

THE THEORY AND METHODOLOGY OF SPEECH SCIENCE AND CONTRASTIVE ANALYSIS

JAAKKO LEHTONEN

University of Jyväskylä

1. Contrastive linguistics is a discipline which can be considered either to belong to pure, autonomous linguistics or to be a representative of applied type of research. All depends on the definition of linguistics, and applied linguistics. In this paper, contrastive linguistics is taken to mean the scientific study of how people use language to communicate in two or more languages and what are the consequences when the two communicative systems clash in the foreign language learning situation. This kind of contrastive analysis of human speech communication is necessarily a branch of applied linguistics.

This does not mean that CA is an activity whose *only* aim is to apply linguistics for some practical purposes outside the scope of pure linguistics. It is a discipline which cannot rely on linguistics alone; it works with similarities and differences in various human verbal, and even non-verbal, codes. In view of what is expected of CA today, it will have to be able to absorb both theoretical perspectives and methods for practical analysis from all branches of the disciplines which deal with language and speech or human behaviour in general. One of the more important neighbouring sciences is the science of speech, traditionally called phonetics. In this paper, an attempt is made to discuss the question of how the science of speech can contribute to CA.

The first chapter deals with some theoretical aspects of the interrelationship between grammar and speech and the nature of the speech chain. In the concluding chapter a brief summary is given of some methods of the instrumental analysis of speech which may have some relevance for CA.

The term *speech science* (or *the science of speech*) is here preferred to 'phonetics' for several reasons. In structural linguistics the term phonetics refers to the output level of phonology (cf. figure 1), which often carries along several

linguistic implications. Some of these will be discussed in chapter 3 below. Phonetics is often defined as a scientific activity whose aim is to reveal and describe how phonological information is conveyed through the sound waves of speech in each language. The term speech science, on the other hand, being free from similar biases and connotations, will be used in this paper to refer to the study of the entire chain of human speech. It includes aspects of communication technology, psycho- and neurolinguistics, and discourse analysis, as well as various other aspects of the study of human verbal and non-verbal behaviour in general. According to Ladefoged (1977:409), the linguist is trying to describe the patterns that occur within each language, whereas speech scientists are concerned with how people communicate: "They want to explain what people are doing when they talk and listen, and perhaps even what they are doing when they think and talk, and when they listen and understand".

The choice in terminology thus, for its part, reflects a distinction between the code-centered approach of linguistic phonetics and the communicative or psycholinguistic perspective in which attention is paid to the communicative behaviour of individual speakers instead of abstracted structures.

2. Expanding the Contrastive Analysis Framework

The limitations of traditional contrastive analysis are evident today (cf. Sajavaara and Lehtonen (1979)). For several reasons its method of describing the grammars of two languages as abstract collections of rules divorced from their users is insufficient. The fundamental role of a human language is to function as a means of communication in human interaction. Accordingly, the task of CA is not a mere parallel description of two grammatical structures but also the description of the chains of communication in the two languages. This means that the analyst must, in addition to the description of the grammatical structures, make an attempt to map the differences and similarities in the processes which take place in the speaker and in the listener during speech communication. Language is used by individuals for definite purposes. Its use is linked with the intentions of the speaker in some specific time and environment. The use of language is always part of human interaction. This aspect of language, the similarities and differences in the rules of discourse in interaction, must also be included in CA.

Contrastive analysis may no longer be the best label for this kind of research. It does not mean the description of the equivalent patterns of two languages but an analysis of cross-language interaction and a search for the reasons for the difficulties and failures in the use of the foreign language by the student. In such a work, contrastive analyses of grammatical structures are still necessary but they must be supplemented by an analysis of the

psychological and sociological aspects of linguistic interaction. The former comprise the study of differences and similarities in the information-bearing features of the acoustic speech signal, differences and similarities in the cues of perception, and sources in the perceptual patterning of speech and foreign accent from the viewpoint of the foreign speaker-hearer and the native hearer, and differences and similarities in linguistic interaction, which will cover the discourse behavior of native and non-native speakers, including features such as tempo, pauses, paralinguistic and kinetic elements, and various pragmatic, structural and semantic paradigms in their communicative function within discourse. The sociological aspects involve the social context, the environmental relations or proxemics, and other factors which affect the choice of the discourse register and modify the attitudes, communicative intentions, and responses of the participant in the discourse.

This description of the objectives of contrastive analysis leads to the fact, mentioned above, that CA is not just a special branch of pure linguistics. In addition to the theory and methods of linguistics, it must inevitably be supplemented by the methods of disciplines such as sociology, psychology and neurology, and applied mathematics, as concerns the analysis and description of pragmatic patterning, cognitive mechanisms, perception, and the information-processing systems in man. The expanded objectives of contrastive analysis are thus, to a large extent, parallel to the goals of the modern science of speech, which has the analysis of various aspects of speech communication as its objective.

3. The Static and Dynamic View of Language

Throughout the first decades of the present century there was an open conflict in the phonetic sciences between instrumentalists, who referred to the physical manifestations of speech, on the one hand, and advocates of 'auditory' phonetics, who based their claims on sophisticated auditory observations, on the other. The observations of these two factions were never identical; however, neither of them was wrong: one described the properties of the physical stimulus, the other observed the result of auditory perception. Early phoneticians never saw the core of the problem: they did not ask what the processes are which lead to the auditory perception of linguistic structures, or how the linguistic information is conveyed by the sound-waves of speech, and thus the conflict remained unresolved.

In a way, the same dichotomy is repeated today in the relationship between phonologists, on the one hand, and representatives of speech science, on the other. Phonologists engage in a dispute among themselves about the reference level of classificatory features: one of the schools claims that the features are principally acoustic, another school is of the opinion that the

reference level should be that of the articulatory settings. The third choice is the view brought forward by the psycholinguist or the speech scientist, that the features are principally neither acoustic nor articulatory but phonological and linguistic; the phonological rules which operate with the feature matrices have as little to do with actual neurological and physiological processes in speech as the rules of generative syntax have with actual processes in the human brain in the production and perception of linguistic messages (cf. Buckingham and Hollien 1978, esp. p. 294). Grammars, including phonology, are always descriptions of structures, not of processes. A grammatical rule, whether it is traditional or 'generative', is only a description of a given regularity in the structure of the language and not a model for the actual processes that are found in the brain of the speaker and the listener (cf. Clark and Clark (1977:190 ff.); For modelling of transformational grammar, see Bresnan (1978:4-5). Therefore, they can never predict all of the interference phenomena that result from the clash of two different information processing systems — for a clash is evidently what takes place in foreign language speech communication.

