying relationship.

3.1. *Prosodic features :* variations in the parameters of frequency, amplitude and speed of utterance (auditorily pitch, loudness and tempo), which permanently characterise speech.

(i) Variations which are relatively easily integrated with other aspects of linguistic structure, particularly grammar and lexis—what are usually referred to as the grammatical and accentual functions of intonation (including such notions as 'contrastive stress', nuclear tone selection and placement within the tone-unit, placement of tone-unit boundaries and structural pause).[14]

(ii) Less discrete, attitudinal variations in intonation, involving variables in the sub-systems of tone, tempo, prominence, pitch range and rhythmicality.

3.2. *Tension :* this is a category which is defined with reference to characteristics of both prosodic and paralinguistic features, and which is therefore most satisfactorily treated as neither. It includes 'slurred', 'lax', 'tense' and 'precise' articulations.

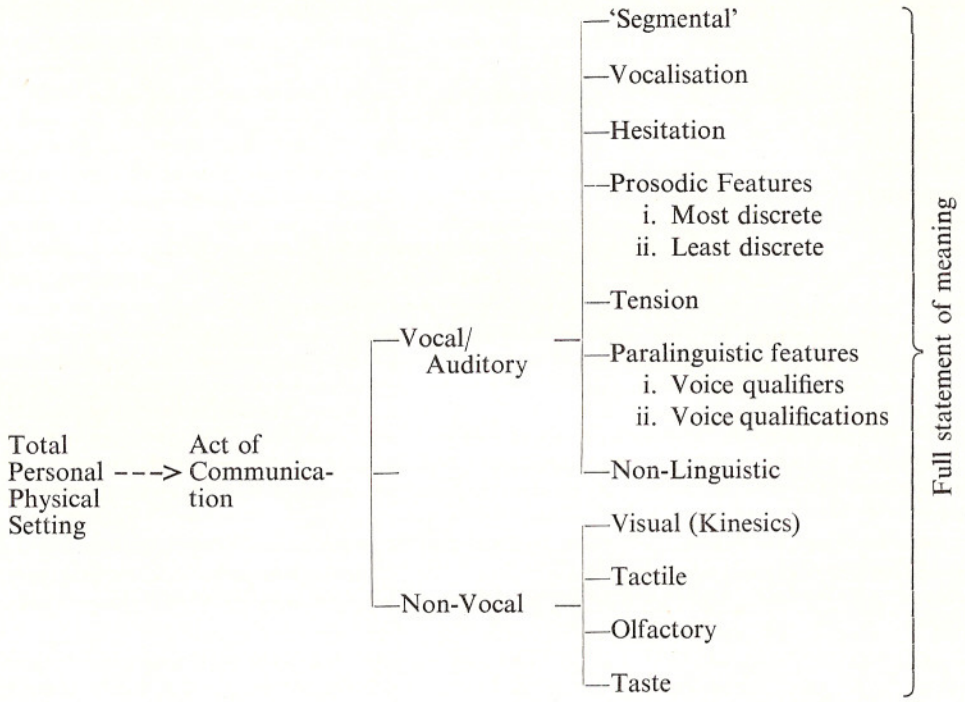
3.3. *Paralinguistic features :* discontinuous vocal effect which is primarily identified with reference to the working of physiological mechanisms in the pharyngeal, oral and nasal cavities other than the vocal cord activity which is of primary importance in the identification of prosodic features. Prosodic features (being permanent features of utterance) of course enter into these effects, but are variables in respect of their definition.

(i) voice qualifiers: whispery, breathy, husky, creaky, falsetto and resonant voice.

(ii) voice qualifications: laugh, giggle, tremulousness, sob and cry.[15]

4. Non-linguistic personal noises: physiological reflexes such as coughs, sneezes, snores, heavy breathing, etc.

The stages of the communication situation as a whole may now be conveniently summarised in the following diagram: the study of the totality is semiotics. (The diagram, being two-dimensional, is of course misleading, but it does suggest the movement of linguistic interest, which is from the top downwards.)



The framework in most universal use that attempts to cope with the problems associated with the status of prosodic and paralinguistic features is that devised by Trager, and others, in the works already mentioned. As is well-known, Trager cuts up the object of study into a central area and peripheral fields. Basically, the field as a whole, 'macrolinguistics', is divided into 'prelinguistics', 'microlinguistics' and 'metalinguistics'. To get these distinctions clear is an important step in assessing the arguments surrounding the status of paralanguage, in which field they are still generally taken as fundamental, though they are not currently held by many scholars elsewhere.[16]

(i) Microlinguistics is glossed by Trager and Smith as 'linguistics proper,' which is glossed in turn as phonology and morphology—and syntax would also seem to fall under this heading.

