BRITISH ACOUSTICAL SOCIETY

SPRING MEETING 5th-7th APRIL, '72.
University of Loughborough

'SPECIAL LECTURE'

Conversation and Mood - an Experimental Approach

Prof. Colin Cherry, D.Sc.,
M.I.E.E.,
Imperial College, London, S.W.7.

In this lecture I shall make no description of finished work, no account of completed researches, so I hope you will not think my remarks out of place at this symposium. All I aim to do is first to take note of various ways by which people in different disciplines are approaching the many problems of human speech; second, to offer some criticism; and third, to press for experimental studies to be made increasingly natural and to be moved away from the artificial restrictions which laboratory conditions necessarily so often impose, largely owing to the traditions of different scientific disciplines.

At present there is no single discipline of "human communication". It is a most diverse field which interests people of very different traditions and scientific methods. Yet increasingly we need to make judgements and assessments which call for some understanding of several disciplines. For example, problems of providing efficient aids for people with sensory defects - the deaf, the blind, the deaf-blind - are not solved by physical understanding alone, because assumptions have to be made about their actual needs. Similarly, even such an apparently mundane problem as providing an efficient telephone service calls for far more than study of speech spectra, voice levels, probabilities and other measurables. It raises very difficult questions of definition, specification and assessment. What elements of speech sounds are important? How important? Under what differing conditions of noise or crosstalk? How good is any particular channel? Does the type of conversation matter to the assessments? There are endless such questions.

Again, the rapid arrival of computers has called attention of many technologists to the possibility of automatic recognition - for example, of handwriting or of speech. A great deal of this work has been naive, and there has been frequent confusion between humans and machines and betwee various meanings that can be given to the word "recognition". At present, the Courts in this country do not recognise the validity of so-called "automatic voice identification". But the threat does hang over us that so-called "voiceprints" may become accepted as are finger-prints or car number-plates today.

I could go on with endless examples of problems which require understanding of several disciplines and which indicate the need for more realistic experimentation. I cannot offer you any revolutionary new experimental method, but I will try to suggest some of the aspects of human speech and conversation that it will need to take into account.

(1) Human Communication is Always an Act of Courage.

In order to emphasise my present point of view, let me start by makin certain observations about the nature of human speech that may seem to some of you far-fetched.

Speech is far more than the mere transmission of physical sounds from mouth to ear. If anyone present tonight is so physically-minded as to doub this, let me suggest that you try a simple experiment. Next time you are i the street, preferably at night, try walking up to the first young lady you see, selected "at random" and, without any introduction or ceremony, utter any sentence used in articulation tests - for example: "the rain in Spain rarely falls in the plains" - or, even better, a so-called semantic-nonsense sentence, such as "Round cows deviate furiously" - and observe what happens.

You will probably end up in prison, where you will have time to reflect upon the fact that conversation is more than mere speech, for it involves one person in selecting another person to speak to.

This act of selection is signified by certain rituals, according to the social customs, such as raising your hat, inclining your body, gazing at a reciprocal gaze, smiling, saying "excuse me", etc., and observing the effect upon the other person. Conversation is not lightly or frivolously started. It involves an element of courage, however slight. You lay yourself open to certain risks, of snub, of ridicule, of embarrassment. The rituals that open conversation may have little to do with the purpose of the conversation, but they serve more to signify the nature of your intended purpose, something of your intended relationships.

Furthermore, it takes positive effort to <u>end</u> conversation, to decide to break off and closing rituals assist you to do so. It is not always easy to say "goodbye".

(2) <u>The Importance of Social Relationship</u>.

Whether in a laboratory, or in the street, the way in which conversation proceeds depends not just upon the partners, but also upon their <u>relationship</u> - in the context of the moment. An experimenter is faced with the problem of how to select partners, for they might happen to be friends or strangers, of the same or opposite sexes, of different ages, social classes, races, educations, statuses, experiences -------- In real life we presumably select a partner to speak to whilst rapidly categorising him in these various ways - though few of us could say just how we do it. In the street, as in the laboratory, the relationship between two people, will depend upon clothes, spectacles, sex, age, cleanliness, length of hair, ------- many visible signs and these cues are all missing during telephone conversation.