One of the crucial questions in the discussion concerning contrastive analysis has been the choice of the reference model: should CA be based on traditional structural or generative grammar, and if so, which variety? In applied contrastive studies the choice is often eclectic: those grammatical theories are exploited which seem to give the best explanation in each problem. The alternative is to describe the structures of the two languages consistently within the framework of a given grammatical theory irrespectively of how efficiently the theory can explain the problems of the cross-language analysis. However, the choice of the reference model does not concern only the model to be used in the description of grammar. It should also comprise the choice between a static and a dynamic view of language or, in other words, the choice between a linguistic and a psycholinguistic description, or between a structural and operational modelling of language and communication. The former (the structural model) aims at describing the abstracted and idealized structure of language, which, in the case of CA, means similarities and differences in the structures, or grammars, of two or more languages. Independently of the choice in the linguistic model, taxonomic or generative, the objective of this kind of description is always the structure of the language instead of the actual processes.

The target of an operational model of language is not an abstracted static structure but language in action. This does not mean pragmatics or description of linguistic interaction only, but also observation of the entire set of communicative means, both verbal and non-verbal, which are applied in human interaction.

The final goal of the dynamic model is not a comprehensive description of

language as a paradigm of rules for structures, where actual messages have only the value of raw-material. The objective of the description and analysis is the vehicle of human communication, both from the perspective of the speaker and the listener, and that of the society. The fact that the description is not focussed on language but also on messages also reflects the content of several basic concepts: a linguist, when speaking of a proposition, for instance, has the meaning of a sentence in mind. In real communication, however, the proper meaning, or the intended logical basic content of the message, may be conveyed through another channel, not language, eg. through gestures, body movements, or paralanguage. There may or may not be a concomitant linguistic utterance, either in accordance with, or contradiction to, the proper meaning. As is well known, the listener normally resorts, in the latter case, to the non-verbal information and thus receives the intended, or propositional, meaning of the message (cf. Knapp 1978:20-26).

4. Modelling of Speech and Language: a New Approach

There is certainly no exaggeration in the claim that, during the last ten years, speech science together with some related branches of science such as experimental neurological and psychological studies of language and speech (ie. neuro- and psycholinguistics) has radically changed the traditional view held by linguists of the structure of actual speech and of the transmission of linguistic information through the speech chain. Figure 1 illustrates the traditional view of the hierarchy in the perception of speech: the speech

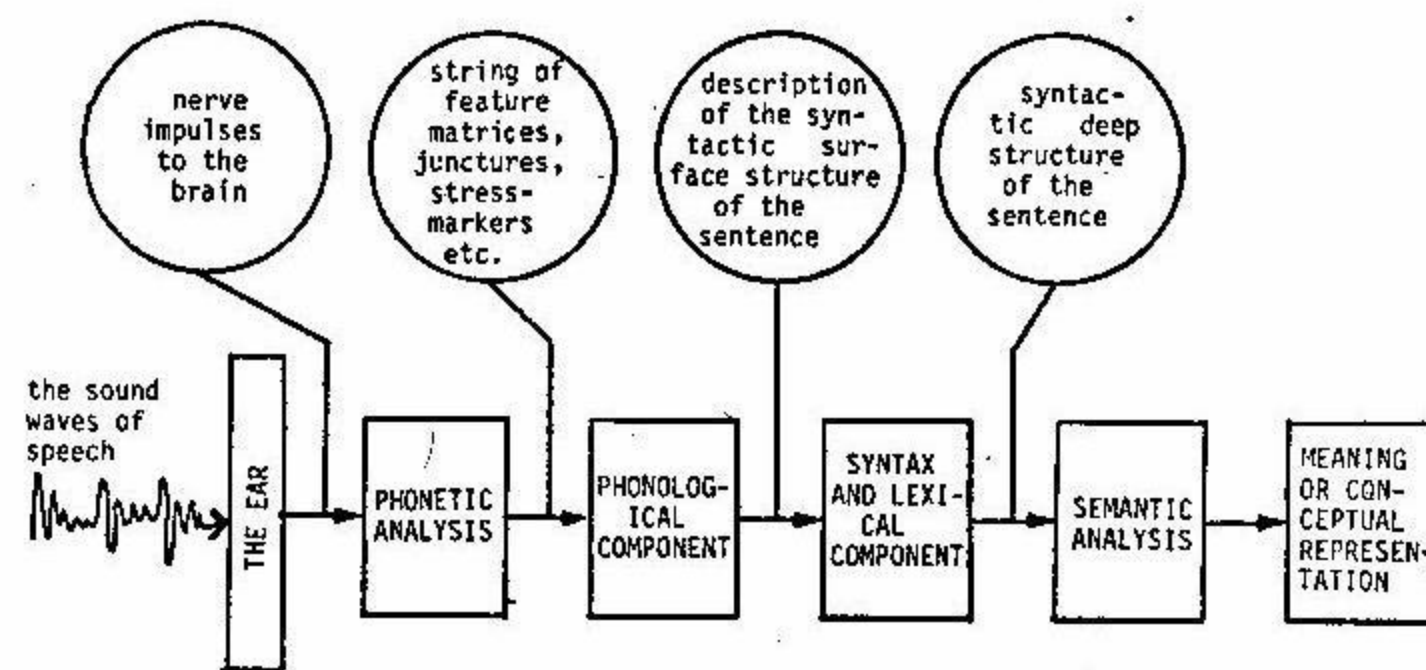


Figure 1. An imaginary 'hierarchical' model of speech perception or a 'model of successive transformations' (this model is not supposed to represent the actual statements of any 'school').

signal undergoes a series of successive transformations whereby the information is recoded into more abstract forms of representation. There is no reason, however, to assume that perception of speech is based on the phonological information conveyed through the acoustic speech signal only or that the recognition process should proceed in a one-way manner from left to right or from sound-wave to meaning. As a matter of fact, there is plenty of experimental evidence that language perception does not work in the way described in Figure 1.

Both the recent data from psycholinguistics and the progress in developing automatic speech understanding systems allow for the rejection of a strict serial organization, and support, instead, the view that the listener processes the message simultaneously, or in some overall way, on several 'levels' of perception.

