Phonetics of talk-in-interaction: a study of the interactional functioning of phonetic detail in everyday talk.

Almost everything we do that concerns other people involves talk-in-interaction. The purpose of this research is to advance understanding of how conversation works. It will investigate the ways in which speakers and listeners make use of the phonetic (sound) resources of language to shape and interpret talk in natural conversation. Linguists have a wealth of knowledge about the way speech sounds are produced and combined together to make words and longer utterances. However, they frequently work on the basis of constructed data or speech produced in experimental settings. In consequence, there is only the most rudimentary information on the ways in which ordinary people use the phonetic resources of language in natural everyday talk. Though we know something of the way the pitch of the voice is used, for example, in asking questions, giving orders or making statements, we lack basic knowledge about the *interactional* work done by other phonetic features such as, tempo, loudness, or the precise way particular sounds are articulated. As a result, we have little understanding, for instance, of how speakers signal that they have made a mistake and are now correcting it, or how they signal that they are willing to yield at turn-at-talk, or why it is that conversationalists treat some talk, which overlaps their own, as interruptive but other overlapping talk as supportive.

The proposed research develops a novel way to remedy this lack of information. Rather than working on constructed data or speech taken out of context, it will study the phonetic detail of how talk is shaped in real conversation. In particular it will investigate:

- how mistakes are signalled and corrected;
- how misunderstandings are clarified;
- how talk which overlaps or completes that of others is shaped and interpreted.

These activities are chosen because they are frequent in natural conversation and they provide rich opportunities to investigate the ways in which participants design their own talk and interpret the talk of others on a moment-to-moment basis.

The data used will consist entirely of audio and video recordings of natural conversation. The investigation will examine the full range of phonetic resources (rhythm, tempo, duration, loudness, pitch, voice quality and the articulation of sound segments) and their individual and combined roles in signalling interactional meaning.

The results of the research promise to be far-reaching. They will inform the work of linguists, sociologists, and psychologists. The development of an *interactional* approach to phonetics will challenge current theories of language use and language understanding. There will be practical benefits, too. The results will be a resource for researchers in computer speech recognition and for speech therapists and other practitioners concerned with the development or remediation of communication skills.

Phonetics of talk-in-interaction: a study of the interactional functioning of phonetic detail in everyday talk.

1. Overview of the research proposal

The purpose of the research is to advance understanding of the ways speakers and listeners use clusters of phonetic parameters in shaping and interpreting talk in natural conversation, with a view to establishing a 'phonology of conversation'. Current phonological analyses of conversation and discourse focus almost exclusively on intonation at the expense of other phonetic aspects and typically trade on analysts' own intuitions in determining the function of phonetic events. The proposed research implements a novel approach which overcomes these limitations. It will employ detailed auditory and acoustic phonetic analysis and the techniques of sequential interactional analysis developed by researchers in Conversation Analysis (CA) to document the *interactional* functioning of *general* (as opposed to simply intonational) phonetic parameters. Following the methodological maxims of CA, particular attention will be paid to the need to justify the proposed analyses by showing that the postulated categories are relevant to the participants themselves.

The outcomes of the research promise to have far-reaching consequences. They will provide a radical reorientation of our understanding of how speakers and listeners use the phonetic resources of their language. The shift to an *interactional phonetics* will impact not only on the conduct of the linguistic analysis of everyday language, but also on the ways in which phonological, sociolinguistic and psycholinguistic theories may be constructed and tested. The research findings will thus make an important contribution to the development of theories of language understanding. There will be practical pay-offs, too, for speech therapists and other practitioners concerned with the development or remediation of communication skills.

2. Background and Statement of Problem

2.1 Motivation for research

Remarkably little is known in detail about the phonetics and phonology of spontaneous speech. Virtually nothing of interest is known of the *interactional* implications of particular kinds of phonetic events in everyday talk: in particular about the ways in which participants in talk deploy *general phonetic parameters* to accomplish specific interactional activities. The principal reasons for this are that

- linguists typically continue to describe the structure of language divorced from its natural site of occurrence — the everyday talk of ordinary people. In the syntactic and phonological domains, especially, this procedure has been elevated to the status of a research paradigm (Abercrombie, 1965a; Schegloff, 1996);
- linguists working on 'discourse phonology' have typically focused on pitch and intonation at the expense of other phonetic parameters (e.g. Coulthard & Brazil, 1981; Hirschberg & Pierrehumbert, 1986; Levelt & Cutler, 1983; Menn & Boyce, 1982; Schiffrin, 1994);
- linguists typically rely on their own analysts' intuitions (or native speaker acceptability judgements) in setting up and explicating functional categories. In consequence these categories are notoriously problematic when applied to unscripted, spontaneous speech (see Campbell, 1995; Couper-Kuhlen and Selting, 1996: Docherty et al., 1997; Ford and Thompson, 1996; Sankoff and Brown, 1976; Schegloff, 1989, 1996).

