# Temporal aspects in conversation[1]

Andersson, Lennart
Department of Linguistics, Gothenburg University
412 98 Gothenburg

## ABSTRACT

*It is important that relatively complex conversational situations are exposed to deeper phonetic investigation. This could then reveal interesting general issues in speech and speech perception. The factors requiring study are largely those which have always been studied within phonetics, e.g. duration, but these known factors can be realised in a number of new and complex ways.*

## INTRODUCTION

Time is one of the dimensions in which speech can be analysed. Within phonetics, segment and utterance length have for instance been studied in detail.

Within the two-speech-project, Andersson (1993a), we try to, amongst other things, investigate how the time dimension kan be organised in conversation between two people. The fact that one is dealing with two people's interactive speech means that it is sometimes other temporal behaviour than that which is usually studied in phonetics which is in focus. Here I intend to go in to two complicating factors. Firstly, there is the question of the temporal relationships which exist between the information communicated by separate senses, simplified below to just temporality from a bimodal perspective. Secondly, I intend dealing with the temporal relationships which prevail when two speakers talk together and interactively.

A central problem is, however, what time precisely is. Phonetic research has shown that an 'objective', physically measurable time, duration, isn't necessarily the same as 'subjective', perceived time, sometimes called quantity.

Without going in to the time dimension too deeply, I would like to point out one interesting aspect. There are certain physical and physiological processes which can recur regularly and thus assist the perception of rhythmicality in speech. This rhymthmicality might be one of the central factors in the production and perception of interactive speech.

## TEMPORALITY FROM A BIMODAL PERSPECTIVE

When studying conversation between a number of people a bimodal perspective becomes particularly interesting, ie. when the person producing and the person perceiving are two different people.

Those speaking in a conversation utilise different modalities, both in production and perception of speech. I shall only discuss the interaction between two sense-modalities here, sight and hearing. These two senses have different physical and physiological requirements and these requirements are of consequence for the organisation of conversation.

From a physiological perspective it is the case that each and every one of the senses is affected by its own important, life-supporting process. The speech sounds which the ear perceives are produced by speech organs which participate, simultaneously with production, in breathing. Breathing is a necessary activity which must be carried out. When one speaks, breathing and speech must be coordinated. This has, of course, a effect on the acoustic modality. In the same way, the eye must blink so that it doesn't dry out. Blinking involves a short interruption of the visual perception and in this way can affect the visual modality.

From a physical perspective it is the case that transmission through air happens much faster for lightwaves (approx. 300,000 km/sec) than for soundwaves (approx. 340 m/sec). In Table 1 below can be seen the time it takes for light- and soundwaves to be transmitted over a distance approximately appropriate for conversation:

[1]Translated from Swedish.

**Table 1:** The table shows the time taken for light and for sound to propogate itself over various distances.

| Distance | Light | Sound |
| --- | --- | --- |
| 1 m | 0.000003 ms | 2.94 ms |
| 2 m | 0.000006 ms | 5.88 ms |
| 3 m | 0.000010 ms | 8.82 ms |
| 4 m | 0.000013 ms | 11.76 ms |

There are phenomena in speech, such as rounding, which are perceptible both visually and auditorily, see Andersson (1993a). The time differences in the table above show that such phenomena can be temporally simultaneous in production, while they are temporally separated in regard to perception. This time difference does not appear to create any especially great problems for perception, in so far as it is a question of 'expected values'. If one, on the other hand, manipulates the values, for instance by putting the sound ahead of the picture in a television interview, then we react immediately and are disturbed by the fact that picture and sound are not synchronised.

However, one can't discuss temporal simultaneity or separation alone. It may also be the case that one sense takes a general and holistic aspect into account to a greater extent, whilst the other takes specificity and particularity into account. One modality could then have a larger time-window than the other from a perceptual time perspective, ie. phenomena in one modality would have greater extent in time than phenomena in the other.

Yet another phenomenon which distinguishes between modalities is that the auditory modality can be silent, ie. a sort of 'zero-line' or point of departure where nothing happens, which can be compared with the remaining situations where something happens, ie. where there is no silence. One can thus organise the auditory modality as a temporal 'string' of utterances, where 'non-silence' and silence alternate. It is difficult to find something corresponding to silence in the visual modality, perhaps blinking, or closing your eyes. Of course, darkness means that the visual modality cannot be exploited, but changes from dark to light are hardly likely to occur in conversation. The visual modality is better described as continuous and simultaneously ongoing utterances, one utterance for each interlocutor, Andersson (1993b).