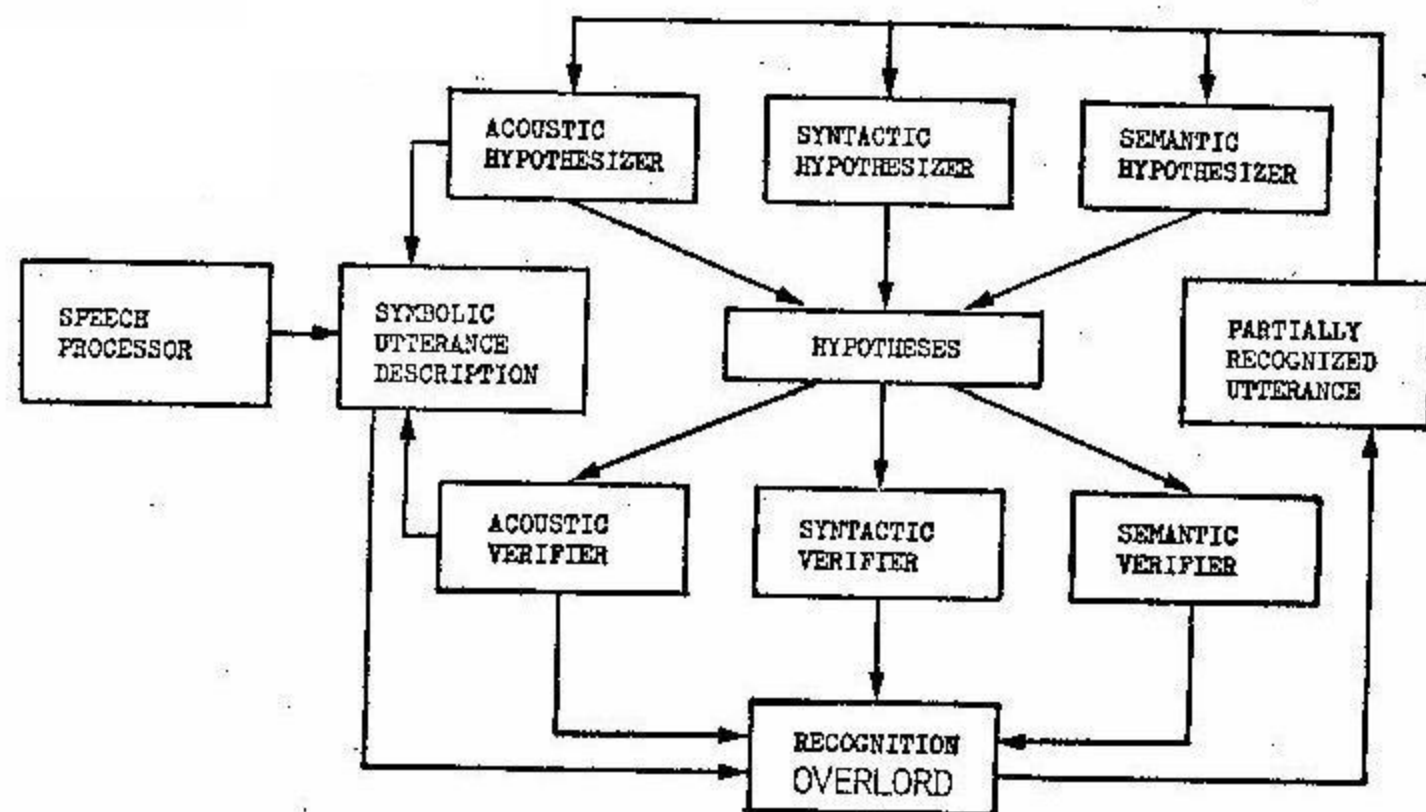


Figure 2. A "heterarchical" model of speech recognition (Ainsworth 1976: 118).

Figure 2 represents an alternative model of the speech recognition system. The figure does not represent a real human information-processing system; it is derived from a work dealing with automatic speech understanding systems (Ainsworth 1976:118). However, it comprises several features typical of recent psycholinguistic models of human speech perception. As compared with the hierarchical or input-driven model given in Figure 1, it is characteristic of the present model that the recognition of the message is not based on the acoustic signal only. Instead, recognition is a result of active guess-work

in which the input signal, i.e. the message received, is compared with the 'synthesized' alternatives. According to this model, understanding is not passive decoding but active processing, where the listener's knowledge of the phonetic, syntactic, semantic, and social constraints and probabilities play an essential role (cf. Marslen-Wilson and Welsh 1978).

The modern view of speech implies that there are in the flow of the sound waves of speech certain cues for various levels of 'grammar' and that the grammatical and semantic information is not present in the speech waves coded through segmental phonology only, but also in the features which directly refer to 'higher-level' structures. Thus, the flow of speech is not just a manifestation of strings of phonemes but also of words, word groups, and sentences as well. But it should not be forgotten that in a normal face-to-face situation language and messages coded linguistically are only one of the channels available for the speaker to transfer information: To a certain extent he can choose between a non-verbal and a verbal channel and between a variety of paralinguistic modifications to the phoneme string of the linguistic signal.

5. Speech and the Observer

Conscious Perception of the Physical Parameters of Speech

Many of our impressions of speech events are — if not theoretically acquired as in the case of linguists and phoneticians — simplified generalizations derived from the categories of the internalized patterns of rules and structures which we call phonology. What is claimed here is that man does not normally hear, or experience, the actual physical utterance but a chain of sound segments which correspond to the expectations concerning the utterance. This is by no means exceptional from the point of view of human perception mechanisms in general: our perceptions are structured by our cultural conditioning, education, and personal experiences, which result in fallacies concerning our observations. Irrespective of what really takes place our observations are affected by preconceived notions about what we will see or hear. Perception is constrained by expectations and stereotypes: we observe structures which we have learned to exist and which we choose to observe. It is also important to remember that there is much inconsistency in the expression of one and the same mental experience: even if two people experience a stimulus (eg. an intonation pattern of a sentence) in a same way, they may express their observations differently (cf. Knapp (1978:381 ff.), see also Öhman 1975).

For the speaker-hearer, language is a psychologically real, abstracted and idealized structure and, due to its categorizing nature, this structure acts in perception like an automatically tuned detector which filters from the

message all information that is not linguistic (or communicative). In other words, language forms for man an earphone through which he can hear only language, and not speech. Each acoustic speech-signal which has been identified as being speech is "heard" as a representative of a given linguistic structure and as a consequence, the listener is able to consciously perceive only a few of the properties of the signal and is deaf to many physical events in it which lie beyond the conscious auditory experience. Many of the acoustic cue-features which actually carry the information distinctive for speech perception are subconscious (notice, for instance, the formant patterns of vowels, consonantal loci and transitions, temporal patterns of the strings of speech-sounds, etc.).

