David Crystal

Paralinguistics

The preliminary prospectus about this series (see p. 5) drew a clear contrast between the study of verbal language in man and the study of the expressive resources of the body as a whole. It was suggested that, compared with the rigorous attention paid to man's linguistic behaviour (particularly in the last two decades), his non-linguistic communicative abilities have been badly neglected. In the light of such an emphasis, it may not be immediately obvious why there should be a paper on a branch of linguistics within a volume ostensibly devoted to what is clearly not linguistic. But it is appropriate that linguistics should make its presence felt here. For one thing, many areas of non-linguistic study derive much of their stimulus and method from the linguistic theories and techniques of the 1940s and 1950s. And there is also the fundamental point that to wholly ignore the linguistic component is to commit as mortal an omission as the one which this series is trying to make good. Understanding man's expressive potential requires the concurrent study of both linguistic and non-linguistic modes of behaviour. Only a distorted picture can result from too rigid a separation between them.

But in any case, this paper is about 'paralinguistics', not linguistics – about 'paralanguage', not language. Paralanguage is in fact generally seen as a kind of bridge between non-linguistic forms of communicative behaviour and the traditionally central areas of 'verbal' linguistic study – grammar (in the sense of syntax and morphology, or accidence), vocabulary, and pronunciation (or, in the case of written language, spelling and punctuation). The study of the pronunciation system of a language is generally

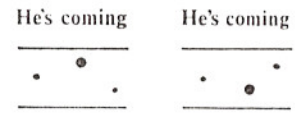
referred to as phonology, and within this the bulk of the linguist's effort goes towards the analysis of the vowel and consonant units, or phonemes, which constitute the identifiable syllables, words and sentences of meaningful communication. Over and above these 'central' properties of speech, however, there remain certain vocal effects which are qualitatively very different from phonemes or words, but which nonetheless seem to have an important role to play in the communication of meaning. These effects are often referred to popularly as 'tones of voice' – a convenient phrase which summarizes a complex functioning of the vocal apparatus, in which pitch, loudness, speed of speaking and many other vocal qualities are used in distinctive combinations – but in the literature on human communication, these features of pronunciation are subsumed under the heading of paralanguage. This term was originally chosen to reflect a view that such features as speed and loudness of speaking were marginal to the linguistic system – 'at the edge of language', as it was once put. It is a view which is no longer universally held, as we shall see, but the term 'paralanguage' has remained in general use nonetheless. Actually, there is considerable difference of opinion as to exactly what should be called paralinguistic in the communicative behaviour of a culture, and how it should be analysed. Not all the vocal effects to be mentioned below would be labelled paralanguage by everyone. Some scholars, also, include aspects of visual communication under this heading – facial expressions, for instance, and characteristics of writing (such as layout and spacing). There is indeed an important overlap between paralinguistic and kinesic function (hence the view of paralanguage as a 'bridge'); but in view of the emphasis on the visual in the rest of this volume, I propose to restrict the present paper to the vocal factors involved in paralanguage, which are a complex enough matter in their own right. I shall, however, be taking the broadest possible view of paralinguistic phenomena, including under this heading *any* meaningfully contrastive sound-effect which cannot be described in terms of the segments, or phonemes, in the sound system of a language, but which extends over stretches of utterance at least a syllable in length.

Paralanguage shares one thing with the study of other forms of body expression – namely, that it has been much neglected, even

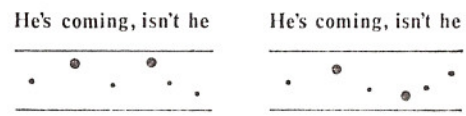
within linguistics. Indeed, the series on linguistics sponsored by the I.C.A. in 1971 made no mention of it.[1] There has in fact been a generally dismissive attitude towards the study of paralinguistic phenomena in the context of communicational analysis – an attitude which is perhaps reinforced by the etymological conditioning of our thinking arising out of the 'para-' prefix. There is the suggestion that tone of voice is a secondary facet of communication – a kind of optional 'extra', which does not affect the basic meaning of an utterance. Some of the reasons for this way of thinking will be discussed below; but before this, it should be emphasized that these attitudes have recently come under attack on a number of fronts, work in social psychology, psychiatry, sociolinguistics and elsewhere coinciding to suggest that the vocal effects called paralinguistic may be rather more central to the study of communication than was previously thought. Birdwhistell, another contributor to this volume (see p. 36), was one of the first to appreciate this point. In a 1959 paper called 'The frames in the communication process', he said: '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.' Certainly, observations of people's everyday reactions to language suggest that paralinguistic phenomena, far from being marginal, are frequently the primary determinants of behaviour in an interaction, sometimes pushing the so-called 'cognitive' or 'denotative' aspect of the utterances used into a secondary role. 'It's not what he said, but the way that he said it which upset/surprised/ . . . me' is the most widely-quoted phrase used in support of this point; but the importance of paralanguage can be similarly shown from a variety of other comments besides: 'Say it as if you mean it', 'You don't sound like a lawyer', 'You can keep that tone of voice for your secretary', and so on. If we begin our analysis of the communication situation by asking what features of the vocal stimulus account for the response behaviour, it is clear from such examples that paralanguage cannot be given anything other than a central role.

