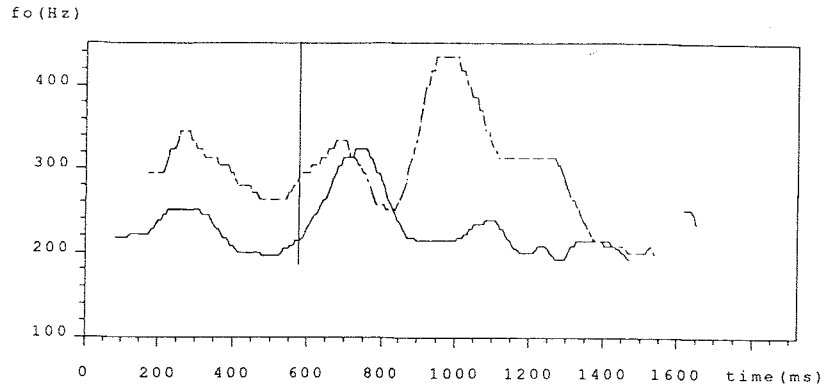


Figure 3. Medially focussed statement and question in child-directed speech. Line-up point in the time domain is the first CV-boundary in the medial word. Statement: solid line; question: broken line.

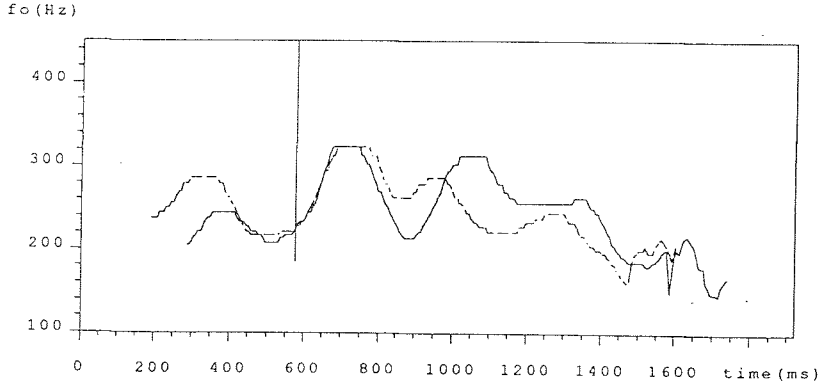


Figure 4. Medially focussed statement and question in adult-directed speech. Line-up point in the time domain is the first CV-boundary in the medial word. Statement: solid line; question: broken line.

Lund University, Dept. of Linguistics
Working Papers 36 (1990), 37-55

On the Analysis of Prosody in Spontaneous Dialogue

Gösta Bruce and Paul Touati

INTRODUCTION

Although by far most phonetic research today (1990) is still done in the sowewhat artificial context of a phonetics laboratory using specially designed, read test material, we are experiencing an increasing trend toward studying spoken language in its natural environment: spontaneous speech and dialogue. It may seem self-evident that the natural environment for studying speech is natural, spontaneous speech and dialogue, but many are the researchers within experimental phonetics who have experienced the overwhelming difficulties of studying real, spontaneous speech. The main difficulty in doing research based on recordings from spontaneous speech is the high degree of variability and thus the low degree of experimental control that can be obtained of whatever feature one may be studying.

One of the pioneers in the phonetic study of prosody in spontaneous speech is Eva Gårding; see for instance her classic study and comparison of Swedish prosody in spontaneous and read speech (Gårding 1967). One lesson learnt by her and taught to us is that studying phonetics, and particularly prosody in spontaneous speech, presupposes studying it in a more rigid, laboratory test setting. The idea is that one has to have fairly specific ideas about what one is going to look for and even what one is likely to find, i.e. one must have a fairly detailed model based on experience from studies of artificial, laboratory speech, in order to be able to extract interesting features of prosody from spontaneous speech. The ideal general methodology would then be some kind of cyclicity between test material and spontaneous speech using feedback from preceding studies.

Our own current interest in prosody and spontaneous speech is at present manifested through a research project called CONTRASTIVE INTERACTIVE PROSODY conducted at the Department of Linguistics and Phonetics in Lund, Sweden. This is a three year project which started in

1988 (see Bruce et al. 1988) and is supported by the Bank of Sweden Tercentenary Foundation. The project is known under the name of KIPROS, an acronym based on the project title in Swedish. The object of study is dialogue prosody in a contrastive perspective in French, Greek and Swedish (two varieties of Swedish: Standard Swedish and Southern Swedish). The ultimate goal of the project is to develop a model for French, Greek and Swedish interactive prosody. For recent reports from the project work see Touati & Bruce 1989, Botinis 1989, Touati 1989, Bruce, Willstedt & Touati 1990, and Bruce & Touati 1990.

Through the KIPROS project we have taken a first step from studying the prosody of more artificial, laboratory speech to examining prosody in natural, spontaneous speech. At the same time, this is also a step from a high degree of experimental control to a much lower degree of control. Sometimes we have felt that this is like going from the swimming pool to the open sea; you may loose control even in the pool, but this is more likely so in the ocean.

Two important general questions that we hope to find an answer to in the project work are the following: 1) Do we find the same, well-known prosodic patterns in spontaneous dialogue as we have met earlier in read, laboratory speech? 2) How are the prosodic patterns observed related to dialogue structure and interactive categories? The first question relates to our "old" research tradition in prosody and the general model of prosody we have been developing in Lund (cf. Bruce 1977, Bruce & Gårding 1978, Gårding 1982, Bruce 1985). Our research on prosody in a spontaneous speech framework will give us an indication of how well we have been able to simulate natural prosody in a laboratory speech environment. The second question is related to the "new" research setting for our study of prosody: spontaneous speech and dialogue. What are the factors that govern the specific choice of prosodic patterns for the speakers involved?