It is possible for the listener to focus on various linguistic structures of the utterance, such as the meaning of the utterance, the syntactic structure, or the phonological structure, i.e. phonemes and phoneme strings. Similarly, conscious attention can be paid to certain postures and movements of the speech organs, but both the way in which they actually operate during speech and the features in the resulting acoustic speech which enable us to understand the message are beyond the scope of conscious observation. Thus, there are features in the speech chain which are outside our capacity to consciously monitor the processing of messages and the differences in the processing between two languages.

All this has certain important consequences: the contrastive analyst and the foreign language teacher must realize that the way they 'see' a certain pattern of the target language may totally differ from the way the student perceives the same thing. The problems of the foreign language student can be understood only if we know how he 'feels': what he attempts to hear, what he actually hears, and what the structures are he actually perceived, and how these differ from the target, i.e. from the way in which the native speaker's perception works in similar situations.

Phonetic Transcription and the Reliability of Auditory Observations

The evaluation of the capacity of traditional phonetics and phonology to describe the phonetic reality must also include a consideration of the phonetic symbolism which is used in phonetic transcription. In linguistic field-work, phonetic symbols are easily given the value of an icon, i.e. they are regarded as images of some physically true articulatory postures or acoustic qualities. And, accordingly, they are often used as 'phonetic evidence' for various phonological rules, for instance. However, the phonetic symbols are nothing but *symbols* of given *perceptual* phonetic categories, i.e. symbols of categories which have been created in the human mind. For the most part the 'narrow' phonetic transcription is nothing but an allophonic description of the alleged phoneme strings of the utterance.

What phonetic symbols are not, and never can be, is *models* of the actual articulatory or acoustic events in speech. Phonetic symbols are not meant to describe the physical reality of speech; they represent the most peripheral level of the perceptual categories in the phonic structure of language. Phonetic transcription could be defined as the lowest level of pure phonological (linguistic) description of language and utterances. The phonetic notions of sound segments, stress degrees, or rises and falls of intonation are, of course as true as any sensations and feelings whatsoever. But they are true in the linguist's subjective reality only. Very often the physical reality differs from his experience: the tongue is not in the position where he thinks it should be, there is no physical prominence to the syllable which he hears as being stressed, and the change in the fundamental frequency of speech may be just the opposite to what he perceives and describes as an intonation curve.

Though there exist, without dispute, some universal tendencies in the perception of the phonetic parameters of language, the reception of phonetic categories is influenced to a great extent by the listener's mother tongue and by the classificatory principle he applies. A Finn and a Pole, for instance, do not 'hear' the same English utterance in an identical way; they perceive the stimuli in terms of the perceptual categories of their own language. What is important from the point of view of contrastive analysis and language teaching is, however, the fact that neither hears the message in the same way as a native listener does.

6. The Speech Chain and Cross-Language Differences

On the level of phonetics the objective of contrastive analysis should be the establishment of the similarities and differences in the way in which ideas (or messages) are conveyed through the chains of communication in the two languages. The contrastive description of the similarities and differences between the communication chains of the two languages consists of two kinds of factors: similarities and differences in the way in which linguistic information is turned into physical speech, and those in the way in which the parameters of physical speech are processed in reception.

Before discussing the special problems related to language contact in foreign language learning which are of interest for contrastive analysis, it is necessary to sketch, in broad outline, the processing of the speech chain from the perspective of modern phonetics. The role of articulatory movements in the transmission of information is that of converting the phonological information, i.e. the string of symbols of the phonologically coded message into sound waves. From the point of view of grammar and phonology, this conversation changes the chain of distinct phonological segments into a blend of various qualities and pretended or overlapping boundaries between them. The result-

ing phonetic variation, whether it is due to an intended articulatory move or some motor constraint, serves as a piece of information for the listener. The properties of the sound waves of speech which convey the linguistic information or which, in other words, are used in perception by the listener to detect a given category are called either phonetic features or acoustic cues of identification. Each phoneme or phonological feature of a language can be described, on the physical level of speech, as a set of cues or phonetic features which bring about the perception of a given phoneme in various contexts. In the light of recent findings in speech perception research, it seems plausible that there are special feature detectors tuned to reveal, in the sound wave, the acoustic features which are used as cues for phonological categories in speech. Evidence for feature detectors has been gathered in tests with, for instance voice onset time and formant transitions. Moreover, the detectors seem to be tunable, which explains their function in languages with phonetically different sound patterns.¹

Identification cues, similar to those for phonological segments, are also found in the speech wave for higher-level information bearing structures. The listener has learned to follow some cue-features in the oscillation of voice which indicate, for instance, a lexical boundary, the phonotactic structure of the word, or some syntactic and semantic structures of a sentence such as constituent structure, topicalization, or emphasis (in form of pitch patterns, sentence rhythm, final lengthening and other phenomena in the temporal organization of the segment strings, etc.).

For speakers of one and the same language the feature detector systems is, of course, matched to the acoustic patterns which correspond to the habits of motor implementation of speech in their own language. In foreign language learning, interference is found between the cue patterns of the mother tongue and those of the target language. It is likely that the differences in the general characteristics of the pronunciation of the two languages (sometimes called the idiomacy of pronunciation or base of articulation) also comprise differences in the cueing systems and difficulties in the perception of the spoken foreign language.

Unfortunately, the capacity of the tests available at present to reveal the 'critical' points in the interlanguage speech channel is rather limited, partly because we still lack an integrated picture of the mechanisms functioning

¹ These terms are often used in modern psycholinguistics in a more restricted sense to refer to given experimentally demonstrated phenomena in the perception of some acoustic parameters in speech, such as voice onset, which have parallels both in the visual feature-detection mechanism of man and in the perceptual mechanisms of animals (cf. Massaro 1978; Cairns and Cairns 1976:143 ff.). In this paper, the concepts of 'cue' and 'cue detection' are used in a more figurative way without referring to any precisely located mechanisms in the process of perception.

in the transfer of information in a normal language communication. The only way of testing speech perception, available for CA, are identification tests based on minimal word pairs and other similar traditional methods of testing pronunciation. These methods, are however, so closely related to the taxonomic view of language that they are suitable only for the testing of certain types of phonological contrasts.