Once we look in detail at the various communicative functions

of paralanguage, this point becomes more cogent. The most widely recognized function is for emotional expression. The traditional view in psychology, for instance, is that verbal language communicates 'cognitive' meaning, whereas the non-verbal code (which covers my sense of paralanguage, amongst other things) communicates 'affective' meaning – anger, sarcasm, surprise, emphasis, excitement and so on. This is certainly an important role for paralanguage, and it is perhaps its most obvious role; but it would be wrong to assume – as some scholars have done – that this is its only function. On the contrary: far more important and pervasive than its affective function is the use of paralinguistic features as markers of an utterance's grammatical structure. The intelligibility of written communication is in large part due to the conventions of spacing and punctuation we adopt; in like manner, the grammatical intelligibility of speech is largely a product of the way in which we organize a stream of noise into structured units (sentences, phrases, words, etc.). Intonation – the systematic use of pitch in a language – is the most important factor here. More than any of the other variables which constitute tone of voice, intonation is used to segment and structure stretches of language, expounding contrasts in meaning which are sometimes almost as clear-cut as the contrasts signalled by phonemes or word-order. For example, the difference between stating and questioning may be signalled by a change in pitch, from falling to rising, as in



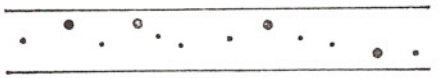
(The interlinear transcription represents the direction of movement of the pitch of the voice. The large dots indicate strong, or stressed syllables.) Another example is in the 'tag-questions' of the following two sentences:



The first sentence has a falling tone on the tag-question, and this

usually means that the speaker is expecting his listener to agree with him. The second sentence replaces this with a rising tone, and the meaning consequently changes: here the speaker is usually making a genuine request for information. These contrasts can be quite unambiguous. What makes intonation different from, say, the word-order differences of syntax (where statement versus question may also be expressed, as in 'He's coming' versus 'Is he coming?') is that sometimes in speech the contrast does not emerge so clearly ('Are you asking me or telling me?', one might hear), and sometimes, for a variety of reasons (not all of which are yet fully understood), one of these intonational tunes may be used with the semantic force of the other, as in the context:

He's coming, isn't he – I'm asking you, damn it!



Here the falling tone is used where, from the above examples, the rising tone might have been expected – presumably on account of the specific attitude adopted by the speaker.

But despite the 'fuzziness' which may surround intonational contrasts, there are enough clear cases to show that pitch *can* be used with a function corresponding to the cognitive use of word-order or morphological inflection in grammar. Moreover, it is not just pitch which works in this way. Increasing the speed of speaking, for example, is an important means of indicating that what one has just said was unintentional – a mistake – and to be replaced by the part of the sentence which was speeded up. In the sentence 'Those of you who aren't happy *aren't ready* for this announcement should . . .', the words in italics would normally be spoken more rapidly (and usually louder) than the surround, to indicate that a restructuring of the utterance has taken place. However, speed, rhythm and other tone-of-voice variations are not used as systematically for this purpose as are contrasts of pitch (and also those contrasts in loudness generally referred to as 'stress'). It is for this reason that some scholars take the intonation and stress systems separately from other paralinguistic characteristics, considering them to be more 'central' features of language.

Paralanguage has other communicational functions apart from the affective and the grammatical. It may be used as an index of our intentions, for instance: if we wish to show that we want to persuade, or irritate, or joke, then it is paralanguage, along with the appropriate facial expressions, which acts as primary exponent. In a similar, though less deliberate manner, paralinguistic effects are of major importance as indicators of social psychological states, such as dominance, submission, leadership, and so on. A great deal of research has been done into the nature and social correlates of such notions as 'brisk' and 'authoritative' voices, and stereotyped interpretations of a number of paralinguistic voice 'settings' have been studied in some detail. The range of paralinguistic effect used in television advertising shows this function very clearly: different types of product correlate with different types of voice – two well-recognized categories are the 'hard-sell' approach, with its dramatic, tense and rapid syllables, and the 'soft-sell' approach, with its gentle, melodious, leisurely tones. And it is this function which underlies Stephen Potter's recommendation about 'plonking' (*Lifemanship*, p. 43): 'If you have nothing to say, or, rather, something extremely stupid and obvious, say it, but in a 'plonking' tone of voice – i.e. roundly, but hollowly and dogmatically . . . if properly managed, the tone of voice will suggest that you can afford to say the obvious thing, because you have approached your conclusion the hard way, through a long apprenticeship of study.'

