

D. Crystal

Prosodic and Paralinguistic Correlates of Social Categories

Research into the linguistic correlates of social categories has been almost exclusively based on the study of lexical, grammatical, and segmental phonetic and phonological characteristics. What are generally referred to as 'speech styles', i.e. modes of speaking restricted to or primarily associated with a particular social group, are illustrated solely with reference to restricted usage of items of vocabulary and of grammatical inflections or structures, and to differences in the articulations of vowels, consonants, and vowel-consonant sequences. There is remarkably little attention paid to one other aspect of speech, which I would hold is of major importance for the linguistic definition of social categories, namely, the *non-segmental* phonetic and phonological characteristics of utterance. These features I shall describe in more detail below; meanwhile, it will suffice to say that they refer to vocal effects due to contrasts in pitch, loudness, and speed of utterance, or to the use of qualities of voice such as nasalization or breathiness in order to communicate specific meanings. 'Intonation', or 'speech melody', clearly comes under this heading, therefore, as does 'rhythm', and what is regularly, albeit vaguely, called 'tone of voice'. Certain aspects of tone of voice are sometimes studied separately under the heading of 'paralanguage'. More precisely, non-segmental effect would include any sound effect which cannot be described by reference to a single segment (or phoneme) in the sound system of a language, but which either continues over a stretch of utterance (minimally one syllable), or requires reference to a number of segments in different parts of an utterance that are all affected by a single 'set' or configuration of the vocal organs - as when velarization of certain sounds produces a cumulative impression and a semantic interpretation that affect the whole of the utterance. Goffman (1964: 133) looks at this area from a different viewpoint,

referring to the expressive aspects of discourse which cannot be clearly transferred through writing to paper – he refers to them as the ‘greasy’ parts of speech! ‘It’s not *what* you say, but the way that you say it’ summarizes the scope of this field, from the formal point of view. Functionally, it is generally agreed that non-segmental phonology provides the main method of communicating emphasis and personal attitudes (sarcasm, surprise, etc.) in language; and that it may also be used with a grammatical role, as when intonation distinguishes between restrictive and non-restrictive relative clauses in English (cf. the use of commas in writing, as between ‘My brother, who’s abroad, wrote me a letter’ – one brother – and ‘My brother who’s abroad wrote me a letter’ – more than one brother). I would add that non-segmental phonology is also one of the main ways of establishing the identity of social groups in speech.

Reasons for the general absence of reference to the function of non-segmental phonology in this area of interdisciplinary overlap are not difficult to find. These features of language are among the most difficult to subject to analysis, being relatively difficult to perceive, transcribe, and measure. Most of us are unused to listening to differences in pitch and loudness, for example, and few people know what kind of phenomenon to look for. Also, the differences between these contrasts are typically less discrete than those between segmental contrasts: the distinction between a [p] and a [b], or even between two vowel sounds adjacent in articulation, is relatively clear-cut; whereas the distinction between a falling tone and a rising-falling tone, or between one and two degrees of stress, is sometimes extremely difficult to hear. Labov (1964: 166, and 176n.7), while allowing the importance of non-segmental phenomena in language, considers them to be essentially unquantifiable at the present stage of study, and therefore omits them from his own work on the ground that ‘we lack the large body of theory and practice in codifying intonation which we have for segmental phones’. There is also less of a tradition of study for these features than for other aspects of language organization. Features such as intonation (‘prosodic features’, as they are usually called) are referred to only sporadically, either as part of an attempt to explain the function of punctuation marks, or to define the phonetic correlates of metre. The simplified

Correlates of Social Categories

patterns used in textbooks for teaching English to foreigners, which have been around since the beginning of this century, are quite inadequate for research purposes, even as a basis for description; and in any case very little work has been done on languages other than English.

More recently, linguists have begun to examine non-segmental vocal effect in detail, but so far there has been little attempt to describe systematically the range of non-segmental features which are in principle operative in a language, or to work out a theory that will define and interrelate them satisfactorily. Work by Pike (1944) and Trager (1958, 1961), particularly the latter's research into paralinguage, the definition of the subject of *semiotics*, viz. the study of patterned, human communicative behaviour in all modalities (see Sebeok *et al.*, 1964), and recent work on voice quality (see Laver, 1968) have all done a great deal to stimulate interest in non-segmental vocal effect, but little of this has so far been used in social anthropology. The very important collection of papers on the ethnography of communication edited by Gumperz and Hymes (1964) provides another move in the right direction. Hymes in his introduction to the volume (and also in his contribution to the present monograph) demonstrates very clearly the need to develop an 'ethnography of speaking' – informally defined as a specification of what kinds of things one may appropriately say in what message forms to what kinds of people in what kinds of situation, and, given a set of alternatives, what consequences stem from selecting one rather than another – and refers to the need for semiotic and other studies. Most of the contributors to the volume underline this point at various places. Albert, for example, refers to the training in tone of voice and its modulation, *inter alia*, for men in Burundi, and shows its relevance to age, sex, kinship, and other relationships, referring to certain highly conventionalized speech patterns such as those used in visiting formulae, petitioning situations, rules of precedence, and respect patterns. Distinctions are made, many of them non-segmental, according to the social role of those present, the degree of formality (especially relating to whether the situation is public or private), and the objectives of the speech situation. 'Together, social role and situational prescriptions determine the order of precedence of speakers, relevant conventions of politeness, appropriate

formulas and styles of speech, and topics of discussion' (1964: 43). But despite this much-needed emphasis on theoretical principles, neither Albert nor any of the other contributors to the volume present any detailed account of the non-segmental phonology involved: the references stay at a maximally general level.

