Relations between the sounds of speech and language's written forms pose many different kinds of question. There is always, from the outset, question about the principle of representation which exists between the two, with, regionally and historically, differing conventions governing scripts. There can be pictographs, for example, in which a picture represents an object, or there can be ideographs, in which concepts are inferred from the pictures—with the two seemingly combined in Egyptian hieroglyphics, five thousand years ago; there are kinds of 'rebus' writing, in which a phonetic transfer is made between one word corresponding to the picture and an existing homophone; then again, there are the systems which transcribe speech, syllabaries and alphabets (Hebrew without vowels, then the ancient Greeks—still around three thousand years ago—filling this omission as regards reproducing the sound shape of speech).

Even such a fundamental idea of a variable correlation between what is spoken and what is written is largely lost for us. The major fact of the regional and temporal particularity of the conventions underlying alphabetic scripts is obscured by habitual reflexes to what is now an extremely distant historical transition, far beyond the development of most of the phonological and syntactic conventions of the language we now recognise ourselves to speak. Much of the later history of relations between writing and speech focuses, instead, on questions of literacy, on access rather than on principles of correlation. Nevertheless, the suggestion has been made that questions of literacy of this kind still contain within them basic issues of script, in the likelihood, for example, that wherever the number of required characters for a writing system is too great for everyone to learn, then even a society's maximal aspiration can only ever be to a form of 'proto-' or 'oligo-' literacy (literacy in the hands of specialists)—with the more socially significant condition of general literacy confined to scripts with an easily learnable number of characters (Goody and Watt, 1960: 321-2; Ong 1982: 87-93).

In the present, questions of speech and writing, and listening and reading, emerge differently again, perhaps now most clearly in medium-dependent features of cultural and artistic forms, such as soundtracks of television and film, or records and tapes. Contemporary communications media compete with, and have to some extent displaced, written texts (the printed book, the telex and microfiche, latterly VDU display), utilising instead spoken forms and modes of interaction (telephones, intercoms, radio, tape). But whereas it is often noticed that information storage and retrieval is undergoing a major reorientation which prioritises the *visual* display of materials, an equally complex and yet virtually opposite transition can also be

development of individual languages *and* of individual speakers, is almost universally acknowledged as the proper object of modern studies of language. The focus of study of this kind might be on the differing relations between ideational and interactive functions in speech and in writing; or it might explore the distributions of deictics in speech and writing, marking their relative degrees of context-dependence; or again; it might seek to identify features in speech which reveal continuous and simultaneous monitoring and planning, compared with writing's invitation to reflective and intermittent forward and backward scanning.

(ii) in more theoretical speculation stemming from notions of the nature of the linguistic sign, and from Saussure's arguments for the primacy of *speech* over writing—ideas virtually inverted in the thinking of Jacques Derrida. What Derrida has queried in particular in Saussure is what he considers a 'phonocentrism', or fusing of voice and an imputed self-presence of consciousness when speech is considered primary over writing. In place of this, Derrida contends that speech too has all the decentered and conventionalised, often alarming, properties of writing; that spoken discourse is as riven by 'textuality' as written discourse. Much of the impetus in Derrida's work, and in deconstruction which is largely derived from it, is accordingly towards demonstrating ways in which the very freeplay of the system of language itself displaces speakers from any central or proprietary position over meaning.

(iii) in cultural-historical argument, especially in anthropology, over relations between 'orality' and 'literacy', often investigating social upheavals linked to changing relationships between them (the anthropologists Goody and Watt, for example, suggest that it is crucially writing which should govern the distinction between anthropology and sociology [Goody and Watt, 1960: 311] ). Studies along these lines might focus on what are called the 'psychodynamics' of orality and literacy; and in this context Goody and Watt in effect concur—though for very different reasons—with the view in Derrida that writing encourages 'scepticism'. What *they* argue is that orality's immediate dependence upon context and environment is displaced with the arrival of literacy into more comparative judgement that discloses division between present consciousness and past recorded experience (Goody and Watt, 1960: 352, 344 and 336-7). As regards the present in Western societies, too, such studies address general questions of the social transitions shaped by modern communications technology, identifying what is claimed to be the new condition of 'secondary orality', or the renewed social force and centrality of speech which is dislodging the traditional prestige of the written word.

### From Orality to Literacy

Ideas of an 'orality' which is '*secondary*' clearly depend on ways of seeing the earlier shifts from oral to literate, and the continuing and changing co-existence between the two. Within the broad range of approaches to what are for Western societies the earlier transformations (anthropological in Goody and Watt, in Ong, and others; sociolinguistic in Labov; literary in Milman Parry and Albert Lord), *three* channels can be distinguished; and it is worth recalling these as a first bearing on more recent movements.

(i) issues of *medium* itself, including necessary technology and institutions (printing, book collecting and selling, library provision, attitudes towards plagiarism and copyright, etc.). Related to these details of realization are properties of language production and processing, such as the massive extension in scale of argument as soon as written thought can extend beyond personal memory limitations, or the enhanced mutual comprehensibility between regions and classes following the emergence of a 'standard' written language with relatively regularised spelling and punctuation. From investigations of modern English, too, it seems that—subject to qualifying constraints of genre—the respective mediums of speech and writing tend to result in different distributions of linguistic elements: spoken language arguably less elaborately structured syntactically (many incomplete sentences, and chains of phrases: less subordination; few passives or cleft-sentences); in speech also a greater reliance on implicit connectives over formal markers of cohesion or logical structure, and more extensive use of prefabricated fillers, such as 'of course', 'if you see what I mean', etc. (see Brown and Yule, 1983: 14-9).