(ii) Prelinguistics studies language sounds considered simply as noise—the why and how of their production (physio-linguistics, or articulatory phonetics, and physico-linguistics, or acoustic phonetics). This includes all sound, whether used for a contrastive purpose in language or not: it therefore also comprehends background voice quality, or 'voice set'.

(iii) Metalinguistics[17] is glossed as "the overall relation of the linguistic system" (*i.e.*, phonology and morphology) "to the other systems of the cultural totality"; or again, "the full statement of the point-by-point and pattern-by-pattern relation between the language and any of the other cultural systems will contain all the 'meanings' of the linguistic forms, and will constitute the metalinguistics of that culture" (*The Field of Linguistics*, p. 7). Trager and Smith are not wholly clear as to just how much of human behaviour and experience this covers: clearly, the relationship between language and literature would fall in here,[18] as would the relationship between speech and writing, psycholinguistics, the relation

between language and society (sociolinguistics, including stylistics in general), and the relation between language and reality—perhaps even non-linguistic reality itself.[19] But Trager and Smith also say that the study of metalinguistic phonology and morphemics covers such matters as value judgments about language, and ideas about status, class, and so on, which would be covered by pragmatics. They sum up metalinguistics as a guide to part of the overt culture (presumably semantics) and to all of the covert structure (or 'sentiment structure'). Paralanguage, in Trager's sense, is covered by this, and is thus outside of 'linguistics proper'.[20]

For Trager and Smith, then, the only valid subject-matter of linguistics is the analysis of language systems, of which micro-linguistics is linguistics proper, and metalinguistics is a secondary concern relating the linguistic form to experience. Phonetics is excluded, on the grounds that the raw data is not systematic. On the surface, the field is cut up in a similar way to that delimited in another, more recent book:[21]

Phonic substance	organised into units of form,	contextualised in	(non-linguistic) situations
Prelinguistics	Microlinguistics	Metalinguistics	
	Macrolinguistics		

One can see these two recent trends as being to a large extent complementary: the first approaches language through the concept of the communication situation, by analysing it into components, one of which may be legitimately called language; the second starts with a fairly rigid and explicit view of what language is (phonetics and morphemics), and then looks elsewhere in human behaviour for analogous phenomena, assessing them against this criterion. In fact, the two views are not entirely compatible; a certain tension exists between them, which comes clearly to the fore when one examines the linguistic status of certain vocal (and non-vocal) features which are not of a phonetic or morphemic character. We may now look more closely at one group of these, the prosodic and paralinguistic features in the above classification.

The fundamental assumption, held by all the scholars interested in this field, is that communicational activity other than the inner segmental-hierarchical core of language is also structured; and most research in this field has been to determine the extent of the structuring and the systematicness. I am not here concerned with the structuring of kinesics, in terms of kines, kinemes, kinemorphs, allokines, etc.[22], but only with the question, "How far can the prosodic and paralinguistic features outlined above be seen as comparable in form and function to language, traditionally-called, and hence a relevant concern of linguistics proper?" And, if there is a scale of 'linguisticness' from completely intra-linguistic to completely non-linguistic (cf. below), then where can one most usefully draw a boundary-line?

The first point which needs to be discussed is that in much of the work which has so far been done, many scholars have accepted as fundamental the Trager-Smith view of language (hence paralanguage) uncritically, even though this has been shown to have important weaknesses and is little followed by linguists in other fields at present.[23] Its central stimulus to the development of paralinguistic studies and semiotics in general, of course, no-one would wish to deny; but a certain complacency seems to have developed around the first analyses which can ultimately only be harmful to progress.[24] It is after all the case that most linguists would criticise any microlinguistic view of language proper which restricts one's view of other phenomena, and forces one to make a binary 'yes/no' decision as to the status of certain features, where perhaps there is none, or where the