It should be emphasized that this collection of papers is quite exceptional in its orientation in this area. On the whole, most fieldworkers, even in linguistics, are still unaware, in principle, of the *kind* of linguistic phenomenon they are liable to come into contact with in this part of language, how they should label phenomena that they hear, or how they should integrate these with other aspects of any linguistic description they may happen to be making. The present paper is therefore an attempt to outline this area of study, so that the functional range of non-segmental features in a social anthropological context may be more readily recognized. But first I ought to indicate the extent to which non-segmental features *have* been noted in the description of social categories, either by linguists or by anthropologists, as this may help to clarify the nature of these features and underline the need for research in this area. A partial survey of the literature in this respect is not all that meaningful, in fact, because, in the absence of any generally agreed theory, there is no guarantee that different scholars are using such terms as 'melody', 'tone', and 'stress' in the same way (it is frequently obvious that they are not); and there have been few attempts to transcribe utterances in order to indicate the frequency of occurrence and distribution of specific effects. But at least some of the references used here may help to provide a context of situation for those not too familiar with the subject.

I have divided the main references into five generally recognized categories (though there is, of course, some overlapping): institutionalized non-segmental correlates (or indications, depending on the point of view) of *sex*, *age*, *status*, *occupation*, and *functions* (*genres*). I shall add some references to English in order to indicate further the kind of information involved.

1. *Sex*

It is probable that there are important non-segmental differences between the speech habits of men and women in most languages, though very little data have been analysed from this point of view. Cf. such informal remarks as 'Stop clucking like an old woman', or references to 'sexy' voice and the like (see Laver, 1968: 49). Intuitive impressions of effeminacy in English, for example, partly correlate with segmental effects such as lisping, but are mainly non-segmental: a 'simpling' voice, for instance, largely reduces to the use of a wider pitch-range than normal (for men), with *glissando* effects between stressed syllables, a more frequent use of complex tones (e.g. the fall-rise and the rise-fall), the use of breathiness and huskiness in the voice, and switching to a higher (falsetto) register from time to time. (This provides an interesting contrast with Mohave, for instance, where a man imitating a woman (or transvestite) does not change to falsetto, but uses his normal voice, and rather imitates verbal and segmental effects (see Devereux, 1949: 269).) According to Ferguson (in Sebeok *et al.*, 1964: 274), velarization in Arabic indicates, among other things, masculinity, whereas avoidance of velarization indicates the opposite. In Darkhat Mongol, women front all back and mid vowels (see Capell, 1966: 101). In Yana, men talking to men 'speak fully and deliberately', whereas when women are involved (as either speakers or hearers) 'a clipped style of utterance' is used (see Sapir, in Mandelbaum, 1949: 212). Also in Yana, to express interrogation, women lengthen final vowels, whereas men add a segmental suffix, *-n* (Sapir, *op. cit.*: 179-180, cf. also p. 211), though it is a descriptive problem whether the length should be interpreted non-segmentally or not. In Chichimeca, where male and female names in the same family may be identical, it is reported that 'tone' may be used to differentiate the sexes being addressed (see Driver and Driver, 1963: 108). Syllabic tone differences may distinguish between sexes in Koasati (Haas, 1944). Sex differences, moreover, sometimes correlate with age. According to Garbell (1965), many female speakers over seventy of Urmi, a dialect spoken by Jews in north Persian Azerbaijan, replace practically all 'plain' words by 'flat' words, i.e. words consisting of 'flat' phones, which in Garbell's metalanguage means such

D. Crystal

features as the strong velarization of all oral consonants, the articulation of all labials with marked lip protrusion and rounding, and pharyngealization. Again, *responses* to non-segmental vocal effects can be a valuable part of a description, e.g. in Mohave, the breaking of the male voice in adolescence is not considered an important, or even a relevant, indication of puberty (Devereux, 1949: 268), whereas, of course, in English it is a feature that is regularly remarked upon. And, as a last example, one could note the training in voice modulation that Burundi men receive, but women do not (Albert, 1964: 37).