Experimental psychologists are very aware of the problems of selecting human subjects. For experiment upon conversation, the social relationship between two subjects may be controlled in certain ways.- for example (a) by the observer choosing the partners according to his knowledge of them (perhaps based upon prior personality tests) or (b) by leaving people free to choose their own partners (c) by conditioning the partners by giving them some preliminary experience such as engaging them in some preliminary "dummy experiment" (d) by contriving situations which give certain statuses to the partners; for example, by staging interviews, or simulating a Trades Union negotiation, both highly ritualised situations, establishing a social hierarchy.(1),(2) But real-life conversations involve people in many other forms of social relationship, often far less ritualised, such as argument, quarrelling, gossipping.

Of these various methods we have found that, for closest simulation of real-life conversation, method (b), (using free choice of partner) and (c) (using preliminary "dummy experiment") are particularly satisfactory. The expression "dummy experiment" here means some form of contrived game which requires the partners to speak to one another, and which they believe to be a serious experiment, whereas the results are ignored by the observer; such activity can serve to motivate the partners to subsequent conversation which becomes the real material of interest. The type of game might be varied so as to affect their attitudes to one another, to make them adopt ce statuses, or to induce certain drives, with corresponding differences in type of conversation.(3) (16;

(3) We Speak with our Whole Bodies

By far the greatest understanding of human speech has come from analys of the speech sounds alone,especially since the introduction of the laborator tape recorder and of the speech spectrograph. These instruments, together

with X-ray and other techniques, have also enabled the articulatory behaviour of speakers to be related in detail to the acoustic properties of speech. We know, in great detail, very much about physical phonetics, about movements of lips, and tongue, vibrations of larynx, nasalisation, closures and all the other articulatory motions. It is also true to say that we know a very great deal about how we hear, or mis-hear, spoken words and sentences heard under different conditions of noise and distortion. The psycho-acoustics of speech is a subject which has been elaborately developed and is well documented.

But, by comparison, we know far less about the process of conversation, which is not only our commonest experience, (with others or with oneself) but is the pre-requisite for self-awareness of man as Man. When we speak, we not only utter sounds, we gesticulate too, we adopt certain body postures, inclinations of the head, smiles, frowns, glares, winks, blushes, shrugs, and a host of other signs, some of which may be inate and some highly dependent upon the culture. We move our whole bodies, leaning forward in eagerness to catch attention, gesturing and grimacing to persuade, stamping and thumping to emphasise. We speak not just with our mouths, but with our whole bodies. Charles Darwin made an early study of facial expression, published in 1872, using the relatively new technique of photography.

When conversing with another person we approach only to a certain distance and stand in certain ways which depend upon our social relationships - our sexes, statuses, degrees of personal acquaintance and others. If there be chairs about we may sit, but we shall place our chairs at a certain distance apart, and inclined at certain angles to one another.(5),(6). The actual configurations adopted depend somewhat upon culture too - the Arabs, for example, seem to sit or stand closer together for conversation, than Europeans do. It is part of the essential ritual.

In laboratory experiments the choice is open - do you constrain the partners to sit in allocated places and positions, by fixing seats there, or microphones, or T.V. cameras, or do you leave them free to set their own seats, or to stand if they wish? The decision may have considerable effect upon the types of conversation or other verbal exchanges that ensue. Not only levels of voice will rise, with increasing spacing, but the extent to which the partners glance at each other will too, glances which part-determine the starting and stopping of utterances[7] [8](See Sec. (4)).

Unless they are deaf, few people are conscious of the extent to which they lip-read, or of their increased reliance upon lip-reading when conversing in noisy situations - e.g. in an Underground train. I believe it is true to say that we know comparatively little about the distribution of lip-reading ability among the deaf and partially deaf, partly owing to difficulties of definition. But work carried out in our own laboratories by Stapley[1] revealed that people may lip-read far better than they realise, when forced to by raising the environmental noise. Stapley used television tape recorders for his work. Lip-reading can become both more conscious and more effective if the vocabulary is small[9] (e.g. confined to spoken numbers).