(ii) more general related *psychodynamics*, dependent indirectly on medium, as a result of the kinds of opportunity and constraint it presents. So long as speech is evanescent, the reasoning goes, and in default of possible appeal to authoritative definition from reference books or dictionaries, sounds are likely to take on magical or ritual potentialities; and for reasons combining ritualistic power and memory limitation, it is proverbs and narratives which serve best as the rhetorically mnemonic cultural or legal record (such narratives being revised to accommodate changed requirements of the history: for example by deleting branches of families from genealogies when they are no longer relevant, so producing what Goody and Watt have called a kind of 'structural amnesia' [Goody and Watt, 1960: 318] ). As is suggested by informant studies with syllogisms, too, experience widely takes priority in oral cultures over linguistic relations in developing or evaluating thought processes. And as regards vocabulary, Sapir-Whorfian notions

of cultural relativity in distinctions encoded within differing languages apply more obviously to cultures which have remained primarily oral than to literate ones, since oral cultures, lacking dictionaries, delete from the lexicon as well as create distinctions within it according to the criterion of current social usefulness.

(iii) still more general cultural *transition,* taken as a consequence of linguistic usage as well as of socially constitutive and reproductive forces of material inheritance, social practice and ideology. In this sense, medium and psychodynamics together reshape forms or genres, and redirect social habit and aspiration. The major claim at this level of argument is that literacy leads to diversification of, and contradictions within, previously homogeneous 'oral' cultures, as readers are differentially influenced by earlier stages of the cultural record, interpret them differently, and use them to support divergent versions of aspiration and intent. In Goody and Watt, this recognition of social contradiction presented by 'literary' forms is balanced against a framework of differentially operative socialising pressures in modern societies: for example, a weighted tension between idealisations offered in advertising and more conservative and realistic peer group and family 'oral' socialising traditions (Goody and Watt, 1960: 342).

## From Literacy to 'Secondary Orality'

The concept of 'secondary orality' has been introduced to refer to this newly-arisen and technologically-created primacy of speech over writing. But the social transition signalled by the concept has been linked to a range of social values (the McLuhanite utopia of global village communications; a rapid and continuing decline in reading standards and the emergence of 'tube' or 'pulp' culture; a restored sense of community which values immediate interaction and so has affinities with earlier oral cultures). Ong himself, who claims to have invented the term (Ong, 1971: 285), seems to exploit quite systematically an ambivalence between a lamentable unreconstructability of pure oral-cultural conventions, and a plain nostalgia *for* them (c.f. his remarks about 'pristine consciousness', Ong, 1982: 15).

Again, the social transition is prompted by a technological stimulus, as earlier with the development of moveable type. This time, it is the invention of radio, gramophone, telephones; magnetic tape; sound and image in film and television. Acknowledgement of the catalytic effect of machines need not commit thinking to technological determinism, however. Indeed, as has already been demonstrated in considerable detail with regard to television, film and other cultural forms such as popular music (see, Williams, 1974; Heath, 1981; Durant, 1984),

there is much to support the view that a complex set of relations are at stake between machines, institutions, and creative initiatives.

It is obvious, in any case, that secondary orality could never be simply a reversal. Quite apart from the fact of changed social circumstance, and co-existence with print, its major element of reproduction is *new*: the new media do not involve live transmission and telecommunications alone, but have established new forms of 'text'. And it is possible, in fact, to see this transformation reflected in changes now taking place in usage of the word 'live', when used to describe contemporary media performance. In presentations of pop or rock music, for example, notice the way that whereas formerly to have a performer 'live' on a programme would signify actually performing rather than miming, increasingly the term is being used to distinguish between 'live' performance as something *mimed in person* for the particular programme in question, and the screening of a pre-prepared promotional video.

To begin to describe the changing forms of presentation at issue in the concept of 'secondary orality', it will be helpful to identify more concretely the differing contemporary systems for relaying and representing speech. Taking up the idea of new forms of 'textuality' in modern media, perhaps the most important initial distinction is one between,

(i) what is for the listener 'impermanent' or 'direct, off-air' programming or transmission, evanescent and only to be heard at the time of broadcasting or telecommunications link,
and,

(ii) reproduced and reproducible sounds for repeated consumption at times chosen by the listener (audio tapes and cassettes, answerphone loops, videotape, etc.)

A second major distinction might then identify techniques used in producing and guiding speech (a distinction likely to be reflected in local features of discourse) —

(i) "SPONTANEOUS":
a) 'live' (simultaneously produced and transmitted), unedited, selected if at all by mixing only (e.g. telephone communication, live sports and news coverage, phone-in shows, etc.);
b) edited after production (involving both isolating or selecting material for use, and also editing out pause-fillers, etc., which are internal to the clip actually used).