2. Age

References to the non-segmental correlates of age are very sporadic indeed in the sociolinguistic literature, though this was one of the most readily demonstrable correlations in the early work in social psychology (see Kramer, 1963; Allport and Cantril, 1934), and one has a perfectly clear intuitive impression of 'old', 'young' voices, and the like. In fact, the only regular references are to baby-talk (i.e. the speech characteristics of adults addressing babies). Kelkar (1964), under the heading of *paraphonology*, refers to the extended pitch and loudness characteristics, and the relatively slow and regular speed of baby-talk in Marathi, and mentions certain general vocal effects, such as pouting and palatalization. Ferguson (1964), with reference to English, also cites the higher overall pitch of baby-talk, the preference for certain pitch contours, and labialization, but does not discuss it further. Ervin-Tripp (1964) refers to some general characteristics of children's play-intonations, and Burling (1966) shows that broad rhythmic similarities exist between samples of children's verse from a number of languages. It is highly likely that older groups are also discriminated by non-segmental features (and not just by grammar and vocabulary, which are the only areas generally cited), but there is no published evidence on the point, apart from a few general remarks, as in Albert (1964) and Garbell (1965).

3. Status

Non-segmental phonology is frequently used to indicate the social identity of the speaker on a scale of some kind (his 'class

dialect', as many would say), or the identity of the receiver in these terms. Certain tones of 'respect' might be conventionalized indications of a particular kinship or caste relation, for instance, or may indicate different social roles. John Boman Adams mentions the importance of stereotyped pitch patterns and tones of voice in order to establish status between participants in one dialect of Egyptian:

'The villager is ordinarily conditioned to give and receive communications whose content is so stereotyped that he pays little attention to it other than to note that it conforms to the norms of traditional utterances and that the speaker is socially acceptable. . . . These statuses are often established in the exchange of stereotyped expressions of esteem and concern that are obligatory whenever two or more persons meet. Since the same expressions are always uttered, interpretations of "friendliness" or "enmity" depend upon meanings conveyed by subtle qualities of tone, pitch, and melody. These qualities, in their different modes, are interpretable to one who is acquainted with their culturally defined meanings' (1957: 226).

In Cayuvava, a rapidly disappearing language in Bolivia, there is a set of nasal phonemes, but nasalization also occurs with 'honorific' stylistic function (according to Key, 1967: 19): an individual of lower social or economic status addresses one of higher rank with a prominence of nasalization for all vowels of the utterance; and similarly with a woman being polite to her husband, or a man asking a favour. Albert (1964) refers to a number of similar examples, also instancing a typical sociolinguistic use of silence in this respect: in Burundi conclave, the silence of the highest-ranking person negates the proceedings, indicating total disapproval (whereas silence of lower-ranking people would have no comparable effect) (p. 41). Gumperz (1964: 144) distinguishes between two forms of the vernacular in Khalapur, *moṭī bolī* and *saf bolī*, the former being used primarily within the family group, the latter being used in external relationships, and refers to particular distributions of pitch glides occurring in the former but not in the latter. Longacre (1957) notes a very restricted formal third person enclitic in Mixtecan, which adds length and nasal quality to the

D. Crystal

syllable. What Shapiro (1968) calls 'explicit' and 'elliptical' codes in Russian are generally distinguished, *inter alia*, by tempo, the latter being faster. Bernstein (1964) makes some reference to intonation in his distinction between restricted and elaborated codes in English.

Many of these oppositions imply a distinction between formal and informal (non-casual and casual, etc.) kinds of speech, which has been frequently referred to in English (e.g. by Joos, 1962). Speed of utterance is presumably one of the features that would distinguish formal from informal speech; and there is, as usual, a fair amount of informal evidence for the existence of status styles in English, e.g. 'How dare you talk to me like that! I'm not one of your employees/students/secretaries . . .', and reference to 'la-di-da' voices, 'talking down', and so on. Hoenigswald (1966: 19) makes the point that in this field it is important to study the *ideals* of speech behaviour cherished by a group as well as the actual speech behaviour used, and an interesting area of research will be the systematic examination of elocution handbooks – not to condemn them, as linguists generally do, but to view them descriptively, as data concerning the desired (real or imagined) correlates of genteel, educated speech and the reverse. Non-segmental effect is regularly referred to here: see, for example, the influential work of the American elocutionist Rush (1827) in this respect. One should also note the vocal effects used, sometimes as mocking forms, when addressing a member of a stigmatized group, e.g. Sapir refers to the 'thickish' sounds of s and j 'pronounced with the lower jaw held in front of the upper' when talking to hunchbacks in Nootka (see Mandelbaum, p. 183), and to the way of satirizing cowards in this language, when either addressing them or referring to them, by 'making one's voice small' (i.e. using a 'thin, piping voice') (p. 184). Cf. children's sing-song cat-calls in English, using the tune



Gumperz, finally, mentions the importance of sentence speed and pause (as hesitation) in the analysis of status, and suggests that it is about time that scholars broadened the range of their linguistic investigations to take account of these matters (in Bright, 1966: 46). The points mentioned in this section are

Correlates of Social Categories

certainly only some of the possibilities. There are hardly any data to illuminate the question of the covariance of change in social status with change in tone of voice (e.g. after marriage, or after some initiation rite), though the existence of such phenomena can hardly be doubted.