2.2 Methodology

The proposed research programme addresses the limitations outlined above by undertaking an *interactionally-grounded* analysis of the phonetics and phonology of everyday talk. One innovative aspect of this line of research is that it integrates detailed phonetic analysis with the rigorously empirical methodology of Conversation Analysis. In consequence it differs from current approaches to the functioning of phonetic parameters in speech in four theoretically important respects:

- the data derives entirely from naturally occurring conversational interaction;
- the approach is one which seeks to locate and identify specific interactional activities and to state the general phonetic parameters which speakers use to accomplish them;
- the approach takes it as axiomatic that for a phonology of conversation, it should be these *interactive* categories which provide the basis for the analysis and such categories must be arrived at from, and grounded in, the data. These categories must be shown to be relevant to the participants in their talk and not be derived ultimately from the analyst's intuitions as a speaker of the language under analysis;
- we demand that the analysis prejudges a little as possible the relevance of particular phonetic details and particular phonetic parameters.

Conversation Analysis

The methodological approach to the analysis of interactive categories I have just outlined has been developed by workers in ethnomethodological discipline of Conversation Analysis. CA researchers have demonstrated by careful sequential analysis of interaction that participants display, in their language behaviour, systematic orientation to features of the talk and that this

systematicity provides a basis for interactive categories (see e.g. the papers in Atkinson and Heritage, 1984; Sacks, 1992; Sacks et al, 1974; Schegloff, 1979, 1982, 1984, 1991, 1992, 1996). It has not been the primary concern of these analysts to state the linguistic exponents of their categories — a task which falls naturally to linguists rather than sociologists. Their main concern has been to explicate the competencies social participants draw upon in producing, understanding and co-ordinating interactional behaviour.

In addressing these issues CA has maintained a rigorously empirical approach to analysis. First, it has required that any analytic claims about social interaction be validated by means of 'participant orientations'. That is, the analysis proposed must be tied to, and grounded in the observable behaviour of participants in the interaction. This stringent constraint reflects an endeavour to make analytic claims commensurate with a participant's analysis. Second, CA has also insisted on the importance of 'sequential' analysis of interaction. The actions which are embedded in conversation take place in sequences, they occupy (particular) positions within sequences and their sequential position is a crucial determinant how such actions are understood and dealt with by co-participants in conversation. The organisation of everyday talk is such that is proceeds on a negotiated turn-by-turn basis. One consequence of this organisation is that any next turn provides an opportunity for its producer to display an understanding or analysis of the prior turn. This is an important resource not only for participants themselves but also for analysts trying to make sense of how talk is functioning. CA thus has important methodological implications for all studies of spoken language in that it provides an interactionally-grounded approach to analysis which can free analysts from traditional reliance on their own intuitions.

Phonetic detail

One key outcome of CA research has been the finding that 'no order of detail can be dismissed, *a priori*, as disorderly, accidental or irrelevant' (Heritage, 1989: 22). Previous work on the linguistic analysis of everyday talk conducted by the applicant has demonstrated that this is certainly true of phonetics in interaction. This work has documented some of the ways that participants in talk systematically manipulate clusters of *general phonetic parameters* (Abercrombie, 1965b) — encompassing rhythm, tempo, loudness, pitch, voice quality, and independent articulatory parameters — in order to structure their contributions to interaction. Different combinations of segmental and prosodic phonetic parameters are used to signal the beginnings, continuations, restarts and endings of turns, and the interactional relevance of other stretches of talk: we have demonstrated, for example, that speakers make strategic use of tempo in combination with particular articulatory and phonatory parameters preceding silence to indicate whether or not they

are going to yield their turn (Local & Kelly, 1986); that different interactional consequences follow depending on whether articulatory and/or laryngeal (e.g. glottal closure) is held or not during intra-turn 'silences' (Local & Kelly, 1986); that speakers systematically manipulate rhythmic, loudness and pitch parameters to display whether overlapping speech as designed to be interruptive or not (French and Local, 1983); that speakers select different vowel, pitch and phonatory parameters to distinguish news-tokens ('oh') which are produced in response to out-of-the-blue informings as compared those which follow question-solicited informings (Local, 1996), and that the fine temporal phonetic detail of the articulation of specific sound segments, such as aspiration, and phonation can signal turn transition (Local, Kelly & Wells, 1986; see also Docherty et. al. 1997).