analysis is more of a 'more/less' kind[25]. Stankiewicz is quite clear on this point in the discussion to his paper in *Approaches to Semiotics*° "It seems futile to approach language as a monolith and to exclude from it those thenomena which are not cut to the same pattern" (266) . . . . "The linguistic status of certain features which signal emotion cannot easily be decided, since they do not lend themselves to the kind of systematization that linguists are used to, nor are they easily correlated with other linguistic elements" (267). He suggests that there may be a "fuzzy periphery" to language, which scholars are now describing more consistently and in "somewhat different terms from the cognitive or discrete elements of language"; but leaves as open questions "the systematicity of the expressive devices and their relationship to the cognitive elements of the code" and the "grading" of expressive components, as opposed to the "discreteness" of phonemes and morphemes (267). While Stankiewicz does not use the term 'paralanguage' in his article, talking more generally in terms of 'emotive language', he does admit to both narrowing and broadening the frame of reference of paralinguistics in this way (272). He suggests that Trager's system may have been stretched too widely. Elements which are retained by Stankiewicz are elements which he considers "a part of the linguistic code, as opposed to those things which have emotive connotations due to the social situation or to the verbal context, or because they are physiological . . . . Everything occurring in a social context can can be endowed with meaning, everything can carry certain overtones, but I try to separate these from elements which have a definite, fixed value in the linguistic system. Coughing, yawning, belching, or whispering and yelling are not part of the linguistic system". His concept of linguistic system here is clearly much narrower than that of many other scholars; he is against extreme permissiveness in defining the subject-matter of linguistics. But his exclusion of 'whisper' and 'yelling' (this being a 'bundle' of prosodic parameters) suggests that his approach is a little too narrow. His main point, however, is valuable, and it is made by others (cf. Lotz's remarks in *Approaches to Semiotics*, pp. 177-8). Thus if only phonemes and morphemes (and syntax?) are linguistics proper, then all else by definition, being judged by these standards, is excluded as non-linguistic; and one is forced to see intonation, for example (if one wants to take it as within language proper) as either phonemic or morphemic in character, There is no other choice—apart, that is, from saying that it is non-linguistic, which most linguists would not wish to do. Now while some case has been made to impose phonemic and morphemic patterns on intonational phenomena (though recent research would suggest that the issue has been largely prejudged[26]), it is a much more difficult matter to force the whole range of prosodic and paralinguistic patterns into this mould, and retain any useful and reasonably homogeneous definition for the terms phoneme and morpheme. Hence the tension: if one wants to include some or all prosodic and paralinguistic phenomena in language (and pitch, stress, and juncture are usually felt to be an essential minimum in this respect), and one does not wish to force them into a microlinguistic framework, then one has no alternative but to revise the system of description and transcription accordingly. Before discussing the reasons for adopting this solution in more detail, however, it is worth referring to three other approaches to the problem which try to reduce this tension.

The first way out is to shelve the question, by calling the features 'extralinguistic': there are important differences between paralanguage and language, though not enough to make a total break, as these features may have a structure that would bring them closer to language, but this has not been ascertained. Mahl and Schulze characterise this position (*Approaches to Semiotics*, pp. 111, ff.): indications of emotion may be culture-orientated and not merely idiosyncratic physiological reflexes, and these may be gradable and transcribable, *e.g.*, there *may* be a shared rate of speeding-up of utterance which is a partial indication of (say) impatience or anxiety[27]; and similarly with all other vocal effects of this type. There is no point in assuming otherwise, that such features are non-linguistic. But as insufficient work has been done to ascertain the extent of this shared systemicness, they cannot be called linguistic as yet, hence 'extra-linguistic' (usually in inverted commas).

This is only a temporary label, then, for the linguist: when evidence emerges that there are codes and grammars in these phenomena, the area will be redefined as a "special field of linguistics; it won't be called 'extra-linguistics' any more" (112).

This terminological 'solution' of course does not get round the central problem at all. Another way out is to see these features as mere modifications of linguistic items (*i.e.*, denying tham any status in their own right), *e.g.*, one section of the contract supporting the Indiana Conference read, "The speaker is free to choose his message. He is not free to choose the code of the message—this is strictly imposed by the language . . . . The speaker is, however, free to *color* his message in certain ways, and these ways are predominantly paralinguistic and kinesic."[28] This is very close to Trager, who puts many effects which he calls 'tone of voice' in stylistics. But this, too, is unsatisfactory. Why should it be the case that the emotive element in meaning is usually taken as secondary in this way, as is implied by such terms as 'colouring' the message? What if the purpose of a particular act of communication was primarily emotive, which involved the use of prosodic or paralinguistic features of a certain kind?[29] This resolves into a funadmental question: "how can we assume that in any total meaning complex, one element is primary (*e.g.*, word meaning), others secondary, and rank their relative status accordingly?" As Birdwhistell has said (in a paper called "The frames in the communication process"): "It is all too easy to assume that there is in any social interchange a *central*, a *primary* or a *real* meaning which is only modified by a redundant surround . . . . Our temptation so to classify certain aspects of a transaction as the central message and other aspects as serving only as modifiers rests upon untested assumptions about communication."[30] Certainly in many actual interpersonal speech situations, the 'expected' priorities are reversed: the 'verbal' side of language becomes subordinate to prosodic and paralinguistic features (which in turn may be subordinate to kinesics).[31] To judge emotional language, then, by a criterion more suited to non-emotional language, distorts the former badly.