## TEMPORALITY IN TWO INTERLOCUTORS

When several people speak with each other they do not tend to talk over the top of each other, but rather order the turns in some way. The various interlocutors' utterances can thus be described on the basis of the others' utterances. Each participant is both listener and speaker, sometimes simultaneously, sometimes in alternation. Both the speakers' and the listeners' demands and requirements affect the form of speech, cf. Lindblom (1983).

It is not the case that one person has complete and simultaneous control over all the processes needed for the functioning of a conversation, in any given conversation. This applies to, among other things, breathing, Andersson (1992b). Consider a hypothetical speech situation:

T1: "å då kom ja in för å se om ja kunde göra nåt åt det då va"
T2: "m"
T1: "men inte fan de dom hade ju kört hela fanskapet i botten"

(S1: "and then I went in to have a look if I could do something about it right"
S2: "m"
S1: "but it didn't fucking work ... they'd stuffed up the whole fucking thing")

In this example, Speaker 1 (S1) requires feedback and thus demands that Speaker 2 (S2) react. S1 determines when it is time for S2 to indicate that he follows. But S1 can

hardly have control over S2's breathing process. One could imagine that S2 might have just exhaled and is without air resources to give vocal feedback when he is required to provide a comment. With this example I have attempted to show that a relatively complex temporal interaction prevails in conversation.

This example is about feedback processes in conversation. The dimension of meaning in feedback has been described by Allwood, Nivre, Ahlsén (1990) and the same authors (1992). I now intend to discuss some of the complications in the dimension of expression in feedback, how feedback is manifested in a physical/physiological reality. If one wants to speculate about these problems, one can imagine three different possibilities:

**1 Interlocutors breathe in time**
This means that speakers in a conversation coordinate their breathing phases with each other and breath in and out at the same time. S1 would then, in the example given above, control S2's breathing as well, and demand feedback only in those phases of the conversation where S2 is capable of giving feedback.

**2 Interlocutors do not breathe in time, and the exact timing of the feedback-givers isn't as important.**
This would mean that S1 above would demand feedback without taking the breathing phase of S2 into consideration. S1 would then wait for a verbal feedback-giver from S2 and then continue. The time relationship between feedback-demander and feedbacker would then be able to vary to some extent.

**3 Interlocutors do not breathe in time, but the exact timing of the feedback-giver is important.**
S1 can demand feedback without consideration for S2's breathing phase, but must at the same time accept that S2 can either answer with vocal feedback or, for instance, with visual feedback which is independent of breathing, such as a nod of the head (of course, speaker 2 can do both at the same time as well).

## CONCLUSION
When studying speech, it can also be important to study "conversational speech", dialogue speech. Even if it is the case that informants can contrast a certain specific opposition in a phonetic test, it isn't necessarily a foregone conclusion that the same informant makes use of the same contrast in a normal conversation. It is thus also important that relatively complex conversational situations are exposed to deeper phonetic investigation. This could then describe interesting general issues in speech and speech perception.

The factors requiring study are largely those which have always been studied within phonetics, eg. duration and intensity, but these known factors can be realised in a number of new and complex ways. Variations in different temporal relationships are used in a number of different ways in conversation. Part of this variation is based on various physical and physiological functions' nature. An interesting question is how this variation relates to the variation of temporal relationships which are exploited in speech.

## REFERENCES
Allwood, J., Nivre, J. & Ahlsén, E. (1990) 'Speech management - On the Non-Written Life of speech', in *Nordic Journal of Linguistics*, 13, 2-48.
Allwood, J., Nivre, J. & Ahlsén, E. (1992) 'On the Semantics and Pragmatics of Speech', in *Journal of Semantics*, 9, 1-26.
Andersson, L (1992a) 'Visuella korrelat till (svenska vokalers) rundning', in *Fonetik 92 - Papers from the sixth Swedish Phonetics Conference*, 103-106.
Andersson, L (1992b) 'Bimodalitet och tald interaktion', presented at OFTI 1992.
Andersson, L (1993a) 'Tvåtalsprojektet - en presentation', in *Fonetik 93 - Papers from the seventh Swedish Phonetics conference*, 33-36.
Andersson, L (1993b) 'Visualitet och talspråk', in *Proceedings from the XIVth Scandinavian Conference of Linguistics*.
Lindblom, B. (1983) 'Förstå och underförstå - Något om de processer som formar talrörelserna', in Teleman (ed) *Språk and tanke*, 147-178.