4. Occupation

In English we are all familiar with the 'tone of voice' that is generally attributed to people acting in their professional capacities, such as the clergyman, lawyer, and undertaker. Phrases such as 'you sound like a clergyman' are conventionally meaningful, and would be interpreted (e.g. in an attempted imitation) as referring to a vocal effect in which pitch-range movements were narrowed, there was frequent use of monotone, rhythm was regular, tempo fairly slow, and overall pitch-height and resonance of the voice were increased. There are many occupations that would be recognized primarily on the basis of the non-segmental features involved, e.g. the disc jockey, barrister, preacher, street vendor, parade-ground commander, sports commentator, and many other kinds of radio and television announcer. Certain of these roles naturally overlap with status to some extent. Miller (1956: 181) talks about *authority roles*, by which he means 'a conceptualized position within a system of interpersonal relations whose incumbent is authorized to perform designated regulative functions for a designated action group during designated activity episodes', and many of those he cites (perhaps all?), e.g. drill sergeant, coxswain, foreman, cheer leader, involve the use of non-segmental features. The notion of a 'professional voice' is commonplace, if ill defined. Lecturers are generally aware of the kind of feature they have to introduce into their voice in order to awaken enthusiasm or promote participation in an otherwise dead class or audience.

Many of these matters have been given some experimental support, and it is not difficult to plan tests in order to verify one's impressions. A great deal of work has already taken place in the related area of defining personality traits, usually in the form of presenting judges with various non-segmental patterns (the verbal side of the utterance having been removed, e.g. by

D. Crystal

using nonsense-words, or acoustic filtering devices, or by articulating the different patterns on a single, neutral sentence – see Kramer, 1963), and asking them to rate the function of these patterns in terms of various traits. ('Trait' is fairly broadly defined in such work, and subsumes age, sex, and certain occupational characteristics.) This research, largely reported in psychology journals between about 1935 and 1950, is methodologically unsatisfactory in many respects (e.g. insufficient attention was paid to the backgrounds (i.e. the preconceptions) of the judges, and there was a blurring of theoretical concepts which should have been kept apart – the difference between voice quality and linguistic contrasts, for instance, which I shall discuss further below), but certain correlations between non-segmental patterns and features of extra-linguistic situations did emerge – and were generally referred to as *voice stereotypes* (see Crystal, 1969, Ch. 2, for a review of this literature). An important theoretical distinction which was not made in this work, but which must be made in future research, is that between *recognition* and *production* stereotypes (cf. the distinction between passive and active in vocabulary study); for example, it is part of our competence that we can discriminate various kinds of radio and television styles of speech, but I would agree with Labov that few speakers are ever directly influenced by such patterns as far as production is concerned (see 1967: 74).

There seem to be few occupational differences involving non-segmental features mentioned in foreign language descriptions. The only area that receives a regular mention is religious and magical language, and this really overlaps my next category, speech functions.

5. *Speech functions*

Particular modalities, or genres of speaking, are generally signalled through the use of non-segmental characteristics, as elocutionists are well aware. In the context of oral literature, one would also expect frequent use of these features, as they would provide an important means of adding further variation to the very restricted, stylized scope of a poem or story. Jacobs (1956: 127) emphasizes that

Correlates of Social Categories

'stylized devices such as connectives, pauses, and vocal mannerisms, to effect transitions from Scene to Scene or Act to Act in a longer story, are invariably discernible in its dictation in the native language. But publications infrequently if ever preserve evidence of these devices.'

In Mohave, when a traditional, memorized text is being uttered, it is delivered in a staccato, rapid manner, which the speakers find very difficult to slow down (sometimes impossible, when the utterance is in front of other people from the same tribe - Devereux, 1949: 269). Henry (1936: 251) refers to changes in force, pitch, vowel quality, aspiration, and pharyngealization as Kaingang story-telling devices which were commonplace rhetorical forms in the language, e.g. 'the Kaingang always raise their voices when they are describing some long drawn-out activity, and their voices even take on what might be to us a complaining tone' (which tone, incidentally, 'was the usual tone to describe the slow climbing of a hill!'). Sapir talks of *styles of recitative* in Paiute, referring to the speech of certain mythological or traditional characters designated by certain sounds and tones of voice (in Mandelbaum, p. 186, and cf. p. 465). Related to this is the tone of voice adopted by a community when imitating another: cf. Sapir's remarks about the Nootka's imitations of other tribes, e.g. adopting velar resonance (speaking 'in a rumbling fashion') for the Uchucklesit (Mandelbaum, p. 193), or speaking in a 'drawling' manner (i.e. 'a somewhat exaggerated rise in pitch towards the end of a sentence') for the northern Nootka (p. 194). The vocal stereotypes adopted by comedians, stage villains, and certain traditional pantomime characters in our own culture would be further cases in point.

Distinct genres also exist in conversation in some languages. In Shiriana, different types of conversation can be distinguished on a non-segmental basis. Migliazza and Grimes differentiate between 'one-sided' and 'balanced' types (1961: 36-37). They illustrate the former by reference to 'myths' and 'narratives', and their distinction is worth quoting at length because of its detail:

'Phonologically a myth is characterized by an initial period in which only lento pause groups occur, a body in which combinations of lento and andante pause groups occur, many of

D. Crystal

which contain ideophonic feet,¹ and a termination in which one or two lento pause groups occur, with extra length on the vowel of the final stressed syllable in the contour and at times a voiced breath intake after the end of the final pause group . . . four pitch levels [adequate for normal speech, DC] do not handle the pitch patterns, which range over a wide area and move largely in long glissandos.'

