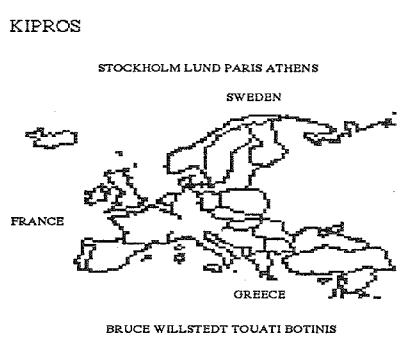
Lund University, Dept. of Linguistics
Working Papers 33 (1988), 23-50



Preliminary Report from the KIPROS Project

Gösta Bruce, Paul Touati, Antonis Botinis and Ursula Willstedt

INTRODUCTION

This paper is about a research project – the KIPROS project –, which started in January 1988 and which is planned to run for three years. The project is called Contrastive Interactive Prosody – with the acronym KIPROS based on the project title in Swedish – and has received financial support from the Bank of Sweden Tercentenary Foundation. The present paper is a status report from the first half year of the research work. In Bruce et al. (to appear) we give a summary of the research plan for the KIPROS project.

In what follows, we will briefly present the purpose and background of the project, give some attention to considerations about method concerning recording, analysis and synthesis and also include some general questions that we intend to find an answer to. After a brief section on prosodic structure in general, the next three sections deal with the three languages involved in the project: Swedish, French, and Greek. Each of these sections will contain an outline of the prosody of the language, a characterization of the speech material collected, and some preliminary observations about dialogue prosody. The final section will contain an outlook on future project work.

Purpose and background

KIPROS deals with how prosody is used in human spoken interaction. The purpose of the project is to investigate dialogue prosody in a contrastive perspective involving a few European languages that display interesting structural differences from a prosodic point of view: French, Greek, and two varieties of Swedish (South and Standard Swedish). One fundamental difference with consequences for the prosodic system of each of the languages is the varying degree and way to which these langauges exploit stress and accentuation (see the language specific sections below). The ultimate goal of our research will be the development of a model for French, Greek and Swedish prosody having interactive perspective. In such a model, we will have to relate prosodic properties to suitable interactive categories and on the basis of the analysis, propose rules for the generation of prosody in dialogues. The study of interactive prosody will also permit the calibration of our earlier research on prosody. By analyzing prosody in French, Greek and Swedish dialogues, we hope not only to elucidate the specific character of dialogue in these different cultures but also and most importantly to increase our understanding of the structure and function of prosody in these languages.

There are three important starting points for the KIPROS project. The first starting point is the research on prosody that has been conducted in phonetics in Lund through the years and that covers many aspects of prosody. Research on Swedish prosody led to the development of the so-called Lund model of prosody (cf. Bruce 1977, Bruce and Gårding 1978, Gårding and Bruce 1981, Gårding 1982, Bruce 1985). See also Gårding 1967 for an early study of prosody in spontaneous and read speech. The second starting point is the contrastive study of prosody conducted in Lund particularly by Gårding and directed towards the development of a general model of prosody. Examples from this research are Gårding 1981, Gårding, Botinis and Touati 1982, Gårding, Lindau, Norlin and Svantesson 1986 and most recently Touati 1987. A third starting point is the great interest in recent years taken in the study of interaction and dialogue analysis within linguistics and child language research in Lund (cf. Sigurd 1986, Söderbergh and Bredvad-Jensen 1987, Wulffson 1987).

Our current, accumulated knowledge of prosody is based mainly on the study of prosody in fairly simple, well controlled experimental situations, i.e. so-called laboratory speech. The kind of methodology that we have been advocating has thus typically been the simulation of a very simple dialogue – a question-answer-paradigm – where our informant has been playing the roles of both interlocutors (cf. Bruce 1977). Against the background of this research we now consider it possible and important to take the step from studying prosody in these simple situations towards investigating the role of prosody in communicatively more

natural and relevant, but also more complex situations such as in interactive speech and dialogue.

METHOD

Recording

In the study of prosody in a dialogue perspective we are as phoneticians faced with a number of new problems. The first major concern of the KIPROS project is the choice of suitable tape recorded material for phonetic and acoustic analysis. A basic requirement is, of course, that the technical quality of the recording, which also includes conversation on telephone, has to be high enough for instrumental analysis to give satisfactory pitch contours, etc. As for the choice of dialogue type we are presently searching along the dimension constructed – authentic dialogue; i.e. on the one hand experimentally arranged and more controlled situations, where the topic and the development of the dialogue is predictable to some extent, and on the other hand situations where this is more or less not the case, for example dialogues recorded from radio and TV programs. The actual choice of dialogue has also to be related to our interest in those dialogues where prosody is expected to play an especially important role for the development of the dialogue.

In a dialogue situation, non-verbal signals of interaction are likely to be given by vocal and prosodic as well as by somatic means. For the right appreciation of the role of prosody in an interactive context it may, therefore, be important to be able to control the contributions to the dialogue of somatic properties. Although we will not be able to do that in all cases, there are at least two possible ways of doing so. One possibility is to make a video recording of the dialogue together with the regular audio recording. Another possibility is to choose a non-face-to-face situation, for example a telephone conversation, where somatic cues cannot be exploited and where vocal and prosodic cues are a priori likely to be more exploited for interactive purposes than in a face-to-face situation. See, however, Schaffer 1984 for data contradicting this expectation.

In the present first period of the project we have been recording and studying different types of dialogue for each of the three languages. We have been doing this in order to explore the field and get an overview and to be able to choose the most interesting types of dialogue for special study. For Greek we started recording a constructed type of dialogue in the laboratory, for Swedish we have been recording a more authentic kind of dialogue from a well known radio program, and for French we have begun studying the most authentic types of dialogue about more existential topics. Although we start at different points for the three languages, we eventually intend to cover the same dialogue types for French, Greek and Swedish.

Analysis

The second major concern of the project is the analysis of the recorded material. The starting point for the analysis is a broad transcription of the recorded material in a mainly orthographic form. In the transcription special attention is given to prosodic categories, while the transcribing of the segmentals will serve mainly as a reference and line-up point for prosodic events. Although we are not covering any details about segments in the transcription, we include information about repetitions of segments/syllables, hesitations, simultaneous speech, etc.

The analysis phase consists of three different parts. The first kind of analysis is a phonological analysis in the form of a selective, prosodic transcription based on repeated listening to the tape recorded material. Our aim here is to cover for each of the three languages involved in the project those prosodic categories – accentuation, phrasing, boundary signalling etc. – that are relevant for the study of dialogue prosody. The basic framework for prosodic analysis in Swedish, French and Greek are given in Bruce 1977, 1985, Touati 1987 and Botinis (forthcoming), respectively. See also Gårding et al. 1982 for a comparative study of intonation in the three languages.

The second kind of analysis that we will attempt based on the transcribed material is an analysis of the structure of dialogue. Our aim here initially is a fairly rough analysis, selecting those aspects that may be revealing for the analysis of prosody. We do not intend to develop a model of dialogue structure of our own, but we will instead have to rely on other persons' work. We are presently considering different models for dialogue analysis. Particularly interesting from our point of view are the following two kinds of analysis emphasizing different aspects of discourse structure: the ideas by Anward 1986 about language games and specifically his suggestions about hierachically organized topic structure and the initiative and response analysis of turn taking presented in Linell and Gustavsson 1987. Other studies that have attracted our attention and that we will consider are the pioneer work within discourse analysis performed by the Birmingham group with special emphasis on discourse intonation (cf. for example Brazil 1985) and the work on discourse structure within computer linguistics by Grosz and Sidner 1986.

We consider the phonological analysis of prosodic categories based on listening to be a necessary prerequisite for the third and most important kind of analysis of the project work: the qualitative and quantitative study of prosodic patterns – mainly tonal and temporal – from acoustic recordings of Fo and speech waveform. The procedure now is to have the recorded material digitized on the VAX 11/730 at our laboratory and analyzed using the API program of the ILS package, where pitch extraction is done with a modified cepstral technique. So far in the project work we have been using pitch information combined with

waveform information for prosodic analysis. The particular kind of problem we encounter here is extracting pitch from telephone recordings and knowing whose pitch is being tracked in the simultaneous speech of two speakers.

Based on the segmentation of acoustic recordings into prominent syllables and prosodic groups, a first part of the qualitative study consists in isolating relevant pitch patterns for accentuation, phrasing, boundary signalling, etc. The next part will be to try to relate these pitch patterns to the phonological analysis of the prosodic transcription – with a possible revision of the transcription – and also to relate these patterns to adequate interactive categories. Starting from the qualitative analysis, we will do a quantitative analysis (measurements) of what may be especially important, particularly for the development of a model for dialogue prosody. So far we have only been doing qualitative analysis.