The complex movements of the tongue, larynx, epiglottis and other articulators can be regarded as gestures, gestures which are normally invisible, but continued by the gestures of the whole face, limbs and body. Speech, in this view, is gestures made audible. Deaf people are then handicapped by their inability to observe the internal gestures of the tongue, larynx, epiglottis etc. Various aids have been proposed for coding the sounds of speech-waves into visible signals which the deaf person might be trained to read (e.g. "visible spectra"), but in our own

laboratories we are taking a different approach and aiming to assist the deaf person to feel certain internal gestures of the vocal organs.(10) The work is at an early stage yet, and we are starting by inducing feel of the speaker's larynx movements.

Although I believe it is true to say that we speak with our whole bodies, it is certainly the case that a listener can misinterpret the facial expression of a speaker. Indeed a speaker can aim to mislead his partner, with various degrees of skill. In particular he may wish to disguise his true feelings or mood. This faculty can be exploited deliberately, as when adopting a dead-pan expression whilst uttering some outrageous comment. We call this sort of thing acting and I want to refer later to deliberate uses of acting for laboratory experiments.(See Sec. (6)).

(4) The Importance of Gaze.

Of all movements of the body, ~~one most important class~~ is that of eye-movements - the gaze. During conversation, the partners glance at each other at certain intervals and sometimes their gazes meet. This intermittent activity serves the whole strategy of conversation rather than the semantic content, by assisting synchronisation of the two partners' remarks.(7),(8). In particular, eye-glances can be used:

(a) To observe a partner's reactions to a remark,

(b) To indicate that our remark is coming to a close (to give "permission" for a reply),

(c) To observe whether the partner is attending or not,

(d) To prevent the partner from interrupting, and perhaps for other strategic controls.

Jean-Paul Sartre has made some searching observations concerning the importance of gaze.(11). Like conversation itself, another person's gaze is a reminder of one's own existence. Sartre observed that we cannot look at another person, in a conscious way, and be looked at the same time.

For, as a looker, we feel as a person, scrutinising the other, whilst we become objects when we are looked at by the other. He was concerned with the question of embarrassment. That is to say, we are conscious scrutineers at one moment, persons, but become objects of scrutiny the next.

The extent to which one partner is able to gaze at the other, in laboratory experiments, can be controlled and it is possible that this may be overlooked when designing seating arrangements, etc. Control can be exerted by fixing seats side-by-side, or by using half-silvered mirrors (letting only one partner see the other) or by uses of two-way T.V. between the partners or other ways. But it must be reckoned with.

We are most accurately aware of the direction of a partner's gaze when it meets our own, eye-to-eye, and experiment has shown that the distribution of correct judgements about this central direction has a standard deviation of only 2 or 3 degrees.(1) (7)

When we are watching somebody speaking, whether as a partner in conversat or as a supposedly dispassionate observer, we are concerned with far more than a mere physical phenomenon - we are observing something of our own natures.

(5) Speech is Always a Social Activity

This last point is an important one and a full appreciation of what it means is essential to the success or the value of any experiments upon human speech and conversation. The observer and the observed are of like natures, for both are human. Any experimental psychologist or medical man knows this full well, of course, but many studies of speech and human communication are today carried out by people raised within the disciplines of physics and mathematics.

Though both human, the observer and the observed are supposed to be playing different roles. In much psychological experiment it is not difficult to ensure this role distinction, but when the experiments are directed towards understanding of human speech, conversation or other communicative activity,we are touching upon the very evidence of that humanity. Good method is essential to the avoidance of certain traps.

It was the great sociologist Emile Durkheim who first argued quite clearly that an individual and his society are two sides of the same coin, that one creates the other.(12) When a baby is born it is part of its own mother and it is by virtue of its mother's teaching, through language, and other sign activities of her culture, that it is created as a separate, self-conscious, thinking creature, aware of itself as a separate person. The process is continued through childhood, by linguistic encounters with family, other children and all its unfolding social relations.

Even when alone, a person talks to himself, or signifies in other ways. Speech and thinking are always social activities, and the hypothetical individual who had lived from birth in utter isolation could have no knowledge at all, not even knowledge that he existed.

In spite of this, much fruitful experimentation has been made using single persons as subjects, carrying out such highly artificial tasks as reading lists of random words, passages from literature, random sentences, and many others. Such work has told us a very great deal about the physical and the physiological qualities of human speech, but it cannot truly study language or communication, for both are social activities.