As an example of 'balanced conversation', the authors cite the *bargaining dialogue*, which is

'delivered at night by a trading partner from one village to his partner from another in the presence of all hosts and members of a trading party, in which each partner's speech has the general characteristics of a monologue, except that the intonation is replaced by a chant form' (p. 38).

They also mention the relevance of *crescendo*, *decrescendo*, and laryngealization for the definition of certain speech styles.

Another well worked-out example is Conklin's study of ways of modifying normal speech patterns for purposes of entertainment or concealment (most frequently as part of voice disguise in customary courting behaviour) in Hanunóo, a language of the Philippines. There may be both segmental and non-segmental aspects of this, but the latter vary independently of the former. There are four types: *yanas* (barely audible whispering), *paliksih* (utterance involving clipped pronunciation, greater speed, greater glottal tension, expansion of the intonational contours, and shortening of the long vowels), *padiqitun* (falsetto), and *paha-gut* (any sequence of articulations during which the direction of air flow in normal speech is reversed, i.e. inhalation).

Malinowski implies the relevance of non-segmental phonological effect at many places in *The Language of Magic and Gardening*, and makes explicit reference to it in his notes about 'modes of recitation' of magical formulae. Fischer (1966: 180-181) refers to the variation in the recitation of magical formulae between Ponape and Truk: the repetitions of the former are as exact as possible, but the latter make great use of expressive variations. Genres of religious speaking are regularly prosodically distinctive. Conwell and Juilland (1963: 30) refer to the distinctive rhythmicality of prayers in Louisiana

French, which is apparently very similar to the rhythmicality of litanies and other liturgical languages in English (see Crystal and Davy, 1969: Ch. 6). West reports that in Mikasuki, the language of the Seminole Indians in Florida, stress and tone differences are minimized in sermons (1962: 90). The introduction of song and chant characteristics into Welsh preaching (*hucyl*) involves markedly different pitch, length, and speed characteristics, as well as such paralinguistic effects as resonance and tremulousness.

A typology of speech functions in language has not yet been established, though there have been numerous attempts at it. The scattered comments collected in this section clearly indicate that non-segmental effect will be a major part of the definition of the physical basis of these functions. In English, phatic communion, routine requests, avoidance ploys, routine format (sports results, weather forecasting, etc.), public-speaking, official announcements, ceremonial language, sports commentary, telephone conversation, television advertising (and other forms of persuasion): these are just some of the areas where intonational and related phonological features are markedly different from those used in spontaneous utterance. For further references to speech functions, see Stern (1956, esp. pp. 382-383) and Frake (1964).

There has, then, been considerable sporadic, impressionistic comment as to the sociolinguistic function of non-segmental effects, though the utility of this has been marred by lack of an adequate theory, inexplicitness of definition, and certain methodological weaknesses. As an example of the latter, it is sometimes difficult to know the extent to which the description of a given effect is intended as referring to a linguistic feature of an individual, a group, or the language as a whole. When data are restricted to the output of one or two speakers of a language, there is always the danger of a lack of perspective causing misinterpretation (a problem that takes an extreme form in the Cayuvava language mentioned above, where there were apparently only six living speakers at the time the description was made). Also, too little reference is made to voice-quality norms for the languages as wholes, e.g. establishing overall pitch-range, loudness, speed, and so on. But enough has been

done to show clear lines of research, and the links that exist between linguistics and related, non-linguistic, semiotic fields. Here one might instance the importance of *speech surrogate* systems, such as the use of conventionalized whistling patterns, which may reflect intonational or paralinguistic patterning in the language (cf. wolf-whistles and rise-fall intonations in English), or the co-occurrence of kinesic features with speech (cf. La Barre, 1964), or the relationship between intonation and primitive music (cf. Herzog, 1934). Stankiewicz (1964) provides further comment on this point.

It is possible that an 'integrated theory' of all the observations made in this area may prove as valuable to social anthropology as Trager's framework has been in stimulating and helping to codify psychiatric research in America (see the references in Sebeok *et al.*, 1964). The descriptive framework outlined in Crystal and Quirk (1964), and developed in Crystal (1969) and elsewhere, allows for the incorporation of all the effects noted in this paper, and groups them into systems on the basis of shared formal characteristics. The following conceptual stages need to be distinguished:

1. *Non-linguistic vocal effects*

(a) *Voice quality*. Speech (or any act of communication) takes place against a personal and environmental background, which has to be identified by the analyst, in order to be discounted. Voice quality is the idiosyncratic, relatively permanent, vocal background of an individual, which allows us to recognize him, as opposed to other members of the group. It may be both segmental and non-segmental in character, but the latter is usually the dominant factor. It is a physiologically determined activity, over which most individuals have little or no measure of control. For a useful model of voice quality, see Laver (1968). In the present paper, I am concerned only with those non-segmental features which display - to however small a degree - a group-identifying function.