Synthesis

A third important concern of the project will be the task of formulating rules for the use of prosody in spoken interaction. We will test these rules perceptually through editing of the relevant parameters of recorded speech and speech synthesis. By isolating stereotypical patterns of interactive prosody through the use of speech synthesis, we intend to be able to evaluate the optionality and suitability of different prosodic markers for interactive categories. One step in this analysis by synthesis may be the use of stylization of Fo contours for the establishment of the relevant pitch patterns, where the equivalence of Fo contours will be evaluated perceptually. This method has been successfully exploited for the analysis of intonation by the Dutch school (cf. for example 't Hart and Collier 1975).

The three phases of recording, analysis and synthesis occur logically one after the other. In the project work they will recur cyclically. We will actually try to come to the third phase – rules and synthesis – fairly soon after having made a careful analysis of a relatively small speech sample. In this way we hope to get an idea fairly soon about the degree of success in our analysis of dialogue prosody.

GENERAL QUESTIONS

The kind of general questions that we are asking now when studying prosody in an interactive perspective are centered around how pitch is being exploited.

One fundamental question is whether the same, well-known pitch patterns in the three languages that are familiar to us from laboratory speech experiments also occur in spontaneous speech and dialogues, and whether we might also come across new pitch patterns that we have not encountered in lab speech situations. Another question that we find important is how successive utterances making up one speech unit are organized as regards pitch by one of the speakers when (s)he

has got the turn and by the two speakers in interaction and cooperation. The next obvious question will then be to ask how these pitch patterns are related to dialogue structure and interactive categories: how are conversation topics organized as regards pitch; what are, for example, possible cues for introducing a new topic, continuing on a topic, rounding off a topic and maybe introducing a new aspect of a topic? Furthermore, how are turns regulated in terms of pitch; what are possible turn-taking, turn-keeping, turn-inviting and turn-yielding pitch cues?

These are only a few of the general questions that we intend to find an answer to in the course of our project work.

PROSODIC STRUCTURE

In our treatment, prosody will be generally divided up into rhythm (stress), accentuation and intonation. For a more detailed account of prosodic structure see Bruce 1985.

Rhythm refers to the weighting and grouping of successive syllables of speech with the dichotomy stressed - unstressed being potentially basic. Among the phonetic correlates of the rhythmical organization of speech, temporal properties are considered primary.

From a functional point of view, stress and accentuation have in common that they express prominence relations. But while stress is a relation between syllables, accentuation appears to have a larger domain, the stress group, which, besides the stressed syllable, may also include related unstressed syllables. This means that accentuation is dependent on rhythmical structure, where stress group boundaries are important points of coordination. The presence versus absence of an accent in a stress group is one way of signalling the importance of the word (phrase) related to the actual stress group. Phonetically accentuation is primarily expressed by pitch relations.

While both accentuation and intonation have to do with the pitch properties of speech, we take intonation proper to mean pitch relations in a speech unit with the exclusion of pitch prominence relations, i.e. accentuation. Intonation may therefore be related to phrasing, boundary signalling, expressing speaker attitude and various textual and interactive categories including topic organization and turn regulation. Intonation is, however, dependent on accentuation in an indirect way. According to this view, overall intonation is primarily made up of interrelations between local pitch gestures related to accentuation and boundary signalling, i.e. the relations between successive pitch gestures including their range and register.

SWEDISH

Outline of Swedish prosody

In the present outline of Swedish prosody we will try to present the basic characteristics of Standard Swedish prosody – especially in contrast to French and Greek prosody – as we know it from our experience about laboratory speech. So far in the project work we have only been concerned with the study of the prosody of Standard Swedish and not yet with the prosody of South Swedish. The purpose is to have this outline as a reference to which we will be able to relate more easily our findings about dialogue prosody.

Characteristic of Swedish prosody is the key role played by stress. Lexical stress in Swedish is not fixed. Although clearly constrained by morphology and phonology, the location of stress in a polysyllabic Swedish word is still to some extent unpredictable. Among the main factors determined by stress are the occurrence of quantity – V:C or VC: – and the phonetic quality of the vowel. Characteristic is also the temporal realization of a stressed syllable in relation to an unstressed syllable with a large stressed/unstressed ratio like for example in English but unlike Greek (cf. Dauer 1983 and below).

Accentuation also plays a major role in the prosodic system of Swedish. In the first place it can be divided up into non-focal, mere (word) accentuation and focal (phrase) accentuation. A stress group in Swedish will normally be but need not be characterized by the presence of accentuation, which may signal for example whether it belongs to an independent word (accent) or is part of a compound word or a larger phrase including lexicalized phrases (no accent).

The most characteristic and well known feature of Swedish accentuation is that an accent has to be one of two possibilities, either accent I (´, acute) or accent II (`, grave), where the choice of accent is largely determined by the phonology and morphology of the actual word form. The two word accents appear to have a distinctively and critically different timing of the pitch gesture relative to the stress, accent I being timed earlier (HL*) than accent II (H*L).

Focal accentuation (marked by capital letters in Figure 1) is an extra prominence given to the accent (and the actual word) and is marked by an extra pitch gesture within the actual stress group (focal accent I = HL*H, focal accent II = H*LH), usually in combination with increased duration of the actual stress group.

The broad pitch characteristics of the accentual categories of Standard Swedish can therefore be described as follows:

accent I <	non-focal	step down in pitch to stressed syllable
	focal	pitch rise through stressed syllable
accent II <	non-focal	pitch fall through stressed syllable
	focal	added pitch rise after grave accent fall

Even if no particular focus or contrast is being elicited in an utterance, the occurrence of at least one focal accent per utterance is still considered (almost) obligatory in Standard Swedish.

Focal accent appears to be decisive for yet another specific characteristic of Standard Swedish prosody: the presence or absence of a downstepping pitch contour. In a pre-focal position, up to the focal accent of a phrase (or a whole utterance) there is typically no downstepping, but instead successive pre-focal accents occur on more or less the same pitch level. However, after a focal accent, the downstepping of successive post-focal accents is a characteristic pitch pattern. While other languages like Greek and French typically have deaccentuation in a post-focal portion of an utterance, post-focal accents of Standard Swedish are still retained and realized as part of a staircase pattern (cf. Bruce 1982).

Speech material

For Swedish we have up to now concentrated on recording and analyzing one dialogue type taken from a popular radio program 'Ring så spelar vi'. It is an entertainment program, where radio listeners' conversation on telephone with the program leader, a well-known radio voice, is broadcasted and the favourite record of the person calling up is played. The conversation also includes a contest, (with the possibility of winning one or several records) where the person calling up has to answer a question about a general topic.

Typical ingredients of such a dialogue are in time order: presentation, ordering of favourite record, getting acquainted, contest, finishing up. Characteristic contextual features of this dialogue type are the following: it is a non-face-to-face situation, i.e. where somatic cues are not used; the length of each dialogue is limited in time; the choice, number and duration of conversation topics is to some extent predictable; the speakers do not know each other; it is likely that it is the program leader who governs the conversation and is the dominant speaker.

We have recorded a few radio programs of this type, and we have so far concentrated on one particular dialogue (6 - 7 mins), where both interlocutors

speak Standard Swedish with special study devoted to a section of about 50 secs. The pitch contour of this section of the dialogue is illustrated in Figure 1.

Preliminary observations

The observations about dialogue prosody in Swedish that we present here should be taken as very preliminary and more as working hypotheses than as findings or results. The material studied so far is very limited, and it is therefore impossible to generalize from these observations. At present our interest is centered around how pitch is being exploited.

Starting from the general questions asked above, we note as a first general observation that the two speakers seem to use the same pitch patterns for accentuation that are typical of Standard Swedish in read, laboratory speech also in spontaneous speech of the kind recorded here. Although our analysis is so far only qualitative, there is nothing to contradict our expectations about the specific timing and direction of change of the local pitch gestures for non-focal and focal accents as described above.

It is also interesting to note that there are several instances of downstepping and non-downstepping pitch patterns, which seem to be triggered by the placement of focal accent in very much the same way as described above. An apparent example of downstepping after an early focal accent is found in Figure 1:1 and an example of non-downstepping before a late focal accent in Figure 1:2.

A point for further investigation in the project is the pitch relations of back channel items. It is a possible hypothesis that the choice of level and range of pitch in a back channel item – for example *ja, mm* etc. – is very much in agreement with the level and range of the utterance (or maybe the last part of that utterance) produced by the other speaker that the back channel item is supporting. An example of such pitch agreement in back channelling – three successive *ja* by speaker B – is seen in Figure 1:2-3.