(ii) "SCRIPTED":
a) learnt, rehearsed, memorized;
b) autocued at time of performance.

(Note: These systems are largely independent of one another, and it is quite possible to have, for example, highly rehearsed and edited material which is nevertheless spontaneous and evanescent for the listener; or, equally, tapes of unscripted and unedited material to be listened to as often as is enjoyable or tolerable).

Cutting across these systems for representing speech, however, are a number of issues which are not reducible either to properties of the machines themselves nor to any specific procedures for recording and production independent of notions of intention and priority. It is these which most of all suggest that investigation of contemporary media should not restrict itself simply to the *technical* properties of any particular representational format.

First, relationships governing speech differ fundamentally between speech and conversation conceived in terms of 'primary' oral interaction and speech considered as organisable 'textual' material. Whilst editing is rarely done by the same person who speaks, the impression of spontaneous discourse persists, with the editing process generally unperceived by listeners. Indeed, the problem generally acknowledged in discourse analysis—that study of speech out of context can distort what were *likely* interpretations in actual, contextualised use by massively increasing indeterminacy of meaning—here acquires a new and additional force, as recording, selection and editing not only dislocate from original context, but calculatedly provide important new contexts for comparison and inference.

Second, since editing is in this way a central resource for representation, new idioms are forged around expectations of unprecedented and often idealized 'appropriate' kinds of discourse. In radio interviews, for example, pauses and pause-fillers (such as 'er..' or 'hum...') are very often removed, creating an image of eloquence, knowledge, and authority personified by professional announcers, appointed spokespersons, academics, and media pundits. More general conventions of usage can be impelled towards change, too, as seems the case whenever edited interviews on television or radio *end* with the voice on a low-rising nucleus, as a result of an edit being made at a point of syntactic juncture which is nevertheless *intermediate* within a larger information structure. More than signalling the loss of a speaker's personal control over reported utterance, this usage takes place so frequently and influentially that it *may* be contributing to a longterm change in the language's systems of intonation.

Thirdly, kinds of attention likely to be given to actual speech and reproduced speech differ in ways not exhausted by the difference between transcription and original. Concretely, what distinguishes these is the precise inscription in recording, but not in conversation, of

inflections and intonations of the voice, movements exact y repre-
sented in a stabilised, reproducible form. One effect of this fixing of
details and anomalies in speech has been a partial redefinition of the
notion of 'recording', from the primarily *mnemonic* sense of preserv-
ing speech which is nevertheless still sponsored by and subordinate to
an earlier actual utterance, into the formative or *prescriptive* form of
creating new texts or works which will be known primarily or even
exclusively in decontextualised, reproduced form. Mechanical repro-
duction of the voice has in this way made possible new creative
opportunities (repeated takes, increased scrutiny and discussion dur-
ing rehearsal and production). New potentialities for listening may be
glimpsed on tape and disc, too, when listening is conceived as a
pleasure of anticipated precise repetition: knowing sounds, move-
ments, and vocal textures by memory and familiarity, rather than
responding to speech-input with processes of information reduction
that distil sensory information into comprehension and memory. Seen
historically, this consideration poses in a new way the traditional
question within certain genres of an aspiration to represent the human
voice as closely as possible (e.g. in cantabile phrasing in music, in
operatic recitative, or again, in poetry which seeks to be 'conversa-
tional', or to relay the 'real language of men'). Comparing such devices
with the fidelity of contemporary audio equipment, it becomes
unprecedentedly necessary to distinguish between, on the one hand,
audio representations of the voice amounting to a technologically
governed artistic *evolution*; and, on the other hand, the idea of an
aspiration to represent the human voice invested on the contrary in the
very enterprise of indirect *simulation*.

Fourthly, mechanical reproduction also leads in the direction of
effects which depend on *limitations* of the technology, such as the way
mishearing can lead to indeterminacies of meaning. Mishearing can
often occur in response to auditory input as the result of competing
homophonic sequences (more precisely partially homophonic sequen-
ces, to follow the terminology of Lyons and others; see Lyons, 1977:
558-69). These effects are in turn likely to be created by features of
continous speech such as by elision obscuring grammatical inflection,
or vowel neutralisation in unstressed positions; by instrusive and
linking 'r'; or by delayed articulation and kinds of assimilation. While
such mishearing is in any given situation *limited* by the fact that
(homophonous) lexical distributions rarely coincide, and by workings
of juncture (word divisions signalled by complexes of pitch, stress and
length as much as by actual pauses in the speech stream), what is
important as regards the new media is that in commonly-encountered
fast-tempo speech or melismatic singing under conditions of poor audio

fidelity, a range of provisional senses are liable to intervene in hearing the spoken text, putting unusually to the test the 'redundancy', or surplus of identifying cues which exist in spoken communication. Effects of temporary polysemy of this kind amount to more than mere stylistic felicity, and can in fact create a potent—and often psychically invested—sub-script which it is left to following sounds to cancel out. Indeterminacy in hearing may indeed suggest that disconnection and suspension of meanings can occasionally be as entertaining or enjoyable as the coherence of an utterance, indicating a whole new modality of imaginative response; and this would clearly perplex any theory of talk and listening which hypothesised a model of comprehension whose only difference from interpreting written discourse results from signal-evanescence and temporally-processive reception.