(b) *Physiological reflexes*, such as coughs, sneezes, or husky voice due to a sore throat, may also occur along with speech, and must also be discounted as background 'noise'.

Correlates of Social Categories

2. Semiotic frame

A model of an act of communication in semiotics is generally viewed as a bundle 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' and a 'communication configuration' or 'network'. It is distinct from the personal and physical background in that (a) it is variable with reference to the biological characteristics of the individual communicator, but is a pattern of behaviour shared by a group, and (b) the activity has always some culturally determined, relatively conventionalized value, or 'meaning'. The sub-systems are five in all, corresponding to the five senses, vocal/auditory, visual, tactile, olfactory, and gustatory; but only the first three are regularly used in normal communication (the latter have little potential structure, but are none the less of some importance as carriers of information to such people as doctor and chemist, e.g. in analysing body odours). The study of patterned, conventionalized, visual human bodily behaviour (facial expressions and bodily gestures) is known as *kinesics*. Non-vocal communicative sub-systems have not been the subject-matter of this paper: their relevance to anthropology is discussed in La Barre (1964).

3. Vocal-auditory component

The vocal-auditory component in communication can be broken down into the following categories:

(a) *Segmental-verbal*. This, the traditional centre of linguistic attention, would in its widest definition cover segmental phonetics and phonology, morphology, syntax, and vocabulary. A sub-set of verbal items is usually distinguished in semiotic literature: these, generally referred to as *vocalizations*, cover such items as 'mhm', 'shhh', 'tut tut', and the like, which are articulated using sounds outside the normal range of phonetic resources in the language. These overlap, formally and functionally, the next category.

(b) *Pause phenomena*. These comprise the various degrees of silence and 'voiced pause' (e.g. the 'ers' of English) that exist

in a language. These features are clearly segmental, from the formal point of view, but functionally silence overlaps non-segmental features, as it enters into the physical definition of such effects as rhythmicity and intonation contour (and is partly subsumed under the notion of juncture, by some scholars).

(c) *Non-segmental features*. These are aspects of the phonic continuum which have an essentially variable relationship to the phonemes and words selected as defined by (a) above. Detailed illustration of all features is provided in Crystal (1969). They may be grouped into two general categories:

Prosodic features

These are meaningful contrasts due to variations in the attributes of pitch, loudness, and duration (which have a primary, but not an identifying, relationship to the fundamental frequency of vocal-cord vibration, amplitude of vocal-cord vibration, and speed of articulation respectively), either singly or in combination. Some values from these three variables permanently characterize speech. Prosodic features sharing a similar formal basis and displaying some mutual definition of contrastivity are grouped together into *prosodic systems*. The following systems have been distinguished:

(i) *Pitch*. There are two systems of pitch, *tone* (referring to the direction of pitch-movement in a syllable, as when it falls, rises, or stays level, or does some of these things in rapid succession) and *pitch-range*. By pitch-range, I mean the distance between adjacent syllables or stretches of utterance identified in terms of a scale running from low to high. Speakers and groups have a normal pitch-level and -range, and they may depart from this in different ways to produce extra-high or -low speech, either in a sudden step-up or -down, or gradually. The normal distance between adjacent syllables may be narrowed (perhaps reduced to monotone) or widened, and different languages display different kinds and degrees of pitch-range variation. The patterns of pitch-movement that occur in a language are referred to as the *intonation*. Connected speech is considered as analysable into a series of

Correlates of Social Categories

units of intonation (variously called *tone-units* or *tone-groups*), which have a definable internal structure, and which function in sequences to produce melodic contours of a more general nature.

(ii) *Loudness*. Degrees of loudness which affect single syllables are generally referred to as degrees of *stress*. (*Accent* refers to a syllable which has been made prominent owing to a combination of both loudness and pitch factors.) Speakers and groups have a norm of loudness, which they may depart from in different attitudes, styles of speech, etc. Over stretches of utterance, there may be *forte* or *piano* loudness, to various degrees. As with pitch-range, the change from one level of loudness to another may be sudden or gradual (as with *crescendo* and *diminuendo* utterance.)

(iii) *Tempo*. Single syllables may be shortened or lengthened (clipped and drawled respectively); stretches of utterance may be faster or slower than normal for a speaker or group, to various degrees (*allegro*, *lento*), and, as above, the change may take place suddenly or gradually (*accelerando*, *rallentando* speech).

(iv) *Rhythmicality*. Combinations of pitch, loudness, and duration effects produce rhythmic alternations in speaking, distinct from the rhythmic norm of the language, e.g. increasing the perceived regularity of a sequence of stressed syllables in an utterance, or decreasing it; clipping certain syllables to produce staccato speech, or slurring them, to produce *glissando* or *legato* utterance. There are numerous possible contrasts here, and of course the physical correlates of each would have to be carefully defined in any description.