Another point that we intend to investigate in the project is the regulation and exploitation of pitch range for interactive purposes, its relation to topic structure, turn-taking, etc. A default assumption seems to be that the introduction of a new topic is signalled by a widening of pitch range (cf. Brown et al. 1980, Hirschberg and Pierrehumbert 1986). In the present section we have an example of an expansion of pitch range apparently signalling the rounding off a topic, while the actual introduction of the new topic is accompanied by a somewhat narrower pitch range (see Figure 1:8) and an example of a clear compression of the pitch range for introducing a new, special aspect of a topic (see Figure 1:9).

It is obvious that variation and change of pitch range may generally be used by the speaker for getting the listener's attention, but if and how there is a more tight connection to interactive categories remains to be examined.

FRENCH

Outline of French prosody

The following is a brief sketch of French prosody in an utterance perspective. The manifestation and representation of prosodic categories in French in an utterance perspective have been presented in detail in Touati 1987.

A negative way of describing French prosody compared with Swedish and Greek prosody is to say that there is no functional lexical prosody. The functional domain for French prosody is a phrase composed of at least one lexical word plus possible preceding grammatical words.

The main relevant categories in an utterance perspective are phrase accent (accent de groupe), focal accent (accent d'insistance) and initial juncture. Phrase accent can be identified with final phrase juncture and has a demarcative function. It is by default assigned to the final syllable of the phrase. The focal accent which is optional can be related to some kind of semantic or pragmatic prominence and in a way to a marked initial juncture. It is typically assigned to the initial syllable of a word or phrase.

The pitch representations of these categories are, for phrase accent, a pitch rise (LH) in utterance non-final groups and a pitch fall (HL) in an utterance final group, for focal accent, a pitch rise (LH) and for initial juncture L. The pitch gesture for phrase accent is accompanied by a marked lengthening of the phrase final syllable only in utterance final position. There is usually no lenghtening of a syllable that is assigned focal accent (accent d'insistance). When focus (or contrast) occurs on a non-final constituent of an utterance, deaccentuation typically applies after focus as in Greek but unlike Standard Swedish.

Speech material

Our aim in undertaking a prosodic analysis of some French dialogues was two-fold; first we decided to adopt a global approach to the understanding of the prosodic stucture of spontaneous speech by investigating several conversational modes and second we tried to solve some experimental problems, i.e. the extraction and representation of long term varying fundamental frequency as found in conversations.

The recordings are made from broadcast conversations (face-to-face situation, public audience). The data is from three different types of conversations: an informal conversation (D1), a casual interview (D2) and a political debate (D3).

Different speakers were operating in the three types of dialogue (modes); the informal conversation was between an adult female and a girl of about five years, the casual interview between an adult male reporter and a ten-year old young actress and the political debate between two political veterans. The use of

different sets of speakers raises some problems about the cross-mode comparison.

The data may be classified in various ways, but the main difference is the conversational mode: the relationship between speakers was very friendly in the intimate conversation, neutral in the interview and antagonistic in the debate. In the informal conversation, the adult woman was the dominant partner (she conducted the conversation by choosing the successive topics); in the interview, there was no direct dominant partner, each speaker knew and played his/her more or less fixed part: the reporter asks the young actress questions about the play and she answers; in the political debate, each speaker attempted to say as much as possible during his speaking time (a turn was more like monologue: a read text or short political speech).

An orthographic transcription of all the three dialogues in their entirety was made. From the 2 minutes and 43 seconds of recording of D1, the 2′37" of D2 and 7′48" of D3, our ILS analysis covered 36 seconds for D1 and 64 seconds for both D2 and D3. Figures 2 and 3 illustrate pitch contours from D1 and D2, respectively.

Preliminary observations

At the present stage of our work, any conclusive explanation about French prosody in different conversational modes will be premature. Preliminary observations suggest, however, that temporal phenomena such as the relationship between speech duration and non-speech duration, especially duration and distribution of pauses, and lengthening at boundaries should been taken into account in order to describe prosodic differences between the modes; in D1, both pitch change and pause were used to indicate the final junctures of the prosodic unit as shown in Figure 2:1-2-3 (for a study of pauses in three discourse modes in French, see Duez 1982). But as noted above, for the moment our interest is centered around pitch.

Our first general comment about pitch is that speakers in a dialogue, even the youngest one, use pitch contrasts found in read speech like rising vs falling contours for phrase accent. Differences in rising pitch gestures also suggest a possible contrast between minor and major rising contours (see for example Figure 2:1-2). Apparently, these contrasts are associated with some kind of continuation versus finality. A point for further investigation is how these contrasts are related to topic structure or/and turn-taking.

Another interesting point yet to be explored is the manifestation and the domain for focal accentuation. Our present project work suggests the possibility that the domain for focal accent is larger than earlier assumed (cf. Touati 1987:28). Moreover, in the political debate, variation of pitch range is exploited

more than in the two other types of dialogue in order to contrast focal with non-focal information. This is accomplished by mainly varying the pitch peaks. No clear case of deaccentuation after focus with extinction of pitch gestures was found in the present material.

An often recurring global pitch pattern consists of an updrifting - downdrifting contour, even for utterances which are semantically questions (see D2 Figure 3:1-2-6). An interpretation of the local pitch movements being superposed on this global pitch pattern seems favourable. Further work on the interplay between local and global pitch patterns is being planned.

GREEK

Outline of Greek prosody

A detailed account of the phonology and phonetics of Greek prosody in an utterance perspective will be given in Botinis (forthcoming). The presentation here covers some main features of Greek prosody in outline.

Stress plays a major role in the prosodic system of Greek as in Swedish but unlike French. The domain of word stress is the grammatical word. Its placement is not fixed but is morpholexically conditioned. It is subject to the trisyllabic constraint, i.e. word stress may not appear to the left of the antepenultimate across the word paradigm – it has to occur on one of the three final syllables of a word – as well as to the monotonic principle, i.e. every word may have only one word stress.

The domain of phrase stress is the grammatical phrase, and it may appear in a word+an enclitic environment. Phrase stress has to take into account the position of word stress. Thus when stress is to the left of the antepenultimate syllable with regard to the phrase boundary it does not move to the right but rather another stress appears, the phrase stress, two syllables to the right of the word stress whether this is on the lexical element or on the enclitic particle. Thus the trisyllabic constraint is applied to both word and phrase level but its effect is on the one hand to move stress to the right at the word level and on the other hand to produce an additional stress at the phrase level.

Both word stress and phrase stress contribute to the rhythmic structuring of Greek spoken language, phrase stress is also related to syntactic phrasing. Stress in Greek will affect the phonetic quality of the vowel only marginally. A stressed syllable in relation to an unstressed syllable – be it word stress or phrase stress – is characterized by increased duration and intensity (the energy integral), with however, a rather small stressed/unstressed ratio compared to Swedish.

A stress group in Greek will normally be accented and accompanied by a pitch gesture signalling the relative importance of the word related to the actual stress group. The pitch realization of an accent in Greek is typically a pitch rise through the stressed syllable in a phrase non-final position, and in a phrase final position, the addition of a pitch fall, which may partially replace the pitch rise.

Additional focal accentuation may occur on any constituent of an utterance, although there are constraints in enclitic constructions containing a phrase stress. However, focal accent does not occur by default in a Greek utterance where no particular focus or contrast is being elicited. The addition of focal accent to a constituent is expressed as a widening of the range of the accentual pitch gesture of the word in focus as well as a reorganization of the whole utterance at the global level. In particular, deaccentuation typically occurs in an utterance after focus as in French but unlike Swedish.

Speech material

This is the first investigation of a series of experiments on the prosody of discourse structure in Greek.

The speech material in this experiment is a constructed dialogue containing five different topics, where the first topic about a book has been extracted for special study. The length of this section is about 50 seconds. The participants in the dialogue were two academic males in their mid thirties, brought up and educated in Athens, one of which was the experimenter. They have had no speech or physical deficiencies and belong to largely the same sociocultural environment. The participants know each other quite well. The recording took place in a face-to-face interaction in a quiet room at the phonetics laboratory in Lund. Although the experimenter had a major interest in the outcome of the dialogue and was governing the dialogue to some extent, both participants appear to have contributed equally towards the development of the discourse. Figure 4 gives the pitch contour of the initial section (first topic) of the dialogue.

Preliminary observations

The familiar pitch patterns with pitch changes (pitch rises as a rule) associated with word stress and accentuation as well as extended pitch rises+falls for focal accentuation are easily recognized in this material.

At the discourse level, pitch range appears to be wider at the beginning and decreases gradually for comparable parts of speech in the course of the topic, quite dramatically at the end of it (Fig. 4). Subtopics within a topic seem to display the same, general pitch structure with wider pitch range at the beginning and a narrower one at the end as well. The phrases by talker B *edéli mu to stílane-alá mu to stílane-ke mu to stílane* with partial repetition of the same words constitute a typical example of pitch range declination through a subtopic (Fig. 4:2-4).