## Aspects of the Speaking Voice

A great deal of work exists in phonetics and phonology into the features of sound used in language, and the combinations these enter into. But much that is most interesting about contemporary communications media falls less into what has generally been considered the central area of research, than into the linguistic penumbra of prosodic and paralinguistic effect, and into aspects of speech outside even these (in 'non-linguistic' features, such as voice quality, or in what are often involuntary vocal reflexes, such as coughing, yawning, etc.).

At their simplest, prosodic systems are those, to follow Crystal in *Prosodic Systems and Intonation in English,* involving features of sound which have an essentially variable relationship to the words a speaker chooses, and thus standing in contrast to those features, such as a language's segments or individual lexical meanings, which have direct and identifying relationships to particular words (see Crystal, 1969: 5, and 126-7). It is not the case that this way of looking at prosodic systems inevitably gives *priority* to the segmental features, and indeed superimposing prosodic or paralinguistic features on a phrase or sentence in uttering it can even override the 'defining' qualities ('it wasn't *what* you said, it was the way you *said* it'). Moreover, Crystal observes that often it is prosodic and paralinguistic difficulties, such as mastering accent and attitudinal contrast, which tend to remain longest in language learning, and so receive most protracted attention (Crystal, 1969: 2).

Distinction can be made here between paralinguistic features and prosodic ones on the basis both of differing articulatory mechanisms and differing behaviour in speech. Paralinguistic features, such as the voice-qualifiers (whisper, falsetto) and voice-qualifications (laugh, giggle, sob), have the following properties:

i) they derive from mechanisms other than the vocal cords (e.g. by oral or nasal cavity variation);

ii) they tend to be less discrete than prosodic systems, and allow more idiosyncratic variation;

iii) they are phonetically discontinuous in connected discourse, and so appear as 'additions' to the normal flow of speech;

(iv) they involve a continuum of potentially infinite gradation in their parameters, while varying from a determinate norm which can be said to be characteristic of non-paralinguistic speech.

Prosodic systems, on the other hand, work with continuously operative dimensions of pitch, loudness and duration, with these generally encountered in some combined form (exemplary in the exceptionally stylised connection between pitch and tempo in horse-racing commentary). Yet whilst prosodic systems show a far greater degree of internal patterning and contrastivity than paralinguistic ones, they still have to be defined relative to speaker norms, with absolute judgements (for example about pitch-range or speed of utterance) tending to obscure the *perceptually* significant systems of contrast.

In English, the prosodic systems include,

(i) tone (type of intonational nucleus, such as simple, compound or complex);

(ii) pitch-range (movement within a specifiable pitch-band used for speaking, variable according to the tessitura of any individual speaker):

(iii) pause (filled or unfilled, of differing lengths);

(iv) loudness (loud/soft variation and crescendo/diminuer.do);

(v) tempo (fast/slow, drawled or held; accelerando-rallentando);

(vi) rhythmicality (rhythmic-arhythmic; staccato-legato).

**The Semiotics of Vocal Texts**

The way phonological features operate in speech to produce contrasts of various kinds is of great interest to linguists. But for analysing speech in modern media, emphasis needs to be given less to the properties of the systems in which such features occur than to the range and potential of their semiotic properties when incorporated into the broader repertoire of representational devices in the new modes of 'text'. To broach this integral function here, it will be helpful to identify some of the more controversial distinctions of meaning and value which surround three aspects of the voice: accent, intonation, and voice quality.

1) Accent

It has been noted that the written system of English underwent standardisation before the spoken, and indeed it is largely only in the course of the nineteenth century—mainly through the influence of English public school education—that one form of accent emerges as a non-regional prestige form, widely known as Received Pronunciation. By accent, what is understood here is a degree of systematic variation between regionally or socially distributed groups of speakers which is evident in their pronunciation but not in their written English. Differences of this kind are identifiable more specifically by reference to the number and incidence of phonemes used in speech, in the phonological make-up of those phonemes, and in their allophonic realization (or the way their phonetic detail is conditioned by local linguistic environment). Understood in these terms, Received Pronunciation itself is most closely connected with the dialect of English spoken in south-east England ("Educated Southern British English"), and in this respect the standardisation in speech follows the pattern established by the earlier standardisation of the written system, influenced as that was by the location of royal court, and political, legal and commercial institutions. Within Received Pronunciation itself, there remain several identifiable varieties. But even if this differentiation is taken into account, it is no longer reasonable to speak of Received Pronunciation as the dominant accent of the B.B.C. or other broadcasting organisations in Britain (though until recently its use for media purposes was defended on grounds of a special likelihood to be understood fully nationwide, and to be the accent most free from regional prejudice—class prejudice seeming to play no role in this debate).

Two areas in cultural forms where accent plays an important semiotic role are worth comment:

(i) In performances and video and audio recordings of Elizabethan drama, poetic recitation, and madrigals and ayres, widespread use is made of Received Pronunciation as a 'standard' accent in relation to which stylistically significant variations are organised (e.g. for comic or satirical effect).