Paralinguistic features

Non-segmental variations *other* than those caused primarily by pitch, loudness, and speed, i.e. where other physiological mechanisms in the oral, nasal, or pharyngeal cavities are being used to produce an effect, are referred to under this heading. Prosodic features, being permanent features of utterance, of course enter into these effects, but they are

D. Crystal

variable in respect of their definition: any of the features listed below can be uttered with variable pitch, loudness, and speed (with one or two minor restrictions). Paralinguistic features are discontinuous and relatively infrequent in speech. They do not display such clear-cut formal and functional contrasts as do prosodic features, consequently the systemicness of their function is more difficult to demonstrate. One possible system would group together the different kinds of tenseness that may occur in a language, e.g. tense, lax, slurred, and precise articulations; others would involve degrees and kinds of resonance of articulation, contrasts in register (e.g. falsetto, chest), degrees of pharyngeal constriction (e.g. huskiness), types of whisper and breathy articulation, spasmodic articulations (i.e. the pulsations of air from the lungs are out-of-phase with the syllables of an utterance, as when one laughs or sobs while speaking, or says something in a tremulous tone), general retraction or advancement of the tongue (e.g. velarization), distinctive use of the lips (labialization), and various kinds of nasalization. A complete description of the possibilities here has not been written, but this cannot really be carried out in the absence of reliable data.

If one examines the data discussed in the first half of this paper in the light of the categories outlined in the second, it will be seen that all the vocal effects cited (or, at least, plausible interpretations of all these effects) can be described in terms of one or more of these categories. A great deal more work is needed before such an approach could be formalized as part of any general phonological theory; meanwhile, it may be the case that even a tentative formulation could stimulate fieldworkers to look more closely at this aspect of language, thereby providing the reliable and wide-ranging data that this corner of linguistics so badly needs.

NOTE

1. An ideophonic foot is a highly conventionalized effect, referring to a rhythm unit 'accompanied either by an anomalous pitch pattern . . . or by a voice quality that stands out in contrast with that of the rest of the utterance

Correlates of Social Categories

(usually laryngealized or breathy in relation to the overall voice quality)' (Migliazza and Grimes, 1961, p. 35).

Lento and *andante* are differentiated partly in terms of speed and partly in terms of the number of contours involved.

REFERENCES

- ADAMS, J. B. 1957. Culture and Conflict in an Egyptian Village. *American Anthropologist* 59: 225-235.
- ALBERT, E. M. 1964. 'Rhetoric', 'Logic', and 'Poetics' in Burundi: Culture Patterning of Speech Behavior. In Gumperz and Hymes (eds.), 1964.
- ALLPORT, G. W., and CANTRIL, H. 1934. Judging Personality from Voice. *Journal of Social Psychology* 5: 37-55.
- BERNSTEIN, B. 1964. Elaborated and Restricted Codes: Their Social Origins and Some Consequences. In Gumperz and Hymes (eds.), 1964.
- BRIGHT, W. (ed.). 1966. *Sociolinguistics*. Proceedings of the UCLA Sociolinguistics Conference, 1964. The Hague: Mouton.
- BURLING, R. 1966. The Metrics of Children's Verse: A Cross-linguistic Study. *American Anthropologist* 68: 1418-1441.
- CAPELL, A. 1966. *Studies in Socio-linguistics*. The Hague: Mouton.
- CONKLIN, H. C. 1959. Linguistic Play in its Cultural Context. *Language* 35: 631-636.
- CONWELL, M. J., and JUILLAND, A. 1963. *Louisiana French Grammar*. Vol. 1, *Phonology, Morphology and Syntax*. Series Practica I. The Hague: Mouton.
- CRYSTAL, D. 1969. *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.
- and DAVY, D. 1969. *Investigating English Style*. London: Longmans.
- and QUIRK, R. 1964. *Systems of Prosodic and Paralinguistic Features in English*. The Hague: Mouton.
- DEVEREUX, G. 1949. Mohave Voice and Speech Mannerisms. *Word* 5: 268-272.
- DRIVER, H. E., and DRIVER, W. 1963. *Ethnography and Acculturation of the Chichimeca-Jonaz of Northeast Mexico*. International Journal of American Linguistics Publication 26. The Hague: Mouton; Bloomington: Indiana University Research Center in Anthropology, Folklore and Linguistics.
- ERVIN-TRIPP, S. 1964. An Analysis of the Interaction of Language, Topic and Listener. In Gumperz and Hymes (eds.), 1964.
- FERGUSON, C. A. 1964. Baby-talk in Six Languages. In Gumperz and Hymes (eds.), 1964.