Pitch range may also contribute to expressing a main ~ subordinate division of a subtopic with larger ~ smaller pitch range, respectively. An example is constituted by the subordinate subtopic *perímena, perímena, perímena pára polí keró fisiká,* which has a narrower pitch range than the main subtopic *edéli mu to stílane-alá mu to stílane* (Fig. 4:3). Besides the pitch range differences for the main ~ subordinate division there appear to be abrupt, local pitch changes signalling the boundaries of this division. The subordinate subtopic may appear in a parenthetical expression like in the above example or to the right of the main subtopic as in B: *éxune paráδosi, ne, γiatí i érevna éxi arxísi eδó ke δío eónes perípu.* (Fig. 4:7).

Focal Accentuation, usually associated with new information, may also occur on shared as well as repeated information as in the phrase *...éxune paráδosi...* repeated by both speakers (Fig. 4:6-7). Moreover, post-focal portions of an utterance are not characterized by complete deaccentuation but appear to be rather varied in pitch according to the discourse requirements.

OUTLOOK

From the present sketchy and tentative suggestions about prosody in a dialogue perspective we will now proceed to a fairly detailed analysis of the data collected. Our strategy at present will be to give a thorough account of a rather restricted speech material concerning relevant aspects of dialogue structure, phonological analysis of prosodic categories as well as phonetic - acoustic aspects of prosody and their interrelations for each of the three languages. This will give us the possibility of formulating specific albeit tentative rules for prosody in a dialogue context. These prosodic rules may then be tested on new speech material by the use of speech synthesis. This analysis by synthesis approach will be one way for us to evaluate the relative success of our work and will also serve as a guide-line for collecting new speech material. We further assume that this modeling of prosody will be helpful in the contrastive study of Swedish, French and Greek interactive prosody.

REFERENCES

Anward, Jan. 1986. 'Emotive expressions'. *Papers from the Ninth Scandinavic Conference of Linguistics,* ed. Östen Dahl, 39-52. Department of Linguistics, University of Stockholm.

Botinis, Antonis. forthcoming. *The prosodic structure of Greek.*

Brazil, David. 1985. *The communicative value of intonation in English.* Discourse Analysis Monograph. ELR, University of Birmingham.

Brown, Gilian, Karen Currie and Joanne Kenworthy. 1980. *Questions of intonation.* London: Croom Helm.

Bruce, Gösta. 1977. *Swedish word accents in sentence perspective.* Lund: Gleerup.

Bruce, Gösta. 1982. 'Developing the Swedish intonation model'. *Working Papers* 22, 51-116. Lund: Dept. of Linguistics.

Bruce, Gösta. 1985. 'Structure and functions of prosody'. *Proceedings of the French Swedish Seminar on Speech*, eds. Bernard Guerin and René Carré, 549-559. Grenoble.

Bruce, Gösta and Eva Gårding. 1978. 'A prosodic typology for Swedish dialects'. *Nordic Prosody,* eds. Eva Gårding et al., 219-228. Dept. of Linguistics, Lund University.

Bruce, Gösta, Ursula Willstedt, Paul Touati and Antonis Botinis. 'Dialogue prosody'. To appear in *Working Papers* 34. Lund: Dept. of Linguistics.

Dauer, Renate. 1983. 'Stress-timing and syllable-timing reanalysed'. *Journal of Phonetics* 11, 51-62.

Duez, Danielle. 1982. 'Silent and non-silent pauses in three speech styles'. *Language & Speech* 25, 11-28.

Gårding, Eva. 1967. 'Prosodiska drag i spontant och uppläst tal'. *Svenskt talspråk,* ed. Gösta Holm, 40-85. Stockholm: Almqvist & Wiksell.

Gårding, Eva. 1981. 'Contrastive prosody: A model and its applications'. *Studia Linguistica* 35, 146-166.

Gårding, Eva. 1982. 'Swedish Prosody'. *Phonetica* 39, 288-301.

Gårding, Eva, Antonis Botinis and Paul Touati. 1982. 'A comparative study of Swedish, Greek and French intonation'. *Working Papers* 22, 137-152. Lund: Dept. of Linguistics.

Gårding, Eva and Gösta Bruce. 1981. 'A presentation of the Lund model for Swedish intonation'. *Nordic Prosody II,* ed. Thorstein Fretheim, 33-39. Trondheim: Tapir.

Gårding, Eva, Mona Lindau, Kjell Norlin and Jan-Olof Svantesson. 1986. 'Final report: 'Phonetic analyses of some non-European languages' (LUCLA)'. *Working Papers* 29, 115-138. Lund: Dept. of Linguistics.

Grosz, Barbara and Candace Sidner. 1986. 'Attention, intentions, and the structure of discourse'. *Computational Linguistics* 12, 175-204.

't Hart, Johan and René Collier. 1975. 'Integrating different levels of intonation analysis'. *Journal of Phonetics* 3, 235-255.

Hirschberg, Julia and Janet Pierrehumbert. 1986. 'Intonational structuring of discourse'. *Proceedings of the 24th Meeting of the Association of Computational Linguistics,* 136-144. New York.

Linell, Per and Lennart Gustavsson. 1987. *Initiativ och respons. Om dialogens dynamik, dominans och koherens*. SIC 15. Studies in Communication. University of Linköping.

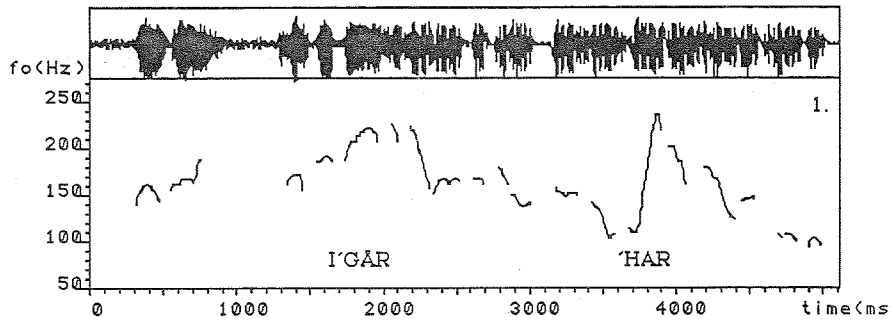
Schaffer, Deborah. 1984. 'The role of intonation as a cue to topic management in conversation'. *Journal of Phonetics* 12, 327-344.

Sigurd, Bengt. 1986. 'Computer simulation of dialogue and communication'. *Proceedings of the 3rd Conference of Scandinavian Computational Linguistics*, ed. Fred Karlsson. Department of Linguistics, Helsinki.

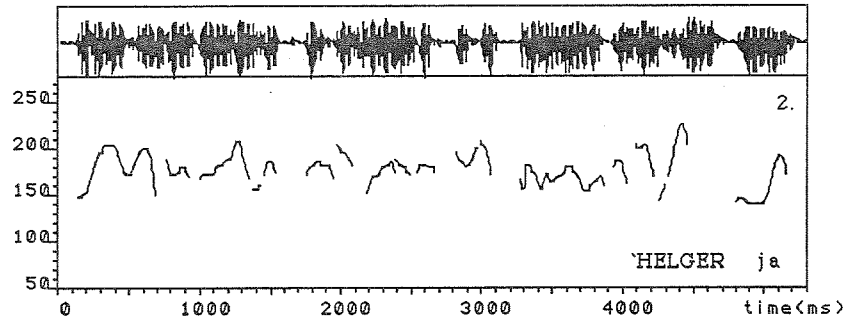
Söderbergh, Ragnhild and Anne-Christine Bredvad-Jensen. 1987. 'A dialogue model and its application in a free play situation'. *The Nordic Languages and Modern Linguistics* 6, eds. Pirkko Lilius and Mirja Saari. Helsinki: Helsinki University Press.

Touati, Paul. 1987. *Structures prosodiques du suédois et du français*. Lund: Lund University Press.

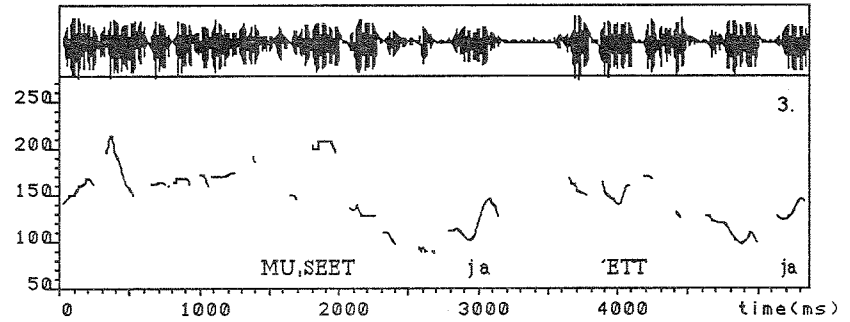
Wulffson, Madeleine. 1987. 'Discourse intonation in Swedish: an exploratory study in communication games'. *Working Papers* 31, 175-200 Lund: Dept. of Linguistics.