What is interesting about this is the way the Elizabethan and Jacobean forms have become closely connected with images of a national culture associated with certain forms of class and education; and it is arguable that one component of this cultural position or status depends on the accents in which the works are produced and reproduced. It seems clear that the general use of Received Pronunciation in this way is at best problematic, however. It cannot be maintained that

the accent serves the purpose of historical authenticity, in the way that usages of replica instruments for early music do. English phonology has changed to the extent that 'authentic' reproduction entails substantial departures from most contemporary forms of pronunciation of English, and indeed can lead to difficulties of comprehension. Yet the other existing justifications seem equally suspect. Among these, for example, is the idea that the accent corresponds in terms of social connotation in the present to that of the Elizabethan classes who performed (simulating not the actual forms of the address, but its social *force* as a class voicing). But this view in fact seems to undermine much in the project of conservatism to which it appeals, neglecting both the regional and class variety in the English spoken in and around the Elizabethan and Jacobean courts, and the fact that attachments of social valuation to kinds of English were themselves during this period undergoing fundamental reorganisation (with a prescriptive consensus as regards pronunciation only emerging on any scale a century later). In retreat from the abstraction of this kind of argument, of course, there is the possible resolution that accents in modern performance do no more than reflect the particular people who are interested in performing the works. But even this appeasing view falters, when brought into comparison with acknowledgements of accent as a major force of representation within virtually all other kinds of performance, and one presently indispensable in all kinds of modern drama and song.

(ii) Such connotative properties of accent can be exemplified in new activities, too, for example in movements in accent and meaning surrounding the extensive simulation in British rock music of American forms of pronunciation, through almost all periods and idioms, but widely challenged and explicitly renounced in much punk and New Wave music following 1976.

Because of regional and social inequalities in power and prestige, indications of speaker-provenance—whether authentic or artificially targeted and modelled (see Trudgill, 1983: 141-60)—signal important cultural patterns of aspiration, resentment and influence. In consequence, reproductions of precise accent and vocal delivery can often elicit quite contradictory understandings of attitude and aspiration. While accents may very often *begin* as indication of local origin, personal identity, or social identification, they acquire through regional and temporal displacement new properties of representation, contributing to the conventions of particular cultural styles or forms.

In this sense, it is interesting to see that usages of American accents have effects which are regionally and socially variable. Not only is it the case that when authentic American accents were first heard on any

scale in Britain, in the dialogue of films of the 1930s the view became widespread that this mode of speech was unintelligible (Gimson, 1980: 90; Trudgill, 1975: 55); to this day, it remains the case that it is the continental contrast between 'British' and 'American', rather than finer intra-national, regional or social distinctions which have broadest semiotic currency. Moreover, when the accent is *simulated* by British rock performers, it leads in either of two, seemingly quite opposed, directions: on the one hand, towards revered American origins for rock and roll, through an imitative respect for 'authenticity'—taking accent as an integral constituent of the original rock and roll form; on the other, towards playing on connotations of a cultural devaluation associated with American cultural currencies relative to traditional, British cultural positions and investments. Of these conflicting directions, one uses accent to bolster a dependence on representations of social relationships elsewhere (and, as regards rock and roll origins, increasingly from another period); the other uses accent to challenge definitions of social relationships in an immediate environment. In both, accents operate as a dimension of form reliant on a technological condition (the precise reproduction of sound on tape and disc), but by changing their significance in altered contexts, show themselves to be equally imbricated within a network of surrounding currencies, connotations and evaluations.

Together, the two instances described here suggest that the orthodoxy, that distinctions created by accent are likely to disappear following increased exposure to radio and television, is unfounded. Rather, it seems that while accents may *alter* in relative prestige and semiotic function, contrasts between them will retain functional importance of one kind or another. The educational consequences of this view seem to be, then, that no pedagogy of egalitarianism of accent and dialect should be substituted for the objective of bi- or multi-dialectalism (for discussion, see Trudgill, 1975). Rather than seeking to abolish accent distinctions, pedagogy might pursue the more modest aim of contributing to the development of initiatives which use less prestigious accents to contest existing stereotypes.

2) Intonation

While it is generally agreed that intonation is very important in signalling meanings, there remains argument over the extent to which it is a part of grammar (for discussion and references, see Crystal, 1969: 42-4, 180-93). Without even engaging with the complexities of this theoretical argument, however, something of the range and importance of the cues which intonation can give when superimposed on a single written phrase or sentence can be glimpsed in Jakobson's fam-

ous anecdote, in 'Closing Statement: Linguistics and Poetics', about how the director Stanislavsky asked an actor from the Moscow Theater to make forty different messages from the two-word phrase, 'Segodnja vecerom' ("This evening"), simply by what Jakobson calls 'diversifying its expressive tint' (Sebeok, 1960: 354-5). Jakobson goes on to recall how, for the altered purpose of linguistic research, this task was replicated—with emotive cues for fifty imagined situations attempted by the voice—and how most of the messages were successfully decoded by Musovite listeners from a tape recording.