A related function of paralanguage is to indicate a speaker's professional background. Most professions in which speech is an integral part of the professional activity have a distinctive paralinguistic style – though some are more distinctive than others. Barristers, undertakers and clergymen are traditionally supposed to be most distinctive in this respect, but there is rarely any difficulty in distinguishing many others on the basis of a sample of tape-recorded speech – drill-sergeant, street vendor, disc-jockey, sports-commentator, lecturer, policeman ('*Your* car, sir?') . . . Professional comedians and satirists are well aware of the importance of paralinguistic features when they 'put on' a particular voice, either that of an individual, or that of a stereotyped social group or class; and the point has long been appreciated by teachers of speech and drama as an essential aspect of routine training.

'Style is the man' is a maxim which was largely viewed in relation to the written language. When we consider speech, it is paralanguage which is the man, as far as social identity is concerned.

More detailed illustrations will be found in the books listed at the end of this essay. But just from the above, it would seem that paralanguage has a complex function in communication, conveying grammatical, attitudinal and social information. If, then, it is such an important aspect of behaviour, why has it been so neglected? The reasons are very similar to those underlying the neglect of other facets of body behaviour. To begin with, there was the difficulty of getting hold of reliable samples of data for scientific study. The real range and complexity of paralinguistic phenomena emerges in informal conversational situations, and tape-recorded material of this kind is by no means easy to come by. Put a tape-recorder in front of the participants in a conversation and their interaction ceases to be normal: their language becomes more formal and less fluent, and their paralanguage alters radically. If the microphone and other equipment is hidden, then the paralanguage stays natural, but of course the quality of the recording may be poor. Quite sophisticated techniques are needed to get around these problems; and it is not surprising, therefore, that progress in this area has been slow. The tape-recorder itself, we must remember, is a relatively recent invention; and tape-repeaters (which provide convenient repeat listening to a piece of language, to ensure maximum phonetic accuracy in transcription) are still not widely used.

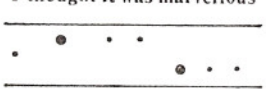
But assuming that some reliable data have been collected, the analyst has to face the problem of how to set about describing paralinguistic effects. And how do you identify and classify a tone of voice? The difficulty here is that linguistic techniques available for language description have been almost completely orientated towards the study of the segmental and verbal units of articulation and construction – the consonants, vowels, syllables, words and so on. The tradition in linguistics which makes the linguist look at language and see it as a sequence of discrete, non-overlapping entities (such as sounds, words) does not readily apply to phenomena such as speed of speaking or intonation. It is not so easy to specify the minimal units which are to form the sequences; and it is not so easy to establish the set of sequences which constitute the

language's permissible pronunciation patterns. What is a falling-pitch pattern, for instance? How high may it start? How low may it finish? And how long may it take to get from one point to the other? Let us imagine a situation in which someone says 'No' in a fairly neutral, matter-of-fact way, in a low pitch-range, and then repeats it in a progressively more excited and emphatic voice, letting his pitch get higher and higher – as if he were reacting to someone who was continually querying the truth of his answer. If the first version of 'no' were then played over alongside the last, there would be a clear contrast in form and meaning – a distinction between 'low' and 'high' falling tones, which we might interpret semantically as 'neutral' versus 'excited' (or in some such terms). But in between, there is a continuum of gradation, which makes it extremely difficult to decide where one meaning ends and the next begins. And if we decide to set up a 'high falling tone', with the meaning 'excited', then we immediately find difficulties. Not only do the physical limits of this tone vary considerably from one utterance to the next, and from one speaker to the next, with no obvious change in meaning; the meaning itself becomes extremely difficult to pinpoint. The same high falling tone may on one occasion help to indicate excitement, but on another it may be part of an attitude of anger, or surprise, or joy. Context conditions our interpretation here in a way which never happens with the segmental and verbal sides of language. There, the meaning of a word or sentence is much more readily definable and much more stable. There too, the formal basis of a contrast is more definable and stable: there is no continuum of gradation between two words, or structures, or phonemes, like that illustrated above. *Is he* does not gradually merge into *He is*. Λ /p/ does not gradually merge into a /b/. It is for such reasons that paralinguistic features have been referred to as the 'greasy' part of speech.