A third way out is to give paralanguage an independent status, but as part of semantics, denying it any formal relevance, This is to avoid one trap, but to fall into another. The first trap was to concentrate excessively on form, as opposed to meaning, by restricting attention to phonemes and morphemes only. Numerous statements on the relevance of meaning in linguistics have produced a more balanced view[32]; but if this is so, then *all* the elements which contribute to the overall meaning of a linguistic item (seen in a particular context) are relevant for study, otherwise there will be an inadequate statement of that item's range of meaning. These elements, of course, include prosodic and paralinguistic features, and without reference to these, too much meaning may be left out. From the point of view of the contextualisation of language form, these features are of a comparable status to other meaning carriers: throw doubt on their status, and one impugns the whole of the linguistic basis of semantics. However, it does not follow from all this that these features should be judged by semantic criteria only, and listed with other semantic statements, as with Trager-Smith, for example. These features have more than semantic relevance, having a great deal in common with other levels of language organisation—as is generally recognised in the instance of intonation (which for some reason is always given a separate place in description, though clearly there is only a difference of degree between intonation and overhigh/low pitch height, for example, as far as attitudes are concerned). Prosodic and paralinguistic features may be identified in language form and substance, in terms of their phonetic shape (degrees of discreteness), structural contrasts (and degrees of grading here, cf. below), their ability to work in sequence and simultaneously, their integrability with grammar, and other distributional characteristics. It is even possible to see in these effects some of the attributes of a closed system: this side of language is certainly unproductive, and one may discuss the prosodic sub-systems outlined above, for example, in terms of both positive and negative 'meanings'. Without going into this further, it should be clear that there is a great deal to be said about these features other than in semantics, and hence they cannot be taken as elements of purely semantic force.

To see these features as modifiers, extralinguistic, or purely semantic, is thus not very helpful, and in the absence of further answers, it suggests that the question itself may not have been well framed, and that there has been too much concern over what is after all a pseudo-problem. The three approaches just referred to have largely in fact assumed the phoneme/morpheme basis of linguistics referred to earlier. By why *should* prosodic and paralinguistic patterning be judged by phonemic or morphemic criteria? The valuable discreteness of phonemic and morphemic definition, with their amenability to clear-cut substitutability tests of an either-or character, usually free from overlap, does not exclude the fact that there are parts of language which are not amenable to such treatment, but are more accurately and realistically covered by setting up scales of contrastivity, which are of a 'more/less' character.[33] If these scales have been accepted in grammar, lexis, and even intonation at times, then why not here also, for the whole range of vocal non-segmental effect? Much more data is satisfactorily accounted for if one postulates a scale of linguisticness for these features: at its most linguistic end are placed items which have a fairly well-defined contrastivity (*e.g.*, tone-unit boundary, nuclear tone type and placement, type of head and prehead, certain types of tone-unit sequence, and so on); at the other, least linguistic end, items which have less discreteness of formal definition but a nonetheless definable contrastivity (*e.g.*, breathiness, creak). Degrees of contrastivity could be determined using techniques similar to that exemplified elsewhere, in terms of perceptibility and replicability by native speakers, or in terms of their integrability with other linguistic levels, especially grammar.[34] The boundary with non-linguisticality would come at the latter end, excluding coughs, sneezes, snores, and background voice qualities, and prosodic and paralinguistic features would be spread over the linguistic scale. There seems to be no real reason why these features should be excluded from the field of linguistics proper, and why they should not be taken as wholly intra-linguistic items, but of a different kind from phonemic and morphemic units, as they are normally understood. Certainly, whichever criteria one sets up to characterise the notion 'linguistic item' or 'language proper', one finds near-identity, *e.g.*, both language and prosodic and paralinguistic features are vocal, consist of minimal phonetically definable units (with a little peripheral 'fuzziness' in the latter's case which awaits further study)[35], are systematic (cf. below), contrastive, meaningful, and to a large extent hierarchical.[36] Prosodic and paralinguistic features also lack productivity, but in this respect are no different from phonology (segmental) and parts of grammar; they are also said to be ambiguous and difficult to define—again, characteristics which no-one has ever suggested language as such lacks! From all this, it would seem that these features may at most be different in degree from language: perhaps some such progression as "grammar——'grammatical' function of intonation —— 'accentual' function of intonation —— 'attitudinal' function of intonation —— prosodic features other than tone —— tension —— paralinguistic features —— non-linguistic noises" shows the interrelationship better.[37]