Intonation appears to serve several distinct functions:

(i) to differentiate between certain kinds of illocutionary act (questions, statements, etc.);

(ii) to mark grammatical, constructional relations, by separating speech into tone groups (changes in tempo also help to do this);

(iii) to contribute to information structure, by distinguishing between elements of discourse being treated as "given" and those being treated as "new" (this function overlapping with kinds of emphasis achieved through movement transformations, cleft and pseudo-cleft sentences, etc.);

(iv) to mark special emphasis, contrast or attitude.

Even relatively straightforward and theoretically non-controversial aspects of intonation, such as tonic placement, can mark nuances of meaning and attitude which are for a particular text quite crucial. Yet despite this, relatively little attention has been given to such issues, even where attention along these lines seems an indispensable component of analysis (as in studies of news bias, of radio, television or film dialogue, or of political rhetorics of all kinds).

To begin with, there are properties of factivity and contra-factivity carried by intonation (as well as by lexical items such as 'know', 'acknowledge'). Compare the two formulations, '/ Mr. SMITH says..../' ('but what Mr. Jones says is...'), and, '/ Mr. Smith *SA YS*.../' (But what is actually the case is...'). Both kinds of formulation are common, even in news-broadcasting, but they are certainly not interchangeable; and there are also intermediate cases between them—the distinctive idiom of newscasting seeming to involve exceptional techniques of 'blurred' or ambiguous tonicity.

There also exist far more complicated questions of co-reference and collocation cued by intonation, and the effects of these are equally crucial, though far less likely to be consciously recognized. Consider, "/ Mr. Scargill says the Miners' Strike might lead to a General E-LECT-ion;/ Mrs. THAT-cher says the country WON'T be held to ransom./"

One effect of this formulation is clearly to contrast Mr. Scargill's view with Mrs. Thatcher's, hence in the recording a shallow fall-rise movement on 'Mrs. THAT-cher', indicating a contrast between the two speakers. But a far more major, and arguably contentious, effect is created by the occurrence in the second clause of the nucleus or tonic syllable on 'WON'T, rather than on 'ransom'. This is not simply because when tonicity occurs on a modal auxiliary it sets up a contrast around the polarity of that modal (in case, for example, anyone expected that the country 'WILL'...). More than this, the markedly preposed tonic (brought forward from what would be its unmarked position, on 'ransom'. the 'last new, information-carrying lexical item'), signals that the unstressed tail which follows ('be held to ransom') is being treated by the speaker (in this case, the newcaster) as 'given information'. Such given information is given either by context or by co-text: either it is information presupposable from the situation and an assumed frame of shared belief, or else it is derivable from the preceding text itself (by being co-referential with something occurring earlier). In this instance, it is clear that the way the discourse is organised *contributes* to the framework of inferencing, in that the phrase 'held to ransom' is treated as given exactly insofar as it is co-referential with the earlier, entire proposition 'the Miners' Strike might lead to a General Election'. The overall effect is one of the two propositions seeming to exist within, to have been 'preconstructed' as, the same semantic set, so establishing a form of equation between them. This effect is created because, in general, information structure signalled by intonation is dependent upon collocational likelihoods, which are determined in turn by semantic sets. So whenever an information structure of this general pattern is perceived, its most likely interpretation will be the one achieved by imputing a semantic identity between items in the structure, drawing them into, or implying, semantic connections between them simply by virtue of the conventions of the information structure itself.

These features of intonation run through all conversation, of course, but gain particular importance when they are embodied in texts that will be heard repeatedly, or which are transmitted as a kind of authoritative discourse, sanctioned by the institutional influence and controls of radio, film, and television. Or again, they can gain peculiar prominence when vocal mannerisms (such as those which collectively signify certitude or assurance) are replayed in situations where any sponsoring context for, or appropriateness of, the certitude is unwarranted. No neutrality of voice in respect of features of intonation and information structure is possible in speech, of course (hence, perhaps, something of the tortuous complexity of news-broadcasting especially).

And for precisely these reasons, it seems an educational priority to alert listeners as fully as possible to the *range* of aspects of speech through which persuasive effects are created—effects as crucial to radio, television and film coverage as the more disputed issues of time-apportionment or camera angle.

## 3) Voice Quality

The term voice quality is used to refer to the impression of a single voice (or underlying background setting) which runs through the whole of a speaker's normal utterance, excluding such variations as deliberate mimicry (see, Laver, 1980). Slightly more technically, it can be characterised as the permanent, non-segmental idiosyncratic factor in a person's speech ('gravelly', 'harsh', 'creaky', 'husky', etc.), a feature which is deemed to be non-linguistic, in that it is contextually random, appearing to have the main function of identifying individual speakers. Voice quality is created by means of, (i) unevenness and irregular movement of vocal folds (aperiodic vibration), as in harsh or hoarse qualities; (ii) additional noise, as in breathy voice; and (iii) unusual resonances, as in throaty, hollow or thin qualities. And its close connection with a speaker's body and automatic neurological and muscular operations of speech can be seen especially clearly in the often discordant effect of film and television soundtracks in the many countries where dialogue is recreated to accompany the image by fresh over-dubbing.