Linguistic studies, in the field, are made either by the linguist endeavouring to sign to the native, or to speak, to inter-act behaviourally with him, and to interpret, or else by listening to natives conversing,

observing behaviour and interpreting. He must observe what the natives do, in addition to speaking and signing, or he can infer no meanings. That is to say, the linguist cannot be a detached observer, in the tradition of physics, but must be a part of his own "experiments". Every newspaper reporter is in a similar position; he cannot report what was actually said, in its context, but must interpret and re-express.

Subsequently, the linguist makes analysis, and observes the regularities of the language, the syntax, which are statistical properties of a whole population and do not tell us much about the speech of any one person engaged in casual conversation. In fact, anyone who has paid close attention to what people actually say when chatting, say at bus stops, - the half-formed phrases, casual interjections, 'ums and 'ers, must sometimes feel surprised that language works at all. On an average, people may conform to the regularities of syntax, the regularities which reflect the regularities of the world around them, or what we call their knowledge of the world.

The fact that the world appears to show regularities and consistences is an epistemological fact. We can make no other assumption, for a world without regularities would simply be unknowable; We would have no language to know it with. That is, knowledge is socially derived, through the medium of human language.

Man is separated from the animals by a gulf, by virtue of language, giving him his immense power to form concepts. As many have emphasised, including Erwin Schrödinger,[13] we come to know the world, to notice some things but to be unconscious of others, each in our own peculiar ways, not so much through our observation, as by being taught to see them whilst we learn our

language. Whenever you perceive-something, you must perceive it as something; you can name it, or signify it in some manner. Speaking, knowing and thinking are all social activities and these facts make experimentation upon human communication specially difficult.

The relationship between the observer and the human subjects he is observing can be controlled in certain ways. Perhaps the subjects may not know they are being observed (e.g. by the "dummy experiment" technique I mentioned or by using subjects whose language, or social class, differs from the observer's, or by deceiving them into belief that the observer is interested in, say, some task they are asked to perform and not in their speech. Situations can also be contrived by suitable ritual - e.g. simulated interviews, negotiations, debates, competitive games, mock trials, etc. But there are whole classes of speech usage which cannot readily be contrived and I should like to end by considering these.

(6) The Great Varieties of Human Conversation and Mood.

It is only too easy to refer to human speech as though it was one thing; similarly, to language. But they are many things, of immense diversity. With language, not only can we discuss with other people amicably but we can quarrel, too. Language not only brings people together, but it can drive them apart. A child must argue with its parents, if ever it is to become a separate responsible being. Dispute is essential to any social change, for a world in which people agreed with one another all the time would be a tyranny, ruled by clichés and platitudes.

With language we can quarrel, argue, plead, negotiate, reason, dispute, enquire, express, and many other things. And each represents people in different moods. They are all very common natural phenomena yet it does not seem easy to do experiments, not only through lack of method but for ethical reasons too.

There are also problems of definition. Dictionaries gives us the common usages of words, and the O.E.D. draws clear distinctions between all these classes of human exchange and many others.(14) For example, to take 3 closely related ones - argue, dispute, quarrel.

Argue ——— maintain by reason; treat (matter) by reasoning;———

Dispute ——— argue (with, against, person); call in question, contend for, strive to win ———

Quarrel ——— occasion of complaint against person or his actions, —— invent or eagerly avail oneself of such occasion to commence hostilities, ——— violent contention or altercation between persons———

Thus an argument is reasonable and concerns a subject matter; dispute is similar but requires a second person, in competition; quarrel is against each other and requires a mutual complaint. These are clear distinctions and give some assistance towards "pragmatic definitions", by which I mean adequate statements of environmental conditions which will produce such clear distinctions of behaviour. Thus to produce an argument requires the two persons to be provided with a subject to argue about, which does not involve themselves expressly, and their statuses to be not very different. The motive could be a reward to be shared equally between them. To make this become a dispute would need the reward to be competed for; one wins, the other loses. For the behaviour to develop into a quarrel, each must be induced to have a complaint against the other. Presumably this means that each must be conditioned into believing that the other forms a threat, or a partial threat, to himself by virtue of some belief that the other holds. This would require some invention or lie on the part of the experimenter.

I am not here suggesting seriously that we try to create quarrels in the laboratory, but have merely used this extreme example to emphasise the great difficulties of experimental study of various classes of human communication.