There are a number of statements hinting at this conclusion in the literature, *e.g.*, Stankiewicz above, and Markel: "Since you can negate a message with a paralinguistic signal, paralanguage is a code. The only way you can negate one code is to use another code that supersedes it" (*Approaches to Semiotics*, p. 133). Now while this is too wide for linguistics, as the term 'code' includes any pattern to which meaning has been conventionally attached (*i.e.*, semiotics in its widest sense), the statement does apply to prosodic and paralinguistic features, and does indicate that they are normally unambiguously contrastive. It would be wrong to suggest that *all* such signals function as clearly as this negating function, but there *is* a large central area for which this systematicness exists. The general impression of course is to the contrary, but to think that such features are fluid and formless is a misconception which seems to derive solely from the lack of research and published material on the subject. This is unfortunate, for it is not the case that prosodic and paralinguistic features are on the whole ambiguous and indefinable. Even intonation is usually mistreated in this way. It is often said, for example, that there is no one-to-one relationship

between form and function in intonation[38]—that it is impossible to define the meaning of the nuclear tone, low rise, for instance, because it can mean any number of things. This is a nonsensical criticism: the low rise should not be discussed in this way. Any instance of a low rising tone has one meaning only *in one context.* One might just as well argue against the word 'table' on the grounds that it is ambiguous. Of course it is, until one takes it in context: to complain about the ambiguity of the term in isolation is unrealistic. Similarly with intonation, where features may usefully be discussed in terms of homophones.

This is not to deny that there are difficulties of precise definition in the field, but this is only a difference in degree from the rest of language. There is a flexibility, of the type "how far does a low fall fall?"; there are difficulties of perception, in particular of separating segmental (*e.g.*, vowel) quality from voice quality, and of obtaining agreement from linguists as to what they hear; and there is also the presence of a certain amount of irregularly occurring formal overlap: but all these troubles turn up elsewhere in language, and such redundancies seem to have no effect on the range of the features contrastiveness—which is just what one would expect. On the whole, prosodic and paralinguistic features in English have a definite, fixed value in the linguistic system (*i.e.*, including whispering and yelling), so long as the items are viewed carefully in their contexts of occurrence. The fact that no-one has yet defined the whole range of contexts clearly, or measured the formal patterns with sufficient accuracy is no ground for relegating the features concerned to non- or extra-linguistic status. Other, more generally accepted linguistic patterns have the same problems. All the evidence, in fact, points the other way: the more one studies the field, the more one finds readily-definable correlations, and an unambiguous, precise formal-functional relationship. So far, the detailed evidence which scholars have uncovered, largely based on observation of stimulus and response, and empirical testing of the use of effects in contexts, obtaining definitions of effects from informants using Osgoodian-techniques,[39] etc., have shown a substantial amount of precision as long as enough contextual information was allowed in. This is what one would expect in principle: if the things are *not* systematic and unambiguous, then how do we ever manage to understand them? From another point of view, experimental research in speech synthesis has shown already that there is a definable acoustic basis for particular auditory effects. (Catford is one who has tried to define an articulatory basis.[40]) Laver, for example, has synthesised voice qualities on PAT, based on Catford's descriptions, most of which are well recognisable, and interpretable in terms of a fairly small number of lexical items.[41] This side, of course, is extremely important, and ultimately the only way one can prove what one is saying in this field: however, a synthesiser with sufficient parameters to cope with paralinguistic complexes is difficult to come by, and research at present is presumably going to be restricted to testing auditory judgments.

This brings me to the most important problem hindering the development of this subject, which is in effect a problem of labels—the way in which informants give linguistic reactions to prosodic and paralinguistic effects. It is the case that the majority of problems associated with the prosodic and paralinguistic field are lexical in nature, and critical attention seems to have been distracted by the labels. It is not prosodic and paralinguistic featres which are amorphous and ambiguous in definition, but the descriptive labels which we try to pin on them which are usually too indefinite and semantically loaded.[42] The range of prosodic and paralinguistic features is fairly readily imitable and recognisable: the vagueness surrounding the definition of a feature is rarely due to our inability to understand its function—we are normally quite clear as to the meaning of a particular form in conversation, and we react accordingly. Only when we bring in the labels, and try to delimit the effect lexically does the trouble start; and the point is, that we should recognise it for trouble *at this level.* There is an almost complete one-to-one relationship between behavioural stimulus and response in practice; but there are many near-synonyms which can be applied to one vocal effect, each of which has a different range of sense-association;