Phonetic parameters such as pitch, rhythm, phonation on one hand, and vocalic and consonantal quality on the other have been traditionally allocated to different, independent phonological systems (prosodic, segmental). One theoretically important result of our work demonstrates that these parameters are best treated as falling into *functional clusters* on the basis of how speakers deploy them to achieve particular interactional goals. If this is done, it becomes possible to document systematically the ways in which speakers and listeners manipulate phonetic parameters in managing the moment-to-moment flow and interpretation of ordinary conversation (Couper-Kuhlen & Selting, 1996; Kelly & Local, 1986; 1989: 30-35; 65-91; 263-286; Local, 1986; 1996)

Integrating CA and detailed phonetic observation

Sacks, Schegloff and Jefferson show how the sequential organisation of conversation can be appealed to for the justification of analyses (1974: 729). Through an inspection of turns at talk directed at a prior turn's talk analysts can access participants' displays of understanding; the latter serving as justification or 'warrants' for analytic claims. The following simple example gives a flavour of the way in such a CA-type sequential analysis can be combined with detailed phonetic observation to justify the functional properties of particular clusters of phonetic parameters. The data is drawn from an investigation of dialect-word recognition in the Tyneside Linguistic Survey (Local 1986). The hypothesis tested is that speakers use particular durational and voice quality features, with a rapidly falling pitch contour (from high to low) and marked loudness on a repeat following interviewers' proffering of a word for recognition. Repeating the word in this sequential position, with this cluster of phonetic parameters counts as doing an understanding check — it is designed to verify that the word has been heard correctly:

Understanding check (McN is the interviewer, ER the interviewee; transcription is given in modified conventional orthography; pauses, indicated in round brackets, are timed in tenths of seconds; the repeat turn of interest is arrowed; a simplified description of the relevant phonetic parameters are given in square brackets)

1 **McN:** er (0.2) varnigh

(3.5)

→ **ER:** varnigh [non-breathy phonation, no extended syllable duration, high falling pitch to low over second syllable, marked loudness over whole word]

(0.4)

5 **McN:** aye (1.5) you know (0.1) for nearly

(1.5)

ER: w I've never heard it I've heard me mother use it

The argument in support of the hypothesis is based on sequential evidence: the following turn by the interviewer routinely contains some acknowledgement or confirmation token (here *aye* in line 5), and the interviewee does not offer any sign of recognition until this acknowledgement is given (as here in line 7 of the transcript). Moreover, when the hypothesised phonetic characteristics are not present (e.g. when the speaker's re-doing of the word has breathy voice, is noticeably quiet, has extended duration or the pitch fall stops mid-way or does not begin high) the sequential treatment is different. In the latter cases, rather than immediately confirming or disconfirming the speaker's rendition of the word, the interviewer routinely witholds a turn until the speaker has offered some appreciation of it. I argue that this cluster of phonetic parameters signals that speakers are 'mulling over' the word. Compare:

Mulling over (McN is the interviewer, GSh the interviewee)

1 **McN:** er (0.8) varnigh

(0.9)

→ **Gsh:** varnigh [quiet, breathy phonation, noticeably extended duration of second syllable, falling pitch from mid]

(1.0)

5 oh yes I've sometimes said varnigh

McN: aye (0.1) uh (0.1) yeah

In this manner the interactional activities and phonetic categories which the analysis proposes are validated by demonstrating that participants in the interaction orient to them in predictable ways.

3. Specific Research Objectives

The proposed research has three key objectives:

- to identify and analyse the clusters of phonetic parameters that participants use in the management of interactional activities in natural conversation;
- to provide an interactionally-situated account of the relationship between the clusters of phonetic parameters and the morpho-syntactic linguistic units employed to construct turns at talk;
- to develop an analysis of the ways in which participants employ the clusters of phonetic parameters in ascribing meaning and intention to other participants' actions.

I These activities provide unique opportunities to investigate the ways in which participants shape their own talk and interpret the talk of others on a moment-to-moment basis. They are highly recurrent in spoken interaction, are readily identifiable and offer opportunities for examining a range of phonetic parameters and interactional activities. Moreover, they have been the subject of systematic investigation within CA (e.g. Jefferson, 1974; Sacks et al., 1977; Schegloff, 1979; 1982; 1987; 1992 Lerner, 1989; 1991; 1996) allowing results of the proposed research to be compared with those which have emerged from these studies.

n developing the analysis a number of specific hypotheses will be tested:

- there are particular clusters of prosodic and segmental phonetic parameters which speakers employ to signal the domain and structure of conversational self-repair;
- speakers use particular clusters of prosodic parameters to signal different kinds of self-repairs (e.g. 'disguised' self-corrections);
- speakers use particular clusters of phonetic parameters to display the interactional status of 'next-turn-repair initiators' (Sacks, et al., 1974) — particularly the rhythmic placement of the initiator relative to the previous turn;
- there are distinct clusters of phonetic parameters which mark particular points in talk-in-interaction as vulnerable to overlap and collaborative completion;

4. Data

The data base for the proposed research is composed of an extensive collection of audio and video recordings and transcripts of naturally occurring interaction made by the applicant and co-workers in the Department of Language and Linguistic Science, University of York, the Department of Human Communication Science, University College London and by other workers in Conversation Analysis in the UK and the USA (archived in the Department of Sociology, University of York). A range of speakers are represented (in terms of age, sex and social class)

and a range of varieties of British and American English (including a number of non-standard varieties — Belfast, London Jamaican, Tyneside, West-Midlands, Yorkshire).