A: ja ja e, ja I'GÅR så `fråga dom mej på 'Rotary så 'här: 'HAR 'du ci'vila 'kläder?

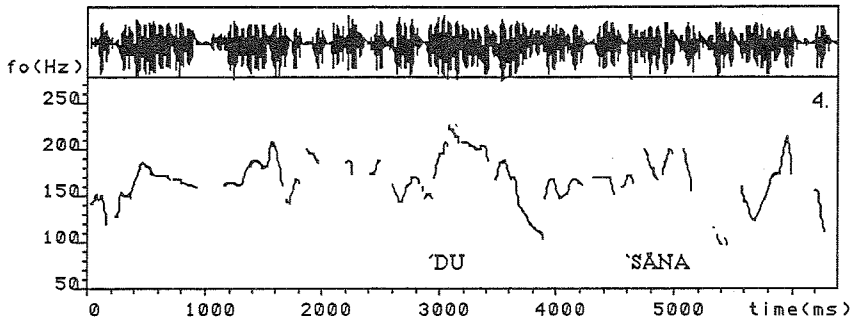


A: o de e `nämligen 'så att ja e `väldit enga'gerad im i`mycke 'här `även på 'HELGER.
B: ja

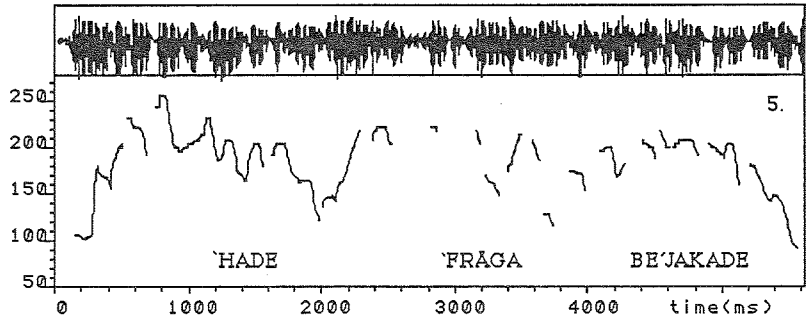


A: o ja e `före,ståndare för GARNI`SONSMU,SEET 'här i 'Boden
B: ja
A: för att ta 'ETT e'xempel `bara
B: ja

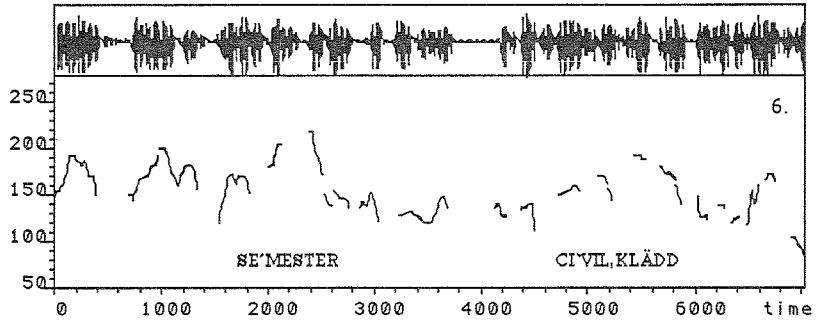
Figure 1: 1 - 3. A section of a Swedish dialogue; waveform (above), pitch contour (centre) and orthographic transcription (below). Key words are aligned with important pitch events.



A: o o de ´gör de att ja e `väldit `mycke i uni´form också
B: ´DU e unifor´merad
A: ä även vi `SÅNA `till,fällen alltså
B: ja just det

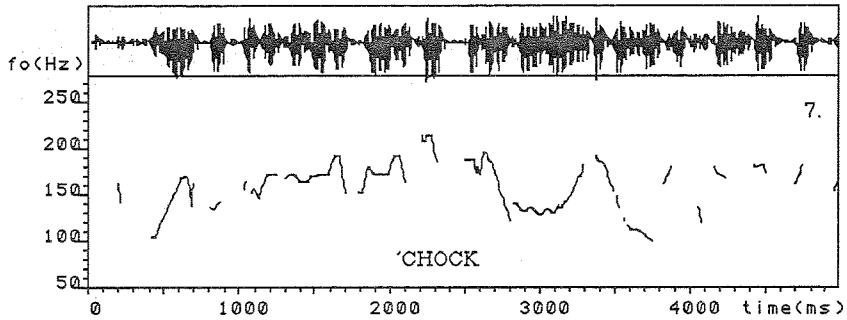


A: ja
B: do dom `fråga om du `HADE ci´vila ´kläder.
A: jo, dom frå `FRÅGADE mej ´de
B: o du sv..
A: därför att..
B: ..BE´JAKADE `frågan.

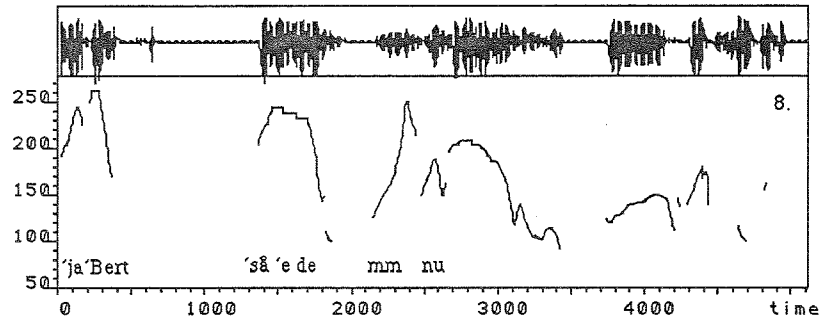


A: jo, ja ja va `nämligen.. hade SE´MESTER den ´här `veckan ´nu ,
o och ´därför så va ja CI`VIL,KLÄDD på `ROTARY SAMMAN,TRÄDET i´går.

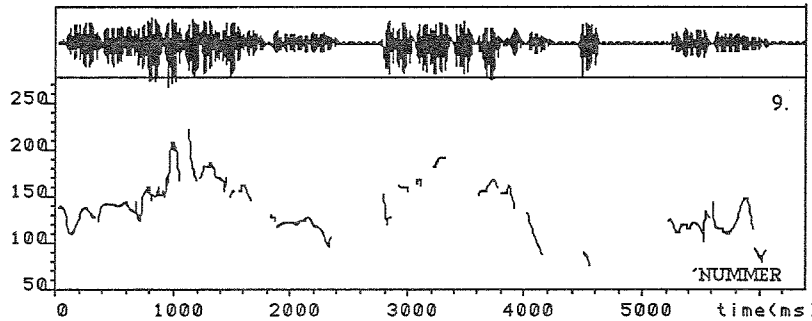
Figure 1: 4 - 6 (cont.)



A: he he
B ja, så de de ´blev en ´smärre ´CHOCK för `del,tagarna.
A: ja, de ´blev de (laughter)



B: ´ja ´Bert, ´så ´e de.
A: mm.
B: nu ska du vinna en `skiva.
A: nja, de e `inte ´säkert.



B: nja, de e ju `ALL.. A: ja kan FÖR´SÖKA. B: ..DELES ´säkert e de väl `inte men
A: ja kan FÖR´SÖKA.
B: ´JA, de ska du `göra.
A: ja
B: va vill du ha för ´NUMMER?

Figure 1: 7 - 9 (cont.)