To some linguists, these difficulties of identification and semantic interpretation are evidence that the effects being described are not matters of language at all – that the prefix 'para-' should be taken literally, as it were. But this does not follow. Just because this area of behaviour is difficult to describe and quantify, it does not mean that it lacks system altogether. Perhaps the reason for our difficulty is simply that we lack appropriate techniques for handling gradience between phenomena, for evaluating affective

meaning and so on. I would in fact want to argue that this is the case. Our understanding of paralinguistic phenomena will not be increased as long as we approach the area assuming that unless we can see the sort of structure that we are used to seeing in verbal language, then there is no structure there at all. This is the kind of attitude that underlies the view, already criticised, that paralanguage has a secondary role to play in communication, that it has a 'merely' affective function, and the like. The argument also takes other forms – for instance, that paralinguistic effects are universal, being a product of nervous tension. It is sometimes said that, as we can always recognize a foreigner when he is being angry (for instance), expressive vocal behaviour must be outside the linguistic system, must be unlearned and culturally neutral. But these arguments embody fallacies, stemming from an over-simplified view of the complexity of paralanguage. We have already seen how paralinguistic effects have other roles than the affective – roles in which nervous tension can have little or no part. Cross-cultural studies, moreover, as they increase in depth of detail, bring to light more, not less paralinguistic difference between communities and cultures. And the fact of the matter is that it is *not* always easy to see that a foreigner is being angry – or sarcastic, or upset, or embarrassed. On the contrary: it is a very common reaction to misinterpret a foreigner's paralanguage – to assume that he is being rude, or belligerent, or sarcastic on the basis of his tone of voice, whereas in reality what we are responding to is unintentional interference from the paralinguistic features of his mother tongue. For example, flat, level tones in English regularly connote boredom or sarcasm:

'I thought it was marvellous'



spoken with a final level tone generally means the reverse! In Russian, on the other hand, the level tone is much more widely used with a neutral, matter-of-fact interpretation. The danger for the Russian learning English, then, is that he produces English sentences with too many level-tone endings: to his ears, the sentences sound intonationally neutral; but to the English listener,

they sound uninterested and often rude. Other examples? In some oriental languages, giggling is a normal indication of embarrassment on the part of adults, whereas in English it either relates to humour, or it is considered childish. In some varieties of Arabic, speaking with the tongue retracted ('velarization', as in much Liverpool or Birmingham speech) is an indication of masculinity; non-velarized speech is effeminate – which can lead to difficulties for the unsuspecting, non-velarizing male tourist!

The complexity of paralanguage can only be seen by attempting to carry out a systematic classification of the features within a number of languages. One's general aim is to set up an 'alphabet' of vocal effects, each of which is capable of altering the (affective, grammatical, or social) meaning of an utterance when it is substituted for another effect within it. Ultimately it would be necessary to classify the effects into categories, based upon their formal distribution and typical functions; but in the first instance, what we are concerned with is to establish the range of effects which are capable of being used by a language with *any* kind of semantic force. Some of these effects will accordingly be more obvious than others; some will be easier to describe than others; but these problems, as suggested above, are secondary.

In this way, it is possible to distinguish a number of variables within the human vocal apparatus which are regularly used in the production of tones of voice. (There is space for the briefest of examples only. To facilitate recognition, I will restrict the examples to English. Further illustration is provided in the accompanying bibliography.)

Pitch. In addition to the examples already given, one might illustrate from words or phrases spoken in a higher or a lower pitch range than normal, as when extended low pitch is used as a marker of parenthesis (e.g. 'My cousin – *you know, the one who lives in Liverpool* – he's just got a new job').

Loudness. Speaking words or phrases louder or softer than normal, in various degrees, is one of the more obvious systems of paralinguistic effect – used, for example, as an indication of rhetorical climax in public speaking, or as a marker of increased emphasis ('I want the *red* one, not the *green* one').

Speed. Words or phrases may be spoken faster or slower than

normal, as when 'Really' is spoken in a drawled, meditative manner, or when an increase in speed of speaking is conventionally interpreted as one of a small set of 'meanings', e.g. that the speaker wishes to forestall an interruption, or to suggest that what he is saying need not be given careful attention.

Rhythm. Pitch, loudness and speed patterns combine to produce contrasts in the rhythm of speech which have paralinguistic force, as when a sentence is spoken with a more marked metrical beat than normal to suggest irritation, e.g. 'I *really think* that *John* and *Mary* should have *asked*.'