It might well be asked: why should we need to do it all? And the only answer I shall give here is that we do need method for producing a far wider range of human speech than results from most conventional methods. It would help us to understand speech better if we could observe angry speech in action, or persuasive, or frightened, or guilty, or despairing speech ------ in all types of mood, with the emotional qualities they entail.

There is no such thing as"neutral speech", for we must always be in some mood or other. And you cannot order or instruct a person to "be in a mood". You may tell him to be angry, but he won't become so.

The voice gives extremely fine and subtle indications of mood - slight tremors, hesitances, changes of pitch and stress, syllabic speeds, stammers. The face gives visible signs too, though we can never be certain whether the speaker is acting, putting on a face. The voice is very difficult to control.

With regard to facial indications of mood, Garneau, in our laboratories, has adopted an interesting technique.(15) He uses a professional actor, who is asked to express a number of named emotions, as though to a theatre audience - and photographic slides are made - sadness, determination, fear, puzzlement, worry, disgust, etc. A group of observers then look at the slides and write down what emotional states they feel are being expressed. The results are therefore highly subjective; how are we to score, for example, if the actor thinks he is expressing fear whilst the members of the group give different interpretations, like terror, horror, fright, cowardice, etc? However, he finds that in many cases a majority of the members give the same word (not necessarily what the actor imagines he is expressing) and that word is then taken as the name of the emotion being expressed. When there is no general agreement he makes no further use of the slide.

Garneau's interest lies partly in finding out the relative importance of different parts of the facial image to different types of recognition. In the present instance, to the recognition of mood. Such recognition experiments, which require the obliteration or obscuring of different facial cues, then form a second series of experiments, using the same group of observers. Although Garneau is at present working with static slides, a similar experimental technique might be used with film or T.V. or even real-life performances. But the work has shown, to some extent, that facial expressions must be made stereotyped, when they are made out of all real situations and environment and for no natural reasons.

I hope that this lecture has not seemed to you to be too far removed from acoustics and physical experiments upon sound. But speech is a very special class of sound, of the greatest human significance. Language and conversation have aroused enormous philosophies yet, when we come to experimentation we are mostly reduced to gross simplification and to such highly artificial material as phonetically-balanced word lists, nonsense sentences, etc. Even such practical work as articulation testing, telephone testing, deaf-aid assessments, public address assessment, raise many problems of difficulty.

After all, we all know the embarrassment of being asked to "speak a few words into that microphone so that I can set the volume level". What do you talk about? It's as bad as being called upon without warning to say a few words, after dinner.

BIBLIOGRAPHY

) Stapley "Visual Enhancement of Telephone Conversation", Ph.D. Thesis, Imperial College, London, 1972.

) Morley, I.E. and Stephenson, G.M. "Interpersonal and Interparty Exchange; a Laboratory Simulation of an Industrial Negotiation of the Plant Level", Brit. Jour. Psych. 60 pp 543-544, 1969.

) Dicks, D.J. Unpublished work, Imperial College, London.

) Darwin, Charles, "Expression of the Emotions in Man and Animals", Murray Ltd, London, 1872.

) Sommer, S. "Leadership and Group Geography", Sociometry, 24 pp 99-109, 1961

) Argyle, M. and Dean J., "Eye Contact, Distance and Affiliation", Sociometry, 28 pp 289 - 304, 1965.

) Argyle, M. "Social Interaction", Methuen, 1969.

) Champness, B.G., "Mutual Glance and the Significance of the Look", Adv. of Science, pp 309-312, Mar. 1970.

) Sumbry, W.H. and Pollack, I, "Visual Contribution to Speech Intelligibility in Noise", J. Acoust. Soc. Am. 26 pp 212-215, 1954.

) Edmundson, W. and Cherry, E.C. Unpublished work, Imperial College, London.

) Sartre, Jean-Paul, "Being and Nothingness", Methuen & Co. Ltd., London, 1943.

) Durkheim, E.

) Schrodinger, E. "My View of the World", Camb. Univ. Press, 1964.

) The Concise Oxford Dictionary.

) Garneau, M. Unpublished work, Imperial College, London.

) Rosenthal, R. "Experimenter Effects in Behavioural Research", Appleton-Century Croft, N.Y., 1966.