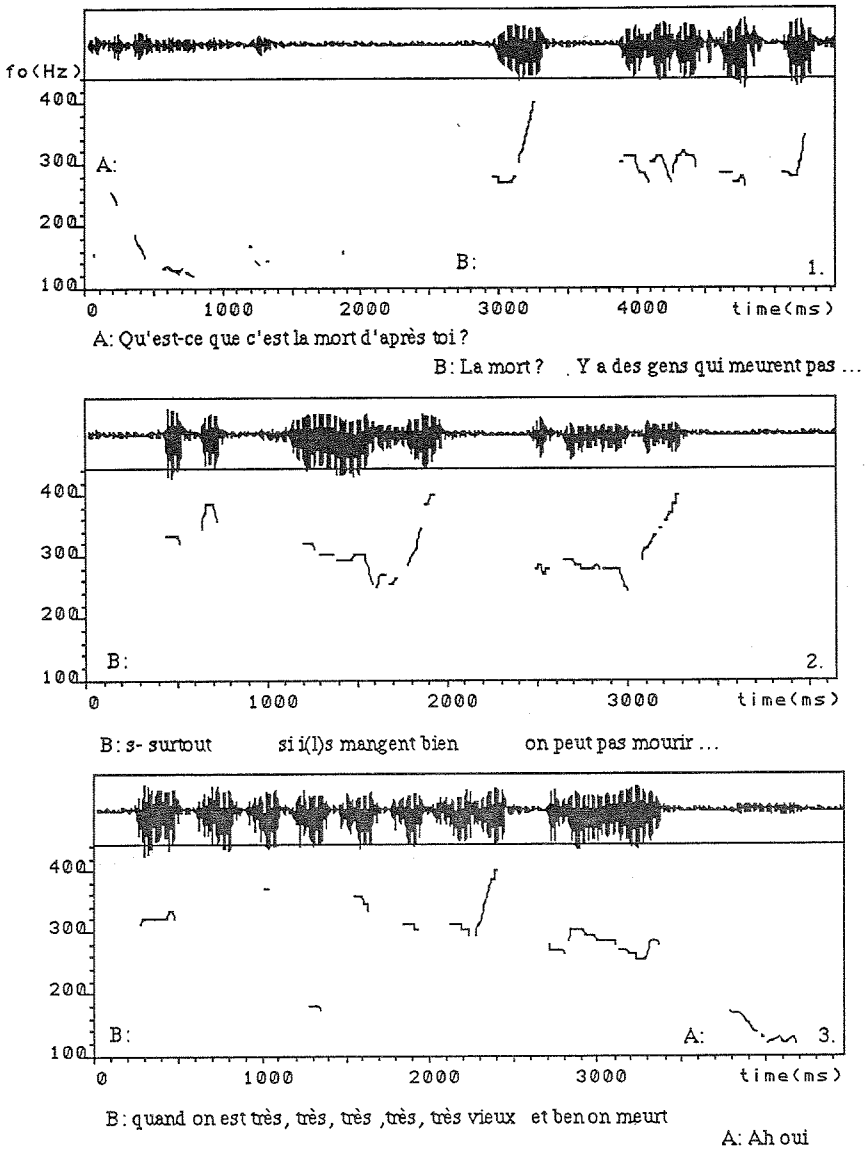


Figure 2. A section of a French dialogue (D1:1-3), waveform (above) pitch contour (centre) and orthographic transcription (below).

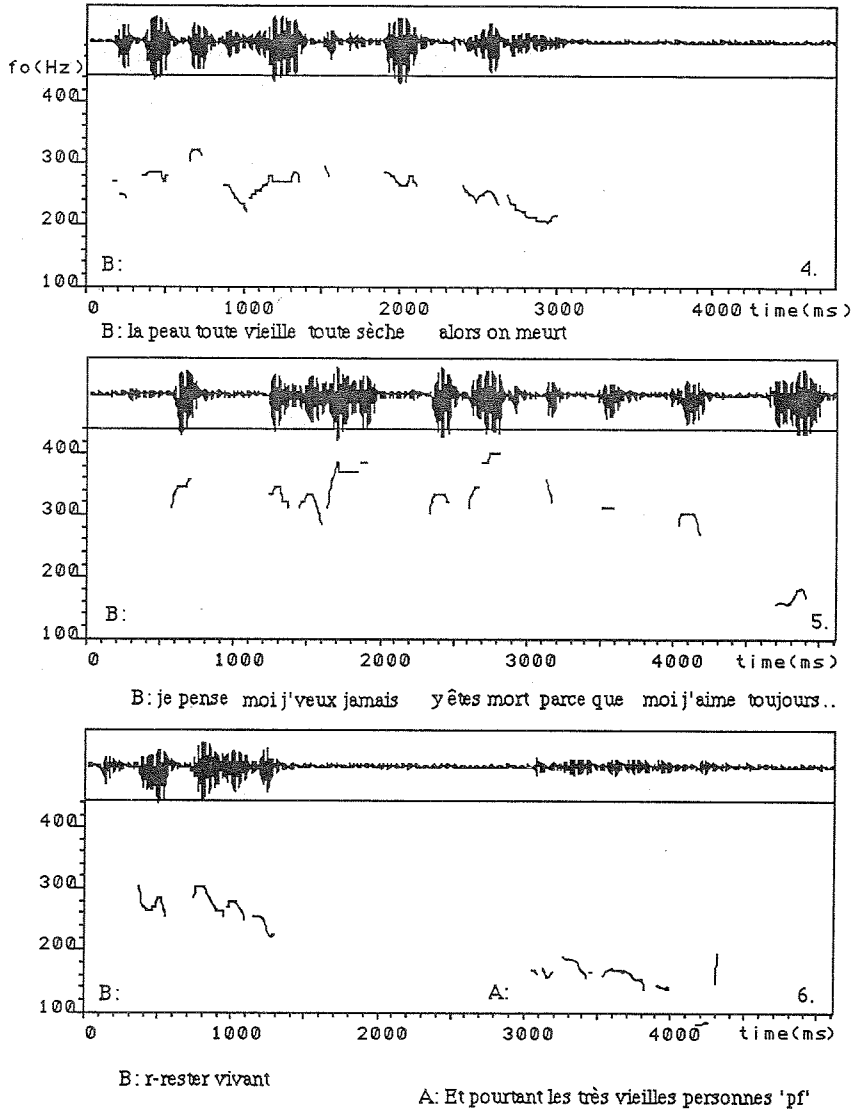
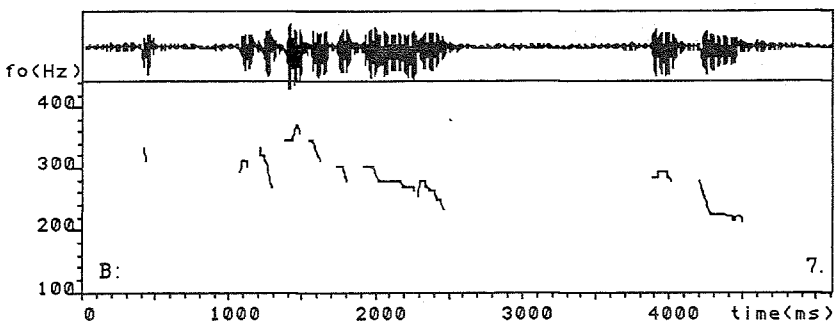


Figure 2 (cont.). A section of a French dialogue (D1:4-6).



B: 'teu' on peut pas s'empêcher de mourir c'est comme ça la vie

Figure 2 (cont.). A section of a French dialogue (D1:7).

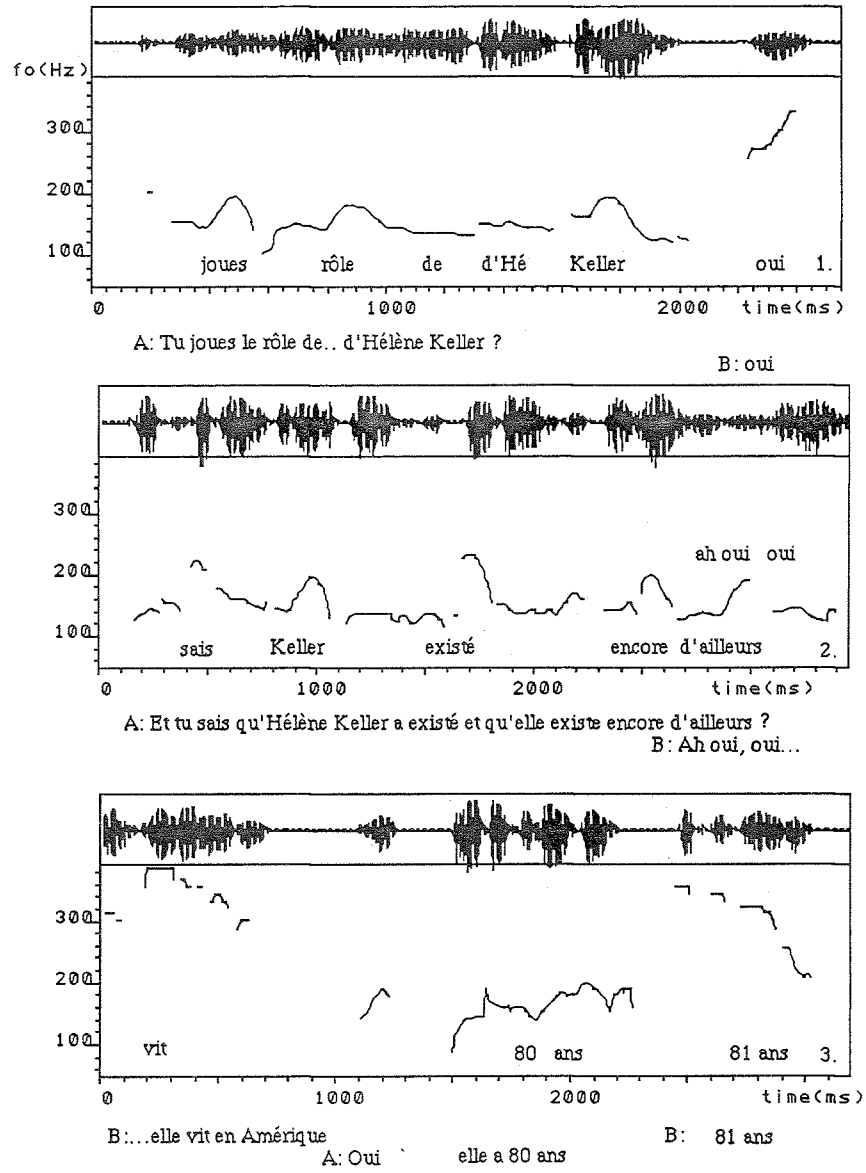


Figure 3. A section of a French dialogue (D2:1-3), waveform (above) pitch contour (centre) and orthographic transcription (below) Keywords are aligned with important pitch events.