and if one is not careful, the extra sense-associations from the term become attached to the vocal effect itself. It may be of interest in this connection to mention a series of experiments currently in progress, which are aiming to plot this indeterminacy. Information about the meaning of prosodic and paralinguistic features will be elicited in two ways: first, by playing a vocal effect to a native speaker, and getting him to define it, *i.e.*, one determines the total lexical range associated with a particular pattern. This is the main way to get round the bogey of definition, by using statistically graded results, plotting uncertain areas, etc. A second way will be to ask native speakers to produce 'grim' utterances, for example, giving them a written sentence and a context, recording the results, and then plotting recurrent patterns, This is a more difficult test to carry out, as reactions have to be elicited individually, but it is an important control on the results of the first experiment. In these two ways it is hoped to minimise lexical uncertainty in labelling attitudes, and thus approach such thorny areas as the semantics of intonation more satisfactorily.[43]

There is one further problem which needs to be mentioned, albeit briefly. There is an area of uncertainty which sometimes makes it difficult for a listener to distinguish the individual's voice set from the conventional prosodic and paralinguistic features of the speech community. Again, it is not a question of being able to see any distinction at all, but of where to draw the line. In the case of paralinguistic features, it is usually easy to distinguish voice set from convention, because of their relatively infrequent occurrence. If such a feature was part of a person's voice set, it would be immediately noticed and allowed for. But there is more of a difficulty with prosodic features, which are always present in speech to some degree. If, for example, I habitually use a falling-rising nuclear tone where other people would normally use a rising tone of some kind, then how long does it take for others to 'tune in' to this relatively subtle idiosyncracy and discount it, so that they realise I am not intending the range of meaning normally associated with the falling-rising tone in that context? And are there idiosyncracies which cause ambiguities which people never tune in to at all? Again, how shared must an item be to be allowed in as a norm? To answer these questions, detailed statistical information, based on the examination of a large amount of connected speech in every day situations, is needed. This is accumulating, but until it is available, it is as well to be aware of this further area of uncertainty in this field.

The Indiana Conference which I have repeatedly referred to certainly did not avoid facing up to problems of this kind: the need for research of a more wide-ranging nature, involving inter-disciplinary activity, is referred to again and again (cf. pp. 49, 112). Certainly such problems should not put anyone off, in view of the great theoretical and pedagogical interest in this research. What needs to be done now is analysis of as much connected speech from different contexts as possible, reducing the data to as many classes as are possible and justifiable from this point of view. But in carrying this out, one must always remember the narrow path along which one has to walk, to avoid impressive theoretical generalisation on the one hand (and an over-wide linguistics as a result, a semiotics), and overminute description on the other (which leads to hyperdelicate anthropological or psychiatric study). A balance between the two may be difficult to achieve, but it is prerequisite if prosodic and paralinguistic features as a whole are ever to become thoroughly acclimatised to linguistics.

REFERENCES

1. But cf. A. Martinet's position, in *Elements of General Linguistics* (London: 1964), p. 93 : "We may regard them (*sc.* prosodic signs, including intonation, cf. p. 76) as marginal phenomena because an utterance is properly linguistic in so far as it is doubly articulated".

2. G. L. Trager, "Paralanguage: a first approximation", *Studies in Linguistics*, 13 (1958), 1-12.

3. As background to the present discussion, see D. Crystal, "A perspective for paralanguage", *le maître phonétique* (1963), 120, pp. 25-9; G. L. Trager, "Paralanguage and other things", *le maître phonétique* (1964), 122, pp. 21-3; D. Crystal, "An approach to a reply", *le maître phonétique* (1964), 122, pp. 23-4.

4. See G. L. Trager and H. L. Smith, Jr., *An Outline of English Structure* (Norman, Okla.: Battenburg Press: 1951); also G. L. Trager, *The field of linguistics* (1949, Studies in Linguistics: Occasional Papers No. 1); and Trager, *op. cit.*

5. "Paralinguistics and kinesics: pedagogical perspectives", in *Approaches to Semiotics*, p. 158.

6. Unless, of course, the point of the utterance is that we *should* pay attention to it, as in mimicry or medical diagnosis.

7. A number of terminological pairs have been used to distinguish physical setting from communicative act in this way, most trying to avoid the unwanted overtones of 'deliberate/non-deliberate' *e.g.*, 'idiosyncratic/shared', 'uncontrolled/determined behaviour/controlled/behaviour', 'non-institutionalised/institutionalised'.