There is, however, an indirect way of approaching the problems of the foreign language speech perception: it is the analysis of the student's speech production. It can be hypothesized that deviancies from the target in the production of the features which are known to function as cues for syntactic or semantic processing of utterances reflect corresponding difficulties in perceptual processing. Similarly, if some native-language features break through the pronunciation of the foreign language, the student may be expected to seek the same cues also when he is trying to understand messages in the foreign language.

7. Instrumental Methods

The concept of the speech chain is, again, useful for the discussion of the problems and the existing instrumental methods for the analysis and physical description of the events in human speech. Analyzing the speech chain means seeking the answer to questions such as how the messages are transmitted from one person to another, or through what kind of transformations the message gets from the brain of the speaker to the brain of the hearer?

The main stages of the speech chain are: the process of speech production, the sound wave (the acoustic signal), and the process of speech perception. The methods of investigation can also be divided into three major categories according to which stage of the speech chain is the object of research. The acoustic signal is the most easily accessible stage; it can be recorded on tape and analyzed by means of several acoustic research apparatus (intensity meters, fundamental frequency indicators, duplex processors, sound spectrographs, digital spectrum analysers, etc.). But when we move from research centred on the sound wave to research dealing with speech production or perception, the task becomes the more difficult the 'higher' the phenomenon which we want to study is in the speech chain. There are certain methods for the study of peripheral phenomena (such as the movements of the organs of speech or the changes of the air pressure in the cavities of the vocal tract), but we still lack methods to study the phenomena in the central nervous system.

Figure 3 illustrates the application of the different methods of investigation to the speech chain. It is not meant to cover all methods of speech research, nor is it possible to describe here all individual methods and instruments of speech analysis. They can all be applied to various analyses with contrastive

orientation as well. Many of them, however, involve some specialization in physiology, acoustics, and the technology of speech analysis, as well as access to a well-equipped speech laboratory, which naturally reduces the contrastivist's interest in the application of such methods in his research.

In addition, there are certain methods within the reach of a contrastive linguist not specializing in phonetics which are relevant in the analysis and testing of 'higher level' linguistic problems such as sentence construction, conveyance of meaning through grammatical constructions and the speech chain, and the progress of discourse in real time. As was pointed out above, it is largely impossible for man to perceive physical speech events objectively. Therefore, one of the most essential applications of the instrumental methods for CA is simply the visualization of physical speech events. Two dimensions in particular are important for the point of view of the analysis and description of linguistic structures larger than sound segments or individual words: these are the time axis of speech and the fluctuation of the fundamental frequency.

Figures 4, 5 and 6 illustrate three fragments of discourse described in a form of an on/off signal on the real time-axis. (For a detailed description of the instrumental method, see Sajavaara and Lehtonen 1978.) The advantages of this kind of description as compared with traditional transcription are evident: In this method, we have access to the analysis of the distribution of speech performance in time, which is an integral element in the linguistic behaviour of man. This information is of special value in the analysis of discourse dynamic and in all tests in which information about reaction time, of hesitation, location of pauses, etc., is needed. The present figures illustrate tracings from a four channel equipment planned for the analysis of discourse at the Phonetics Laboratory of the University of Jyväskylä. It makes possible the recording and analysis of simultaneous speaking turns (eg., simultaneous starts, feedback moves of the listeners, or completions) as well as the chronemics of the discourse in general, which has so far been a parameter neglected in the analysis. Figures 7-9 illustrate tracings of a fundamental frequency meter as applied in the analysis of discourse intonation.

One of the methods in speech research whose possibilities are far from being exhausted is the testing of the linguistic reactions of the speakers of the two languages. The stimuli for the tests can be either natural speech, instrumentally processed natural speech, or synthetic speech generated by means of a speech synthesizer and/or a computer.

Figure 10 illustrates the results of a preliminary test of certain phonetic cue features whose stimuli were produced by means of a speech synthesizer. There are, of course, lots of problems involved in the composition, execution, and interpretation of such tests as well as in the use of synthetic stimuli, but they can yield results which are unattainable through conventional methodology.