(The range of effects thus far outlined is sometimes studied separately from all other paralinguistic variables under the heading of 'prosodic features' – a term which reflects the traditional interest in the study of metre, where stress and syllable-length in particular were considered to be central. Those who make use of the distinction between 'prosodic' and 'paralinguistic' effects use the latter term for such other variables as the following.)

Larynx effects. Whispered speech is one of the more obvious paralinguistic effects originating in the larynx – one of its most conventional interpretations being to indicate a 'conspiratorial' situation. Another example would be 'husky' speech, in which the throat is constricted to produce a hoarse effect, commonly used to connote disparagement (as when 'Never!' is spoken forcefully in a low pitch-range).

Oral effects. Increased lip-rounding ('labialization') is an important feature contributing to a number of paralinguistic effects, e.g. dislike, scorn or (most obviously) as a feature of intimate vocal play (as in talking to babies or animals). A muscularly tense, precise mode of articulation is commonly used as an indication that the speaker is becoming increasingly irritated.

This is by no means a complete classification. In addition, one could refer to various kinds and degrees of resonance of articulation, contrasts in register (e.g. normal versus falsetto voice), spasmodic articulations (e.g. giggling, tremulousness), nasal effects, and many more.

What should be clear from even this brief illustration is that it is important not to underestimate the range and subtlety of the

para-language that we can all intuitively interpret and produce ourselves as mature speakers of a language. This, as we have seen, is particularly evident during the process of foreign-language learning, where, characteristically, these are the features of one's mother tongue that it seems most difficult to eradicate. Even highly motivated foreign-language learners seem to find it almost impossible to replace their own paralinguistic system by that of another language. Why should this be? Doubtless it is something to do with the early age at which these features are learned. Paralinguistic features seem to be among the first language specific vocal contrasts produced by the child. It is normally assumed that a child begins to communicate in its own language when its 'first words' appear – usually around the end of the first year. But for several months before this, the child has already been using certain of his language's paralinguistic features. At around seven months, biologically conditioned babbling ceases to be random and undifferentiated: the vocalization becomes gradually organized into 'sentence-like chunks'. Long before one can identify specific vowels, consonants or words, there is an impression of organization and meaningfulness in the babbling, recognized by the parent in such comments as 'Baby always says that when his brother comes into the room'. The basis of this parental awareness is in the emerging intonational, rhythmic and other patterns which the child is introducing into its utterance. Babies respond to adult tones of voice very early indeed – from around two months; and it is these which are the first effects to emerge in their own production – from as early as seven months. It is at about this time that one can begin to tell children from different language backgrounds apart. Of course, the child takes many months to learn, control and use the whole range of his language's paralinguistic features; but the basic point is that during the first year he is well on his way to becoming an extremely competent paralinguist – which makes it hardly surprising that such features, being learned so early, are the most difficult to uproot later.

But much of this is speculation. We are still a long way from the stage of being able to state with confidence the facts of first- or second-language learning. It is but recently, after all, that paralinguistic phenomena have been studied at all in sufficient detail to warrant the formulation of testable hypotheses about their acqui-

sition. But progress is being made. It is at least now possible, using a general phonetic framework, to define and classify all the tones of voice that the human vocal apparatus can produce, and a number of descriptions of particular languages are well under way. Progress in studying the variety of functions in paralanguage is much slower, but even here information is accumulating, as the references below make clear. There is still a considerable gap, however, between our intuitive ability to recognize and interpret paralinguistic effect – our 'natural' sense of linguistic appropriateness and taboo – and our ability to state in clear terms what it is that we perceive. The spectre which still haunts papers on paralanguage, including this one, is the extraordinary difficulty of putting into words and diagrams what it is that we hear, in order that the effects described be as meaningful as possible to the reader. Nor is it at all obvious, at present, how paralinguistic information is to be correlated with the data derived from the study of other modes of expressive behaviour. But at least, these days, we are beginning to have some precise ideas about exactly what it is that has to be correlated.

1 Published as *Linguistics at Large*, ed. N. Minnis, Gollancz, 1971; Viking Press, New York, 1971; Paladin Books, 1973.

Crystal, David, *Prosodic systems and intonation in English*, Cambridge University Press, 1969.

Crystal, David, 'Prosodic and paralinguistic correlates of social categories' in E. Ardener, ed., *Social Anthropology and Language*, Tavistock, 1971, pp. 185–206.

Lyons, John, 'Human language' in R. A. Hinde, ed., *Non-Verbal Communication*, Cambridge University Press, 1972, pp. 49–85.

Sebeok, T. A., Hayes, A. S., and Bateson, M. C., eds., *Approaches to Semiotics*, Mouton, The Hague, 1964.