8. They are given full discussion in the sessions following the main papers at the Indiana Conference.

9. Which is not to say that ambiguities over the identification of prosodic and paralinguistic features do not arise, cf. D. Crystal and R. Quirk, *Systems of Prosodic and Paralinguistic Features in English* (the Hague: 1964), pp. 33-4, cf. also below, p.00.

10. Cf. Stankiewicz's viewpoint—*Approaches to Semiotics*, p. 265, ff.

11. Cf. Hayes, *Approaches to Semiotics*, p. 153; C. F. Hockett, "The origin of speech", *Scientific American* (Sept. 1960), where characterising features of human language are discussed and compared to other systems.

12. Not all of these categories are set up by the American scholars already mentioned. The classification here is in fact that explained in detail in Crystal & Quirk, *op. cit.*, and is so far based only on an examination of British English.

13. In English, types of voiced hestiation are more varied than is usually maintained: apart from ə(:)(m/n(:)), to take just two examples, there is a tense, unaspirated voiceless bilabial (often trilled), and I have frequently heard in informal conversation a voiced lingual 'inter-labial' trill, both functioning clearly as 'hesitation noises'.

14. Cf. p. 00, and R. Quirk, *et al*, "Studies in the correspondence of prosodic to grammatical features in English", *Proceedings of the Ninth International Congress of Linguists* 1962 (the Hague: 1964).

15. This is therefore a fairly narrow definition of the term 'paralanguage'. Trager & Smith's term covers

some of what I call prosodic features and some of the above paralinguistic features, but not all in each case. Others of the above they consider in a different connection (as stylistics, headed 'distortion', though the status of this is unclear). Paralanguage is also not being defined in the original sense of Welmers, covering the non-linguistic as well as the linguistic function of pitch, intensity and duration (see "Non-segmental elements in foreign-language learning", in *Report of the 5th Annual Round Table Meeting on Linguistics and Language Teaching*, ed. H. J. Mueller (Washington: 1954), 130-6), nor in Hill's sense, to include kinesics (cf. A. A. Hill, *Introduction to Linguistic Structures* (New York: 1958), pp. 408, f.) There is quite a bit of uncertainty over the term in the literature: different authors have slightly differing lists of features. Most distinguish paralanguage from intonation, for example, but cf. Ostwald, *Approaches to Semiotics*, p. 17, who talks of "paralinguistic acoustic cues, *e.g.*, variously intoned forms of 'oh' and the nuances which support or belie overt meanings of words". It should also be noted that the term 'intonation' is being used in this paper as a general label for a set of vocal effects caused by the combination of a number of prosodic variables; and that the term 'tone of voice' is no more than an *ad hoc* label, used from time to time for economy of reference.

16. *E.g.*, semantics is these days generally considered to be wholly intra-linguistic.

17. The term 'metalanguage' is of course different here from its use in logic, or indeed by linguists in their more recent work, where it refers to language used to talk about language. Other names for parts of this field have been proposed, *e.g.*, 'ethno/socio-linguistics', 'exolinguistics'. In Britain, the tendency has been to distinguish semantics from institutional linguistics, and both these from ethnolinguistics.

18. Cf. Hill's microliterary study, in "An analysis of *The Windhover*: an experiment in structural method", *PMLA*, 70 (1955), 972-3, and elsewhere.

19. Cf. A. A. Hill, "Linguistics since Bloomfield", *Quarterly Journal of Speech*, 41 (October 1955), 253-60, reprinted in *Readings in Applied English Linguistics*, ed. H. B. Allen (New York: 1958), 14-23, see especially p. 18, f.

20. "The very essence of the analysis (of paralanguage) is in the recognition that paralanguage is **not** a part of language, and that therefore the statements about it are **not** linguistic statements" (Trager, "Paralanguage and other things", p. 23.)

21. M. A. K. Halliday, A. McIntosh, P. Strevens, *The Linguistic Sciences and Language Teaching* (London: 1964).

22. For which see the useful survey by Hayes in *Approaches to Semiotics*, esp. pp. 157, ff.

23. For an early review, see J. Sledd, *Language*, 31 (1955), 312-335.

24. *E.g.*, Markel, *Approaches to Semiotics*, p. 141: "I think we will find, if we go over the list, that everything that we want to call paralanguage is covered by Trager's system. We would also find those things that are not covered are not paralanguage; they are psycholinguistics or metalinguistics . . . ." Trager himself takes a more reasonable view, and is fully aware of insufficiency in his system, but nonetheless maintains that "as a guide to observation, the system presented by me works on all kinds of materials" (Paralanguage and other things", p. 22). It is doubtful whether the differences between his system and the one outlined here can be wholly accounted for by the difference between American and British English, however.