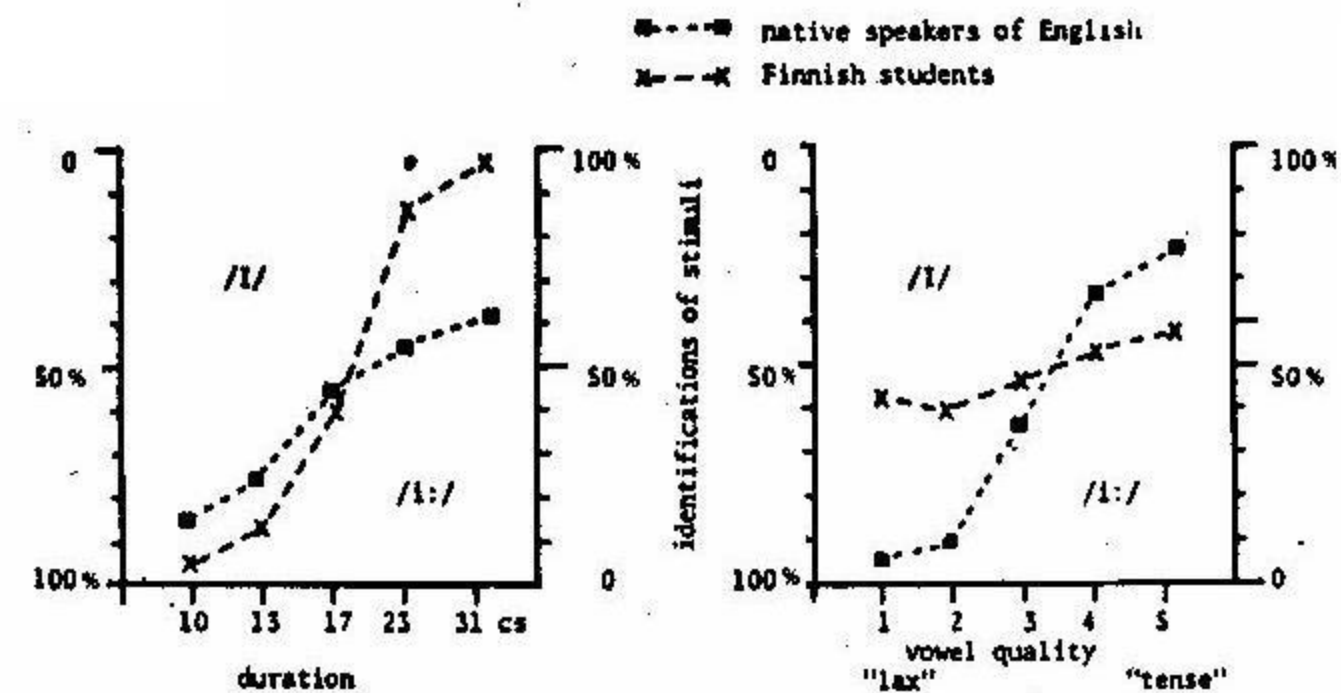


Figure 10. The responses of native speakers of English and Finnish students of English to various cues of the tense-lax distinction. The chart on the left shows the effect of the change of vowel duration on vowel identification. The chart on the right shows the respective responses to changes in the acoustic structure of the vowel. The diagrams show that a change in the quality or formant structure was for the native speakers of English the stronger of the two cues tested. However, the change from a "lax" to a more "tense" vowel quality hardly affected the identifications of the Finnish students of English. On the other hand, changes in the duration of the vowel, which affected the identifications of the native speakers only a little, have a dramatic effect on the identifications by the Finns. The test thus reveals that a Finn does not apply the correct cues of identification when discriminating between tense and lax vowels. The diagrams are based on studies with synthetic speech by Raimo and Suomi (1979). For further discussion and test results, see Lehtonen and Sajavaara, eds., (1979).

REFERENCES

- Ainsworth, W. 1976. *Mechanisms of speech recognition*. Oxford: Pergamon Press.
- Bresnan, J. 1978. "A realistic transformational grammar". In Halle, M., Bresnan, J. and Miller, G. A. (eds). 1978. 1-59.
- Buckingham, H. and Hollien, H. 1978. "A neural model for language and speech". *Journal of phonetics* 6. 283-97.
- Cairns, H. and Cairns, C. E. 1976. *Psycholinguistics, a cognitive view of language*. New York: Holt, Rinehart and Winston.
- Clark, H. H. and Clark, E. V. 1977. *Psychology and language: an introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.
- Cohen, A. and Nooteboom, S. G. (eds). 1975. *Structure and process in speech perception*. Heidelberg and N. Y.: Springer.
- Davis, S. M. 1978. "Audition and speech perception". In Schiefelbusch, R. L. (ed.). 1978. 44-66.
- Dechert, H. W. and Raupach, M. (eds). 1979. *Towards a cross-linguistic assessment of speech production*. Bern: Lang.

- Halle, M., Bresnan, J. and Miller, G. A. (eds). 1978. *Linguistic theory and psychological reality*. Cambridge, Mas.: The M. I. T. Press.
- Hurme, P. 1979. "Neurophonetics: aspects of the neural control of speech production". In Hurme, P. (ed.). 1979. 49-57.
- Hurme, P. (ed.). 1979. *Fonetiikan päivät - Jyväskylä 1978. Papers from the Eighth Meeting of Finnish Phoneticians*. Papers in Speech Research 1. Publications from the Institute of Finnish Language and Communication. University of Jyväskylä.
- Hurme, P. (ed.). 1980. *Puheentutkimuksen alalta - Papers in Speech Research 2*. Publications from the Institute of Finnish Language and Communication. University of Jyväskylä.
- Knapp, M. 1978². *Nonverbal communication in human interaction*. New York: Holt, Rinehart and Winston.
- Ladefoged, P. 1977. "Communicative and linguistic aspects of speech production". In Sawashima, M. and Cooper, F. S. (eds). 1977. 409-10.
- Lehtonen, J. and Hurme, P. 1978. Tutkimuslaitteet ja niiden käyttö. *Puheentutkimuksen menetelmät*, toim. J. Lehtonen. Jyväskylän yliopiston suomen kielen ja viestinnän laitoksen julkaisuja 17. University of Jyväskylä.
- Lehtonen, J. and Hurme, P. 1980. "The speech chain and the theory of speech". In Hurme, P. (ed.). 1980. 1-28.
- Lehtonen, J. and Sajavaara, K. (eds). 1979. *Papers in contrastive phonetics*. Jyväskylä Cross-language Studies 7. Department of English. University of Jyväskylä.
- Marslen-Wilson, W. D. and Welsh, A. 1978. "Processing interaction and lexical access during word redognition in continuous speech". *Cognitive psychology* 10. 29-64.
- Massaro, D. 1978. "A stage model of reading and listening". *Visible language* 12. 3-26.
- Öhman, S. 1975. "What is that we perceive when we perceive speech?". In Cohen, A. and Nooteboom, S. G. (eds.) 1975. 36-47.
- Raimo, I. and Suomi, K. 1978. "Pari kontrastiivista kuuntelukoetta". In Tikka, T. (ed.). 1978. 63-68.
- Sajavaara, K. and Lehtonen, J. (eds). 1979. *Papers in discourse and contrastive discourse analysis*. Jyväskylä Contrastive Studies 5.
- Sajavaara, K. and Lehtonen, J. 1979. "The analysis of cross-language communication: prolegomena to the theory and methodology". In Dechert, H. W. and Raupach, M. (eds). 1979.
- Sawashima, M. and Cooper, F. S. (eds). 1977. *Dynamic aspects of speech production*. Tokyo: University of Tokyo Press.
- Schiefelbusch, R. L. (ed.). 1978. *Bases of language intervention*. Baltimore: University Park Press.
- Tikka, T. (ed.). 1978. *Fonetiikan päivät - Joensuu 1976*. University of Joensuu.