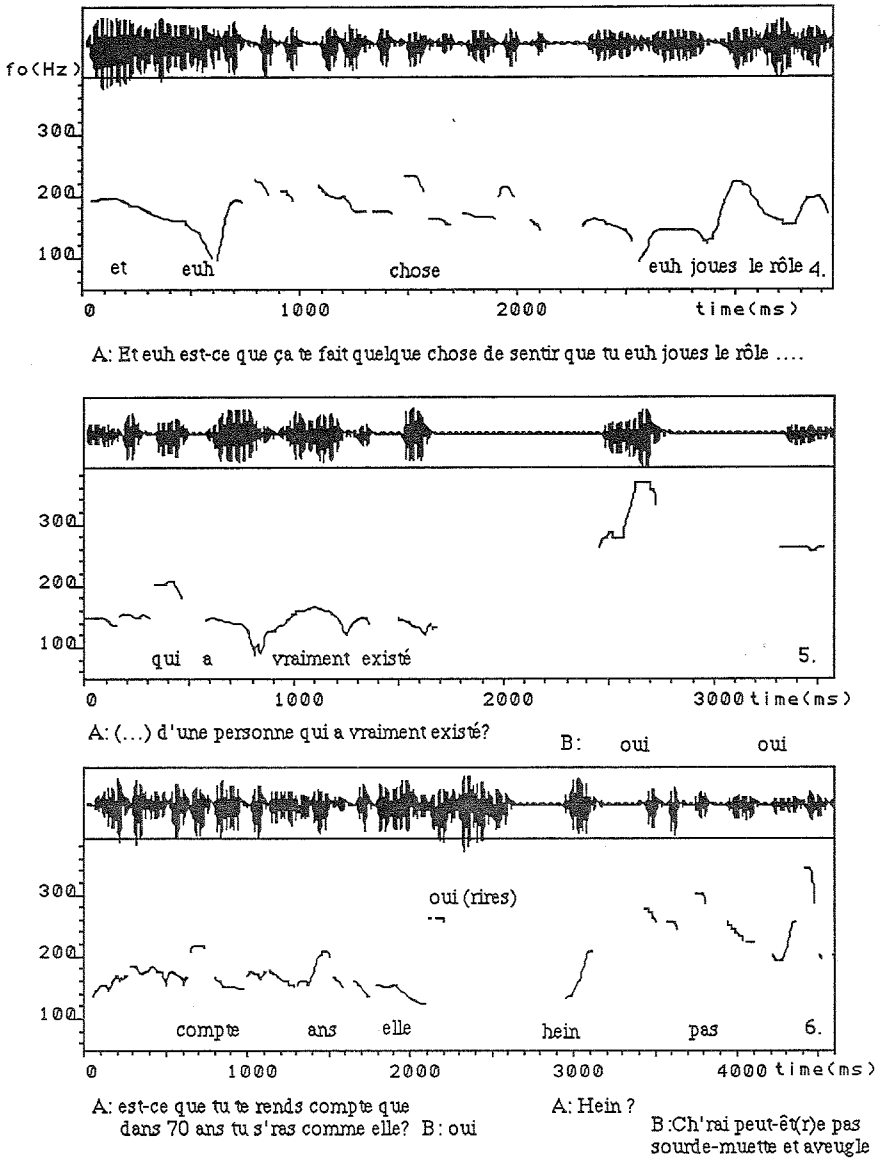


Figure 3 (cont.). A section of a French dialogue (D2:4-6).

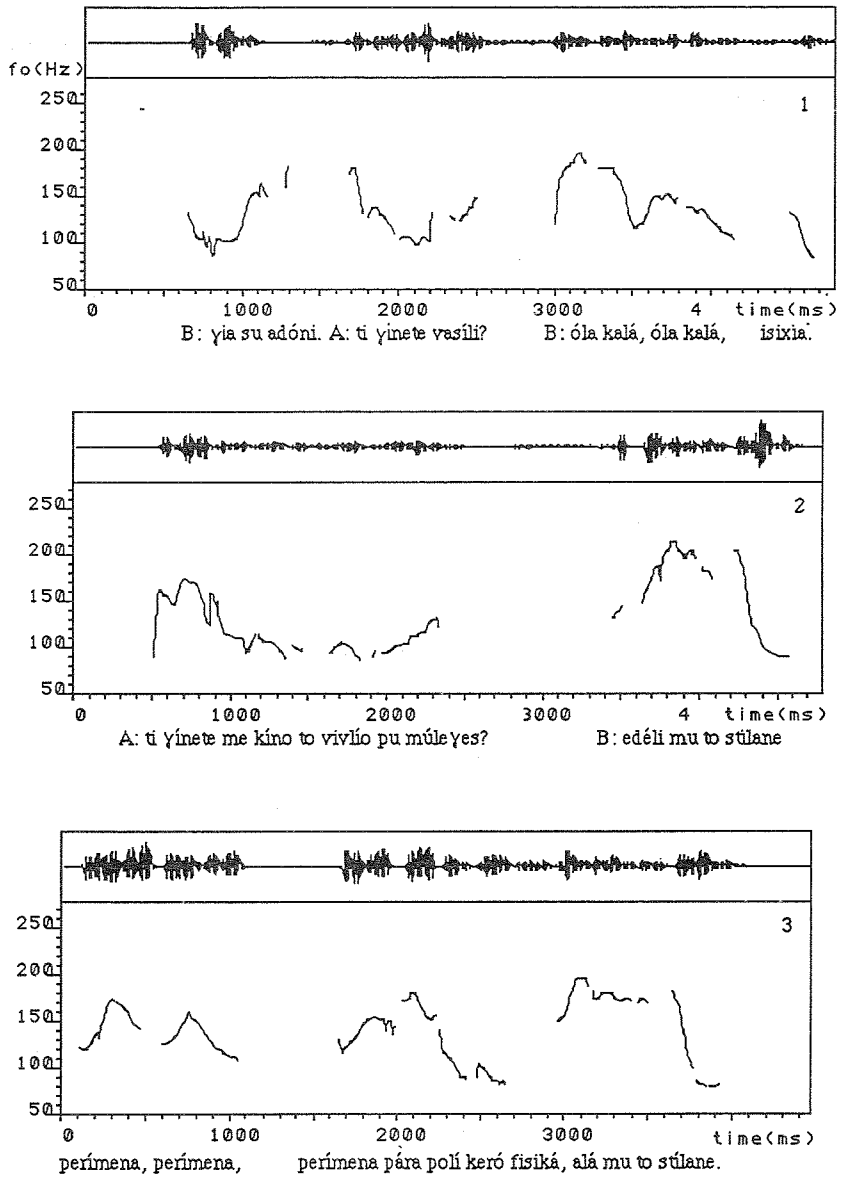
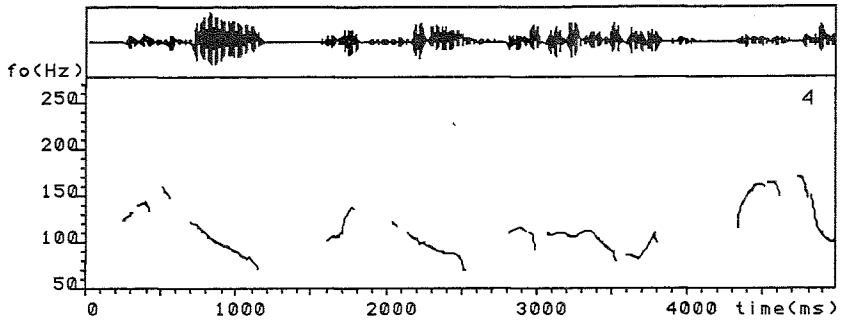
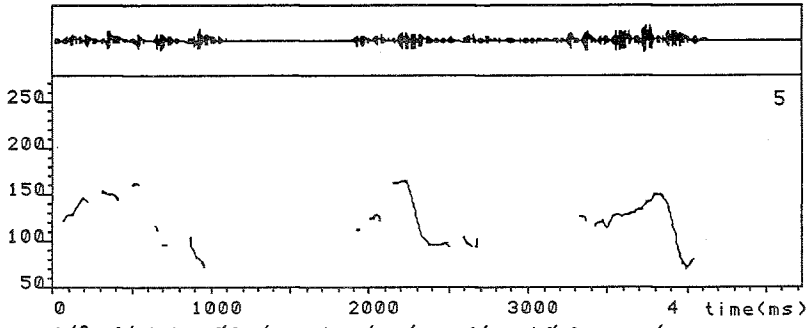


Figure 4:1-3. A section of a Greek dialogue; waveform (above) pitch contour (centre) and phonemic transcription (below) aligned with pitch contour and waveform.