25. Cf. M. A. K. Halliday, "Categories of the theory of grammar", *Word*, 17 (1961), pp. 252, 275; Crystal & Quirk, *op. cit.*, p. 36, and below, p. 16.

26. P. Lieberman, "On the acoustic basis of the perception of intonation by linguists", *Word*, 21 (1965), 40-54.

27. Cf. C. F. Hockett, R. E. Pittenger and J. Danehy, "*The first five minutes*", *A sample of microscopic interview analysis* (Ithaca: 1960), who survey some of the conditions under which a speaker may utter a passage unusually fast. Hockett had earlier set up two tempo phonemes, each constituting a single morpheme, in a discussion of "alternations conditioned by tempo" in Chinese ("Peiping morphophonemics", *Language*, 26 (1950), 63-85).

28. Cf. Hill, "Linguistics since Bloomfield", *op. cit.*, who talks of voice qualifiers 'modifying' meaning and thus contributing to it (p. 20).

29. Cf. Hendrix, *Approaches to Semiotics*, p. 128, "frequently paralinguistic and kinesic behavour can express the presence of the highest intellectual attainment that an individual has arrived at, at that moment".

30. Paper presented to the American Society of Clinical Hypnosis (1959), quoted by A. S. Hayes in *Approaches to Semiotics*, p. 161.

31. Cf. the oft-quoted "It's not what he said, but the way that he said it", and like remarks, which testify to this.

32. Cf. J. R. Firth, "The main concern of descriptive linguistics is to make statements of meaning" (in "Modes of meaning", *Essays and Studies* (1951), p. 118.

33. Cf. M. A. K. Halliday, "The tones of English", *Archivum Linguisticum*, 15 (1964), p. 2; R. Quirk, "Descriptive statement and serial relationship", *Language*, 41 (1965), pp. 205-17; E. Uldall, "Attitudinal meanings conveyed by intonation contours", *Language and Speech* (1960), pp. 223-234; D. L. Bolinger, *Generality, gradience and the all-or-none* (the Hague: 1961). On some fundamental differences between phonemes and intonation, see Bolinger, "Intonation and, analysis", *Word*, 5 (1949), 248-54. The "expressive prosody" of S. S. Newman is also "not necessarily capable of the same type of systematization as that which is applicable to usual kinds of morphemes "("On the stress system of English", *Word*, 2 (1946), 172).

34. Cf. Quirk & D. Crystal, "On scales of contrast in connected English speech", *In Memory of J. R. Firth* (London: 1966), pp. 359-69. Also see D. Crystal, *Studies in the prosodic features of educated British English, with special reference to intonation* (unpublished University of London doctoral thesis), where the co-occurrence of prosodic features with each other and with grammatical structures is examined in detail.

35. Hockett, *op. cit.*, states that paralanguage has "largely no" discreteness, which is a premature decision that seems largely untrue.

36. Cf. Halliday's approach to intonation, *op. cit.* (1964), and Quirk & Crystal *op. cit.*

37. Failing this, one can always fall back on the view that these features are linguistic because they are studied by linguists!

38. Martinet, for example, states that it is "exceptional" to find clear-cut signifiés for contours (*A functional view of language* (Oxford: 1962), p. 30), though he admits (p. 37) that there has been insufficient research to display regular form-function correspondences.

39. Cf. C. E. Osgood, G. J. Suci, P. H. Tannenbaum, *The Measurement of Meaning* (Urbana, Ill: 1957).

40. See J. C. Catford, "Phonation types: the classification of some laryngeal components of speech production", *In Honour of Daniel Jones* (London: 1964), 26-37.

41. J. D. M. Laver, "The synthesis of voice-quality" (Mimo).

4 . There are such a large number of them: a quick count of those used in O'Connor & Arnold's *Intonation of Colloquial English* (London: 1961), A. C. Gimson's *Introduction to the Pronunciation of English* (London: 1962) and M. Schubiger's *English Intonation* (Tubingen: 1958) produced the fantastic total of 269 different labels.

43. Similar techniques have been used by psychologists, but rarely within a satisfactory linguistic frame of reference. For an introductory review of the literature, see E. Kramer, "Judgment of personal characteristics and emotions from nonverbal properties of speech", *Psychological Bulletin*, 60 (1963), 408-20.

Department of Linguistic Science, Faculty of Letters,

The University of Reading