The issues provoked by voice quality in performed and reproduced texts are not merely ones of taste, however. In styles of speech and singing, unusual textures or grains of voice have been widely experimented with, and much enjoyed. But beyond offering accentuatedly clear speaker identification (and so possibly contributing to an image of 'personality'), such inscriptions of the voice carry distinctions of meaning and value alongside aesthetic pleasures; and it is this aspect of them which most clearly displays the theoretical problem of dividing up the continuum which runs from individually idiosyncratic features of voice, through voice stereotypes (such as 'plummy', 'sexy', etc.), to what is generally known as regional, occupational or class accent. Consider rock and popular music as example once again. Widespread use has been made of extended or improvised vocal techniques of 'strained' voicings and conventionally unaesthetic or dissonant singing, creating new orientations of social position and value through implied rejection and counter-identification; and it is such distinctions around meaning and value which cut through aesthetic dimensions of vocal performances by Bob Dylan, Marvin Gaye, Joni Mitchell,

Robert Wyatt and many others, establishing kinds of stylistic contrast through connotations of the respective voice qualities.

Moreover, there are complexities within the *aesthetics* of the voice itself, including issues surrounding an empathetic, or erotic participation or attraction in speech and singing (Roland Barthes refers to an 'identificatory fantasy of performance' which overlaps with the more traditional, phonetician's conception of articulatory empathy, as outlined by Abercrombie [see Barthes, 1977: 149-54; and, Abercrombie, 1965: 16-25] ).

In speech and singing, body and meaning intersect, and interest in the voice divides between meanings and a level of sounds as primarily acoustic material, while properties of vocal sonority interlock with points of social circulation and reference. It is largely by reason of this intersection that performance of operas in the original language or in translation has remained a crucial debate: whether the language of the libretto is made up mainly in dramatic meanings translatable between languages, or in a patterning of sounds, clausulae, and inflections peculiar to a language of original conception.

Most often, the arguable surplus of forms over meaning, and related freeplay of meanings across given forms, has been valued in assessments of written forms, though, as has been pointed out above, Jacques Derrida and others have argued for the existence of an analogous condition of textuality in *speech*. Perhaps the most detailed consideration of the play between body, subjectivity, and meaning in the sounds of the voice has been given by Roland Barthes, in the essay, 'The Grain of the Voice' (Barthes, 1977: 178-89). Contrasting the voice in that essay as bodily sound generator, and as assured mechanism of expression, Barthes proposed a distinction between two different kinds of voicing, a distinction formulated in terms of a taste for the singing of Panzera, and a lesser enthusiasm for that of Fischer-Dieskau. It was the proposed importance of this distinction that a certain kind of vocalisation, in which processes of enunciation are prominently displayed, has disappeared or is disappearing under pressure from a kind of singing in which the body finds itself replete with cultural meanings, individually and socially represented. (The actual point of aesthetic division is between an expressive or affective voicing—meaningful pauses and breathing, attention to stylistic or generic detail, what Barthes, acknowledging Julia Kristeva, calls 'phenotext'—and a kind of voicing with 'grain'—voicing of the body in language, beyond particular cultural meanings or intelligibility, the voice eroticized, for Barthes 'genotext'.)

Linking questions of voice quality both with concerns of subjectivity as something organised in movements of language on the body, and

with aesthetic forms and criticism, Barthes' essay has arguably been as influential as other, primarily descriptive, phonational investigations conducted within linguistics. And to distinguish between kinds of pleasure in performance and listening has been important. But while vocal performances can indeed be divided between Barthes' categories, the condition of enunciation to which listening might aspire is always also social. The 'genotext', and the bodily history it speaks, can only exist within wider circumstances. The central problem Barthes' essay introduces into cultural criticism is less one of disengaging the 'genotextual'—hearing everywhere new graduations of inflection or finenesses of performance—than of looking at the ways in which the conjunctions and contradictions between the two vocal axes have been managed in the past, in order to reflect on them in the present. What can be gained is seeing the 'genotextual'—the aesthetic or erotic aspect of vocal production—as a constituent of contemporary artistic activity, while still acknowledging that such activity is always also social both in its forms and in its changing and changeable limits.

### Speech, Media and Education

This discussion of three areas of semiotic interest and potential controversy in speech does not finally have very much to say about how texts within modern communications media are likely to be interpreted in any given instance. Rather, what it indicates is local areas of concern which problematise medium-dependent properties of language, and in doing so overspill definitions of media technology while falling short of the anthropological generality of the concept of 'secondary orality'. Nor is it the case that, if textual criticism of modern cultural forms were extended to incorporate a much extended semiotic and kinesic machinery, this would lead to an exhaustive, overall 'communications text-grammar'. Rather, what seems clear is that each area involves socially *competing* or contradictory claims on interpretation; and while analysis in any one instance must certainly make appeal to the specific detail of medium, it must also refer to ideas of intention and style, *and* to circumstantial patterns and pressures of social structure. Consideration of a new 'communications rhetoric' cannot be limited *either* simply to technical specifications of the machines, *nor* to a listing of tropes and figures: rather, it must seek to offer accounts of particular initiatives by cutting between several concurrent levels: beyond questions of medium themselves, there are issues of technology and of institutions governing the application or use of that technology; beyond these, there are cultural forms or idioms, which *embody* within themselves the technical and institutional constraints of the machines; beyond these again, there are the

arguable psychodynamics, cultural ramifications and 'erotics' or sexual dimension of the new forms—elements which are *called into play* when speaking and listening articulate the body as well as social experience and linguistic convention.