THE SPEECH CHAIN

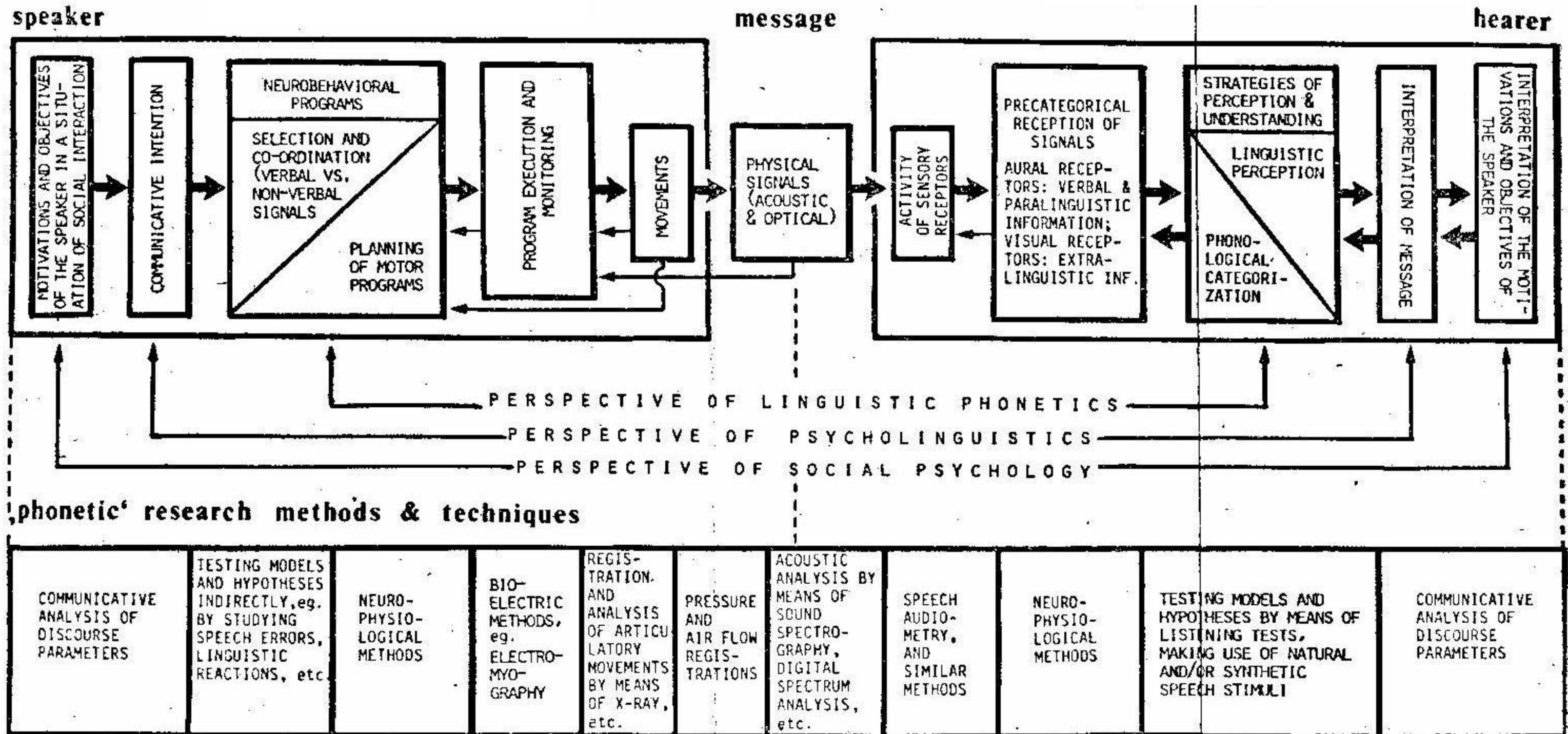
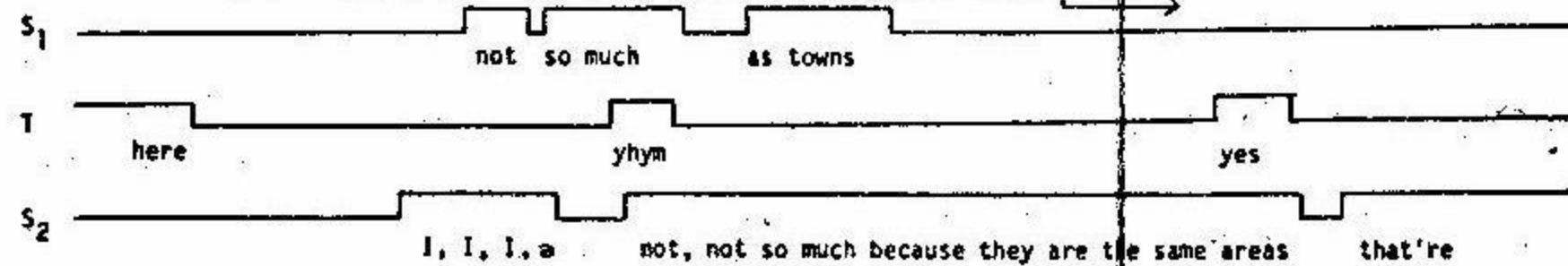
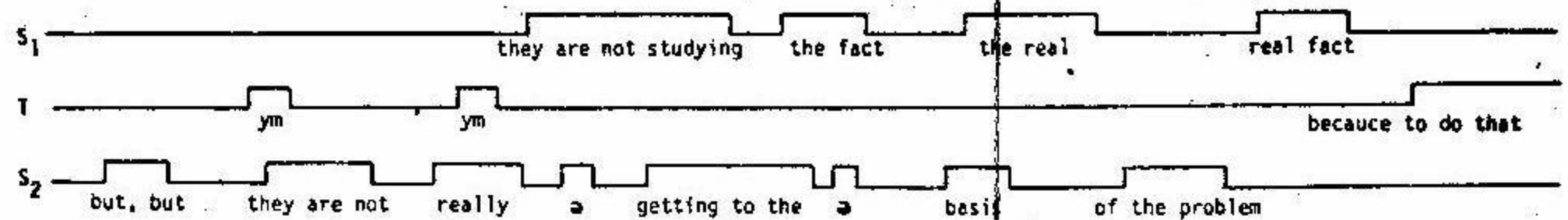


Figure 3. A model of the speech chain (from Lehtonen and Hurme 1979) and the experimental methods available to the analysis of each 'link' in the chain. The innermost perspective is typical of linguistically-based experimental phonetics. Expansion of the perspective adds aspects such as intentions and motivations of the speaker and hearer and, finally, the social situations as well as the communicative needs and expectations of the community. The arrows pointing from left to right stand for the feedback mechanisms on various levels of the processes. The entire chain should actually be bidirectional because in normal discourse the roles of speaker and hearer change continuously and, in addition, because the information which the speaker receives from the feedback behaviour of the listener also plays a role in the planning and processing of messages.