5. Dissemination

The results will be incorporated into the proposed book, in the form of case studies of the ways in which speakers and listeners manipulate phonetic parameters in managing the moment-to-moment flow and interpretation of everyday conversational interaction.

References

Abercrombie, D. 1965a. Conversation and spoken prose. In Studies in Phonetics and Linguistics. Oxford: Oxford University Press. 1-9.

Abercrombie, D. 1965b. Parameters and phonemes. In Studies in Phonetics and Linguistics. Oxford: Oxford University Press. 120-124.

Atkinson, J.M. & Heritage, J. 1984. Structures of Social Action. Cambridge: Cambridge University Press.

Campbell, N. 1995 From read speech to real speech. In K Elenius & P. Branderud (eds.) Proceedings of the XIIIth International Congress of Phonetic Sciences. Vol. 2. 20-27.

Coultard, M. & Brazil, D. 1981. The place of intonation in the description of interaction. In D. Tannen (ed.). Georgetown Round Table on Languages and Linguistics 1981. 94-112. Washington, D.C.: Georgetown University Press..

Couper-Kuhlen, E. & Selting, M. 1996. Introduction. In E Couper-Kuhlen and M. Selting (eds.) Prosody in Conversation: Ethnomethodological Studies. 1-56. Cambridge: CUP.

Docherty, G., Foulkes, P., Milroy, J., Milroy, L. & Walshaw, D. 1997 Descriptive adequacy in phonology: a variationist perspective. Journal of Linguistics. 33, 2. 275-310.

French, P. & Local, J. 1983. Turn-competitive incomings. Journal of Pragmatics 7. 701-715.

Heritage, J. 1989. Current developments in conversation analysis. In D. Roger & P. Bull (eds.). Conversation. Clevedon: Multilingual Matters.

Hirschberg, J. & Pierrehumbert, J. 1986. The intonational structuring of discourse. Proceedings of the 24th Meeting of the Association of Computational Linguists. 136-144. New York.

Laver, J. 1994. Principles of Phonetics. Cambridge: Cambridge Univerity Press.

Levelt, W.J.M. & Cutler, A. 1983. Prosodic marking in speech repair. Journal of Semantics. 2. 205-217.

Local, J. 1996. Some phonetic aspects of news receipts in everyday conversation. In E Couper-Kuhlen and M. Selting (eds.) Prosody in Conversation: Ethnomethodological Studies. 177-230. Cambridge: CUP.

Local, J. & Kelly, J. 1986. Projection and silences: notes on phonetic and conversational structure. Human Studies 9. 185-204. 1986

Local, J., Kelly, J. & Wells, W. 1986. Towards a phonology of conversation: turn-taking in urban Tyneside speech. Journal of Linguistics 22: 2. 411-437.

Sacks, H., Schegloff, E.A., & Jefferson, G. 1974. A simplest systematics for the organization of turn-taking for conversation. Language 50: 696-735.

Schegloff, E.A. 1982. Discourse as an interactional achievement: some uses of 'uh huh' and other things that come between sentences. In D. Tannen (ed.) Georgetown University Round Table on Languages and Linguistics. 71-93. Washington, DC.: Georgetown University Press.

Schegloff, E.A. 1984. On some questions and ambiguities in conversation. In J.M. Atkinson and J. Heritage (eds.). Structures of Social Action: Studies in Conversation Analysis. 28-52. Cambridge: Cambridge University Press.

Schegloff, E.A. 1989. Reflections on language, development, and the interactional character of talk-in-interaction. In M.H. Bornstein and J. S. Brunner (eds.) Interaction in Human Development. 139-153. Hillsdale, New Jersey: Lawrence Erlbaum.

Schegloff, E.A. 1991. Conversation analysis and socially shared cognition. In L. Resnick, J. Levine and S. Teasley (eds.). Perspectives on Socially Shared Cognition. 150-171. Washington, DC: American Psychological Association.

Schegloff, E.A. 1992. Repair after next turn: the last structurally provided defense of intersubjectivity in conversation. American Journal of Sociology 97: 5 (March 1992), 1295-1345.

Schegloff, E.A. 1996. Turn organization: one intersection of grammar and interaction. In E. Ochs, E.A. Schegloff and S.A. Thompson (eds.). Interaction and grammar. Cambridge: Cambridge University Press. 52-133.