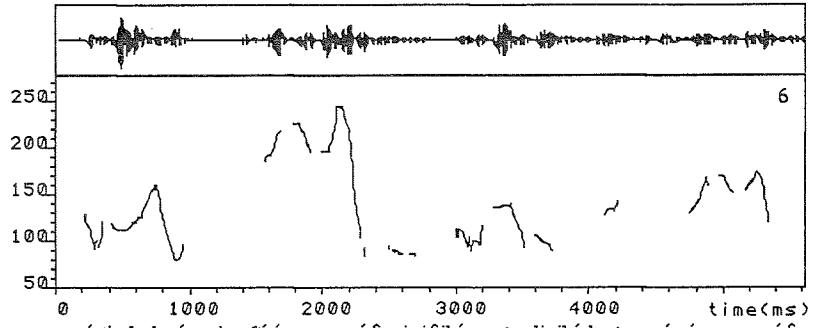
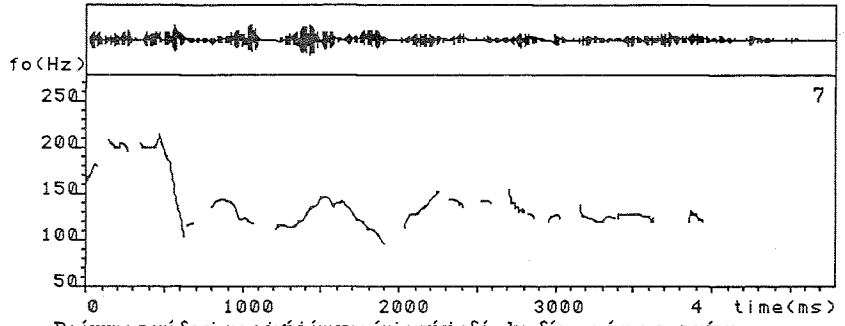
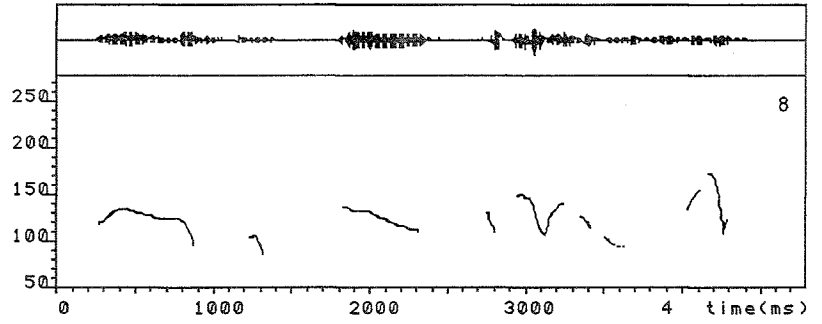


Figure 4:4-6 (cont.)





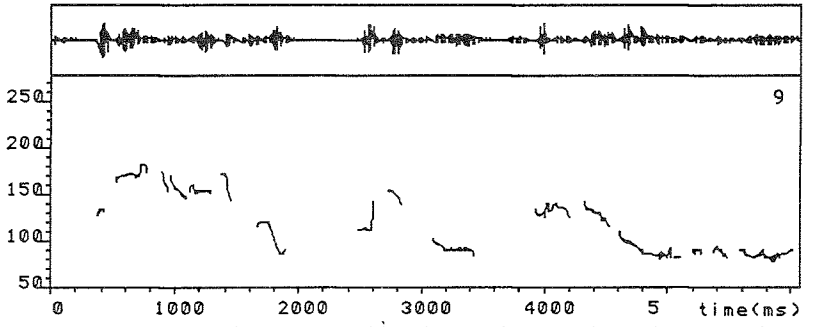


Figure 4:7-9 (cont.)

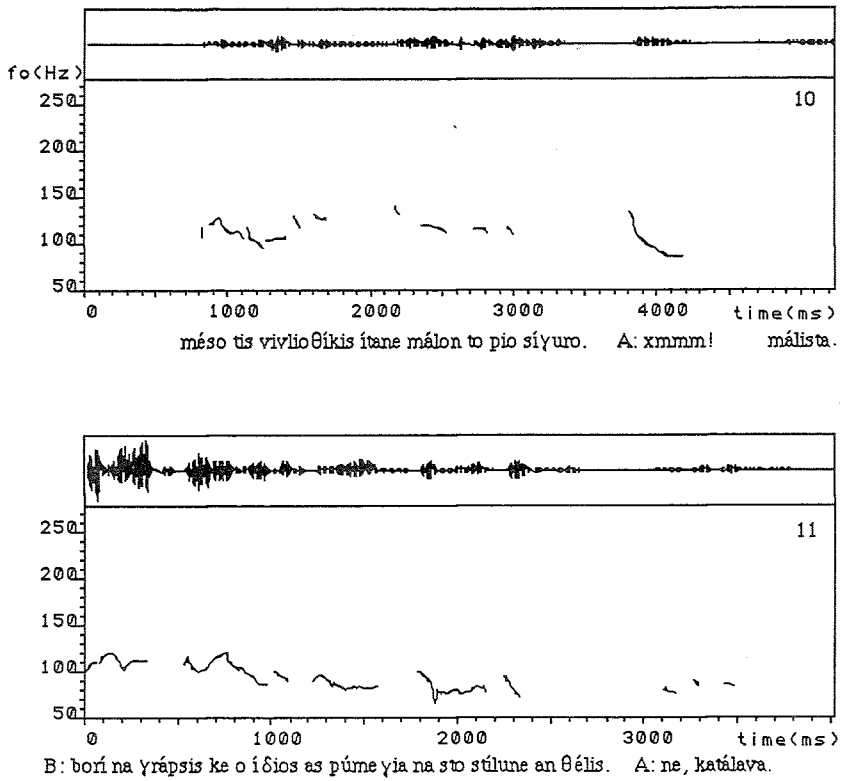


Figure 4:10-11 (cont.)

Lund University, Dept. of Linguistics
Working Papers 33 (1988), 51-64

Some Problems for Machine Translation

Barbara Cairns

Abstract
An analysis of the translation of a Swedish magazine article text into English. Special emphasis is on the type of semantic problems which could arise in machine translation, as envisaged by a human translator. Some syntactic problems of word order are also discussed.

INTRODUCTION

The purpose of this paper is to analyse a text from the point of view of a human translator and to present some of the linguistic obstacles a machine might encounter in carrying out the same translation.

The Swedish text chosen is a magazine article, since this represents a relatively neutral use of language. Since the magazine – Tempus – aims at a general readership, the language is not too specialised or tailored for a specific group of readers. A more literary text would require a totally different kind of analysis and it is now generally acknowledged that the translation of literary texts is beyond the scope of any machine in existence today, as opposed to the naive and unrealistic enthusiasm of earlier machine translation work:

> And now we must come to a question which has long lain in wait for us. Will the machine translate poetry? To this there is only one possible reply – why not? (Delavenay 1960:109)

Admittedly, newspaper language also has its own distinctive style both with regard to syntax and semantics but that does not necessarily present the translator with any extra problems. Journalistic language exists within all languages whose culture includes newspapers and magazines. A machine could presumably be supplied with a specialised programme for the journalistic framework.

The emphasis here will be on problems at the semantic level since it seems that this is where most difficulty would arise for a non-human translator.

THE HUMAN TRANSLATOR

There is a radical difference between the way in which a translator deals with the language of the text he is working on and the way a listener processes the language he hears as part of a spoken dialogue. The translator is already in possession of the "whole" in advance: in other words he has a holistic overview of a