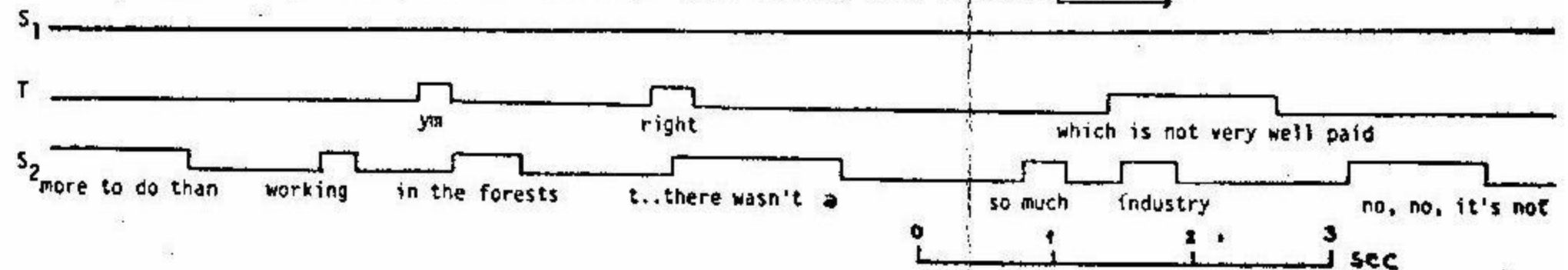
4. (discourse history:)
T summarizes S₁'s opinion on competition between urban and rural life-style. S₂ points out that circumstances in Finland differ from those in the article they had read; there are not such areas that people would consider bad or areas of poor people. Then T goes on "So that there is not a problem of, of areas that rise and fall - in terms of wealth here..."



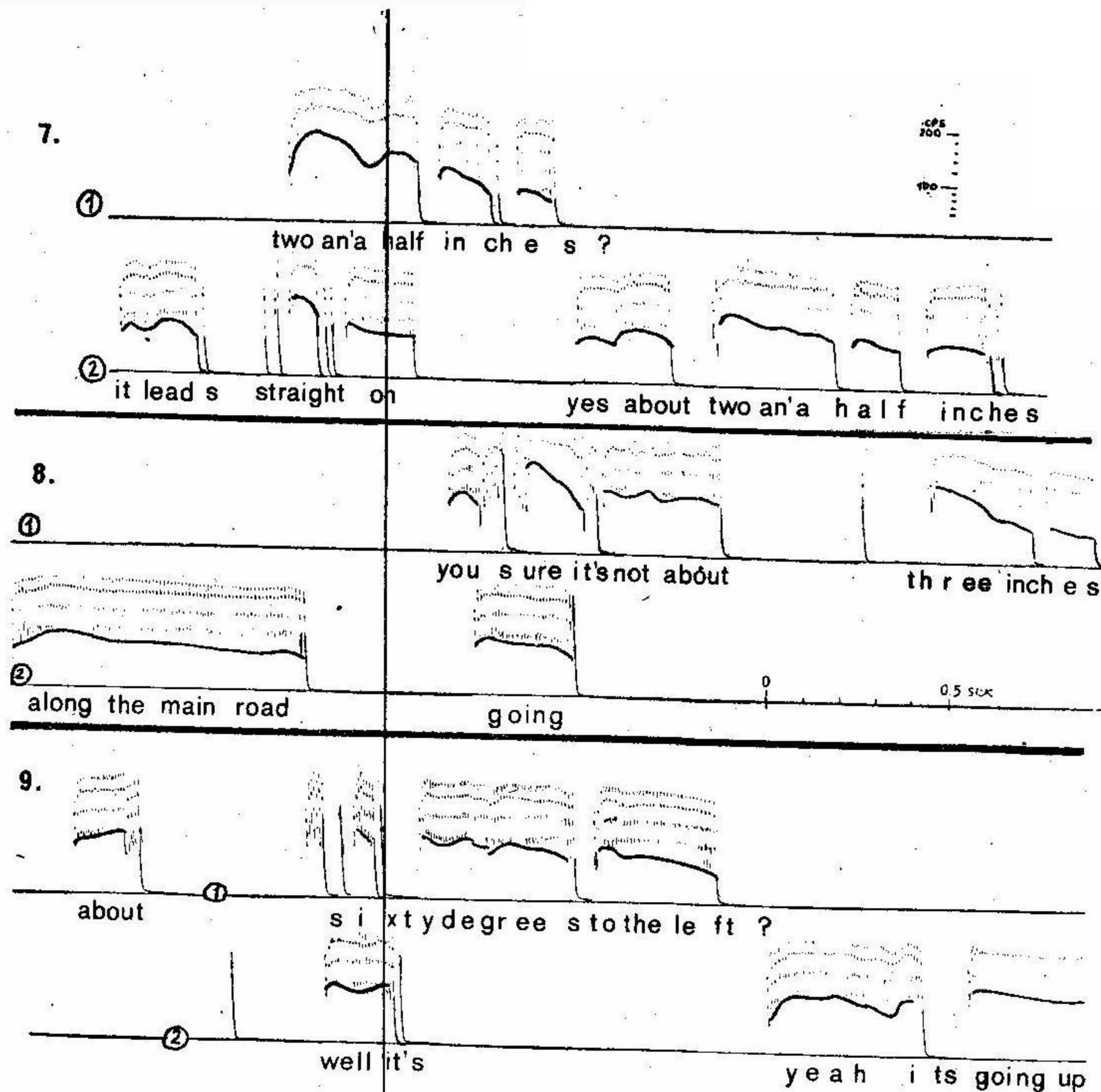
5. (discourse history:)
They are talking about investing programs stating their opinions on the way in which the writer of the article has outlined the programs. S₂ goes on "Yes but I, I think he tries to say that they are, you know, they are just making programs to show that they are doing something but..."



6. (discourse history:)
They are talking about agriculture. S₂ says that farmers stop keeping cows. T mentions unbalance of resources and thinks that the remaining problem is what the people in the farms do instead. S₂ goes on "Yes, yes it is, because t... on those areas there is, there isn't much more to..."



Figures 4, 5, 6. Three fragments of a discourse analyzed using a method in which the speaking turns of each participant are transformed into a binary on/off signal and record on paper by means of a level recorder. S₁ and S₂ are two Finnish students of English, T the British teacher.



Figures 7-9. Three fragments of a fundamental frequency recording of a spontaneous telephone discourse of two native British male speakers. In each fragment the participants speak simultaneously, which is a rather common phenomenon in natural discourse. Further on, in each case the move of speaker [7] must be interpreted as a yes-no question where a special intonation contour, "tune two", should be applied according to textbooks. Speaker (1) does not, however, use any rising fundamental frequency pattern but a falling pattern of a "declarative" sentence.