Clearly there is a need for educational thinking to respond to the cultural repercussions of technical redefinition and transition. Yet historically, it seems it is *idealised* versions of past transitions which have tended to dominate modern perceptions, depriving us of important frameworks of reference and precedent. It is a well-established habit, for example, to locate foundations of modern Western culture in the writings and recorded thought of the ancient Greeks, alongside religious inheritances from Hebrew works and traditions. But there is a quite concrete reason, in addition to the undoubtedly important mythological inheritance, for taking these two cultural precedents as primary historical bearings: a major linguistic determinant alongside the more unaccountable cultural ascendancy (e.g. 'the genius of the Greeks'). The Greek and Hebrew civilizations to which appeals in respect of cultural origins are so often made coincide with the first consolidations of phonetic representations of speech as alphabetic writing—with, that is to say, the first uses of writing to transcribe speech rather than represent objects in the world. Yet the crucial material circumstances and conditions of this important transition— whereby the formulaically-composed Homer becomes the patron of subsequent Western Literature as well as some of the first actually written 'literature' (see Lord, 1960)—are regularly obscured by a distilled image of idealised tradition and cultural priority (for discussion, see, Goody and Watt, 1960: 330-8).

So too with cultural upheavals associated with literacy, and with historical resistances to it. Such upheavals involve major problems, including *access* to written materials, including social acquisition of an adjustment to the 'psychodynamics' of literacy, and including the development of *analytic* capacities for reflecting on the creation and control of meanings; but very often these pressing realities are submerged in forms of cultural study which prioritise a line of authors and works over the circumstances of printing, readership, patronage and income which made them possible.

In the present, again, attacks on film, television, and radio—as at best trivial and at worst degenerative when compared with Literature— now co-exist with defences of these same media as sites of great artistic achievement. But for so crucial a comparison between cultural forms, it is important to realise that terms of personal preference, stylistic comparison and established canon will always be too narrow. Simply enlarging traditional principles of literary criticism, such terms obscure

the fact that 'full literacy' in respect of the new media requires a degree of understanding of conditions of production and control that runs far beyond any definition of autonomous conventions, artifice or codes.

It would take much more than the few observations gathered here to argue fully the case for a reorientation of educational priority towards work with speech. In any case, that reorientation should not be isolated as a conclusion from other reforms which appear to follow from the same and related arguments. Nor is it possible to *predict* with any confidence the outcome of the processes of contemporary change surrounding the new 'secondary-oral' communications media, and future developments would need to be included in any formulation of educational policy as they happen. But it is worth at least remonstrating against what is a complacent notion of the second coming of interactional 'communicative-transparency'. Quite contrary to this, it seems that the only solid basis for developing skills with and responsiveness to language lies in widely encouraged comparative analysis and experiment; and for these to have any leverage on texts in the new sound-media, massively increased familiarity with the semiotic properties of speech is a crucial but often undervalued educational imperative.

REFERENCES

Abercrombie, David (1965), 'A Phonetician's View of Verse Structure', in *Studies in Linguistics and Phonetics* (Oxford University Press, London), 16-25.
Barthes, Roland (1977), *Image-Music-Text*, essays translated by Stephen Heath (Fontana, London).
Brown, Gillian, and Yule, George (1983), *Discourse Analysis* (Cambridge University Press, Cambridge, England).
Crystal, David (1969), *Prosodic Systems and Intonation in English* (Cambridge University Press, Cambridge, England).
Durant, Alan (1984), *Conditions of Music* (Macmillan, London).
Gimson, A.C. (1980), *An Introduction to the Pronunciation of English* (Edward Arnold, London, third edition).
Goody J. and Watt I. (1960), 'The Consequences of Literacy', in Pier Paolo Giglioli, ed. *Language and Social Context* (Penguin, Harmondsworth, 1972, reprinted 1980), 311-57.
Heath, Stephen (1981), 'The Cinematic Apparatus: Technology as Historical and Cultural Form, in *Questions of Cinema* (Macmillan, London), 221-35.
Lord, Albert B. (1960), *The Singer of Tales* (Harvard University Press, Cambridge, Mass.).
Lyons, John (1977), *Semantics*, 2 vols (Cambridge University Press, Cambridge, England).
Ong, Walter (1971), 'The Literate Orality of Popular Culture Today', in *Rhetoric, Romance and Technology: Studies in the Interaction of Expression and Culture* (Cornell University Press, Ithaca and London), 284-303.

_____ (1982), *Orality and Literacy: The Technologizing of the Word* (Methuen, London).
Sebeok, Thomas, ed., (1960), *Style in Language* (M.I.T. Press, Cambridge, Mass.).
Trudgill, Peter (1975), *Accent, Dialect and the School* (Edward Arnold, London).
_____ (1983), 'Acts of Conflicting Identity: the Sociolinguistics of British Pop-song Pronunci-
ation', in *On Dialect* (Blackwell, Oxford).


Williams, Raymond (1974), *Television: Technology and Cultural Form* (Fontana, London).