D. Crystal

- FISCHER, J. L. 1966. Syntax and Social Structure: Truk and Ponape. In Bright (ed.), 1966, pp. 168-183.
- FRAKE, C. O. 1964. How to Ask for a Drink in Subanum. In Gumperz and Hymes (eds.), 1964.
- GARBELL, I. 1965. *The Jewish Neo-Aramaic Dialect of Persian Azerbaijan*. Linguistic Analysis and Folkloristic Texts. Series Practica III. The Hague: Mouton.
- GOFFMAN, E. 1964. The Neglected Situation. In Gumperz and Hymes (eds.), 1964.
- GUMPERZ, J. J. 1964. Linguistic and Social Interaction in Two Communities. In Gumperz and Hymes (eds.), 1964.
- 1966. On the Ethnology of Linguistic Change. In Bright (ed.), 1966, pp. 27-38; discussion pp. 39-49.
- and HYMES, D. (eds.). 1964. *The Ethnography of Communication*. Menasha, Wis.: American Anthropological Association.
- HAAS, M. R. 1944. Men's and Women's Speech in Koasati. *Language* 20: 142-149.
- HENRY, J. 1936. The Linguistic Expression of Emotion. *American Anthropologist* 38: 250-256.
- HERZOG, G. 1934. Speech Melody and Primitive Music. *Musical Quarterly* 20: 452-466.
- HOENIGSWALD, H. M. 1966. A Proposal for the Study of Folk-Linguistics. In Bright (ed.), 1966, pp. 16-20.
- JACOBS, M. 1956. Thoughts on Methodology for Comprehension of an Oral Literature. In Wallace (ed.), 1956, pp. 123-130.
- JOOS, M. 1962. *The Five Clocks*. International Journal of American Linguistics Publication 22. The Hague: Mouton; Bloomington: Indiana University Research Center in Anthropology, Folklore and Linguistics.
- KELKAR, A. 1964. Marathi Baby-talk. *Word* 20: 40-54.
- KEY, H. 1967. *Morphology of Caywava*. Series Practica 53. The Hague: Mouton.
- KRAMER, E. 1963. Judgment of Personal Characteristics and Emotions from Non-verbal Properties of Speech. *Psychological Bulletin* 60: 408-420.
- LA BARRE, W. 1964. Paralinguistics, Kinesics, and Cultural Anthropology. In Sebeok, Hayes, and Bateson (eds.), 1964, pp. 191-220.
- LABOV, W. 1964. Phonological Correlates of Social Stratification. In Gumperz and Hymes (eds.), 1964.
- 1967. The Effect of Social Mobility on Linguistic Behavior. In Lieberman (ed.), 1967, pp. 58-75.
- LAVER, J. 1968. Voice Quality and Indexical Information. *British Journal of Disorders of Communication* 3: 43-54.

Correlates of Social Categories

- LIEBERSON, S. (ed.). 1967. *Explorations in Sociolinguistics*. International Journal of American Linguistics Publication 33. The Hague: Mouton; Bloomington: Indiana University Research Center in Anthropology, Folklore and Linguistics, Publication 44.
- LONGACRE, R. E. 1957. *Proto-Mixtecan*. Bloomington: Indiana University Research Center in Anthropology, Folklore and Linguistics, Publication 5.
- MANDELBAUM, D. G. (ed.). 1949. *Selected Writings of Edward Sapir*. Berkeley, Calif.: University of California Press; Cambridge: Cambridge University Press, 1950.
- MIGLIAZZA, E., and GRIMES, J. E. 1961. Shiriana Phonology. *Anthropological Linguistics* 3: 31-41.
- MILLER, W. B. 1956. A System for Describing and Analysing the Regulation of Coordinated Activity. In Wallace (ed.), 1956, pp. 175-182.
- PIKE, K. L. 1944. *The Intonation of American English*. Ann Arbor, Mich.: University of Michigan Press.
- RUSH, J. 1827. *The Philosophy of the Human Voice*. Philadelphia: Lippincott.
- SAPIR, E. 1910. Song Recitative in Paiute Mythology. In Mandelbaum (ed.), 1949, pp. 463-467.
- 1915. Abnormal Types of Speech in Nootka. In Mandelbaum (ed.), 1949, pp. 179-196.
- 1929. Male and Female Forms of Speech in Yana. In Mandelbaum (ed.), 1949, pp. 206-212.
- SEBEOK, T. A., HAYES, A. S., and BATESON, M. C. (eds.). 1964. *Approaches to Semiotics*. The Hague: Mouton.
- SHAPIRO, M. 1968. *Russian Phonetic Variants and Phonostylistics*. University of California Publications in Linguistics 49. Berkeley: University of California Press.
- STANKIEWICZ, E. 1964. Problems of Emotive Language. In Sebeok, Hayes, and Bateson (eds.), 1964, pp. 239-264.
- STERN, T. 1956. Some Sources of Variability in Klamath Mythology, 3: Style and Elements of the Myth. *Journal of American Folklore* 69: 377-386.
- TRAGER, G. L. 1958. Paralanguage: A First Approximation. *Studies in Linguistics* 13: 1-12.
- 1961. The Typology of Paralanguage. *Anthropological Linguistics* 3: 17-21.
- WALLACE, A. F. C. (ed.). 1956. *Men and Cultures*. Selected papers of the 5th International Congress of Anthropological and Ethnological Sciences. Philadelphia: University of Pennsylvania Press.

D. Crystal

WEST, J. D. 1962. Phonology of Mikasuki. *Studies in Linguistics*
16: 77-91.

© D. Crystal 1971