Our research strategy so far in the project work has been to study a fairly restricted sample of speech material in relative depth and from different angles. The choice of dialogue type for our study of prosody has been governed by a number of different criteria which we have discussed in earlier papers (Bruce et al. 1988, Bruce et al. 1990). We have been conducting three different kinds of analysis: 1) analysis of the structure of the dialogue itself without specific reference to prosodic information, 2) auditory analysis in the form of a prosody-oriented transcription, and 3) acoustic-phonetic analysis centered around the examination of pitch.

In the present paper we will concentrate on the two latter aspects of the analysis of dialogue prosody: auditory analysis and acoustic-phonetic analysis. Our focus will be on methodological aspects, while the contrastive analysis will be treated in future papers.

AUDITORY ANALYSIS

One part of the project is concerned with the methodology for the prosodic transcription of speech corpora represented by dialogues. The current study is restricted to dealing with transcription of Swedish (Standard Swedish) and French dialogue prosody. Our prosodic transcription has been discussed in outline in Bruce et al. 1990 and in Bruce & Touati 1990.

We feel that the auditory analysis in terms of a prosodic transcription is a useful categorization of the basic prosodic structure of the utterances of a dialogue. This structure then serves as a basis for the acoustic-phonetic analysis and our search for regularities in prosodic patterning there. Another function for the prosodic transcription is as the input to the text-to-speech synthesis used for simulating dialogue prosody (see the section analysis-by-syntesis).

We try to keep the prosodic transcription distinct from the analysis of dialogue structure and interactive categories. Therefore, our prosodic transcription does not contain categories such as question intonation, continuation tone etc. It is only at a later stage, when we are relating the auditory prosodic analysis - as well as the acoustic-phonetic analysis - to the analysis of the structure of the dialogue itself, that we may establish such potential categories, for example a strong interactive initiative frequently being combined with a characteristic intonation pattern.

The intention of the prosodic transcription system is to provide the three languages involved in the project with a common coding system in order to achieve an auditory analysis of dialogue prosody in the respective languages. Although this system should essentially be considered as a tool for auditory prosodic analysis, it allows us to establish an indirect contrastive analysis. The coding system is thus working as a *tertium comparationis*, as an implicit model.

Before presenting the details of the transcription system, we will first consider some basic principles of transcription in general and specifically prosodic transcription.

One first consideration concerns the dependence of the transcription on the linguistic interpretation. It is clear that the prosodic transcription to be

presented here is based on our own way of conceiving of prosody, i.e. on our particular prosody model. It seems to be true of any transcription system - and particularly true of prosodic transcription - that it must be model-based, presuppose a linguistic interpretation, and consequently has to be based on at least some knowledge of the language in question; cf. Grønnum-Thorsen 1987 for a discussion of prosodic transcription.

Another consideration concerns the abstractness of the transcription. It is true of the prosodic categories of our transcription system that the notation is fairly abstract. Typical examples are stress (and accentual prominence) as well as prosodic grouping (and phrasing). Both categories have a complex phonetic cuing, but the notation is categorical. Thus, although the phonetic correlates of, for example, prosodic grouping involve relations of pitch, intensity, duration, voice quality etc., these are not notated separately but are combined in the perception and notation of a prosodic phrase boundary. The notation of the categories stress and grouping is also abstract in the sense that the actual composition of correlates making up the cuing appears to vary between languages and even dialects. It is thus different for French and Swedish, but has the same symbolization.

A final consideration is related to the phonetics-phonology issue. To take the example of stress and accentual prominence, it is clearly the case that our notation for this is basically phonological. We are transcribing what can be considered as distinctive levels of prominence, and we are consequently categorizing according to these levels, thus disregarding finer degrees of prominence that may be apparently present. A purely phonological notation is, however, not adhered to for all prosodic categories of our transcription system. Below we will try to show for each prosodic feature the level - phonetic or phonological - on which the notation takes place.

In the following, the transcription system will be presented, and we will exemplify how the system can be used to transcribe the prosody of Swedish and French. It should be noted that the prosody-oriented transcription described here is under development and should not be considered in any sense final. Basically it is an orthographic transcription of what has been recorded, although with the actual colloquial forms, repetitions, hesitations etc. represented. To this segmental transcription are added prosodic features selected from our model of prosody. While it does not contain potentially very interesting features such as change in speech tempo, loudness and voice quality, our system does encode five prosodic features: accentual prominence, phrasing, pitch range, boundary tones and pausing.

THE PROSODIC CATEGORIES

Stress and accentual prominence

For Swedish the analysis of prominence levels was made in terms of three binary features: 1) The lowest level of prominence (apart from unstressed), mere stress, coded [,x] as in a so-called secondary stress position of a compound, 2) A higher level of prominence, word accent, coded ['x], and equivalent to the primary stress position of a word, where in Swedish either of the two word accents (accent I or accent II) is manifested, 3) The highest level of prominence at the phrase or utterance level, focal accent, coded ["x] and attached to one or more of the word accents of the actual prosodic phrase. The symbolization of accent I is left unmarked and is implied by the use of the symbols ['x] or ["x], while accent II has the additional symbol [`] above the relevant syllable nucleus. In French only two levels of accentual prominence seem to be relevant (see Mertens 1987 and Touati 1987): a default accent assigned to the final syllable of a prosodic phrase ['x] and an optional focal accent - usually referred to as accent d'insistance - assigned to the initial syllable of a prosodic word or prosodic phrase ["x]. The notation of accentual prominence for both languages is phonological, and the symbolization is in accordance with the new, current IPA system 1989.

Prosodic grouping and phrasing

In the analysis of prosodic phrasing in Swedish, we have so far only recognized one kind of prosodic phrase boundary [‖]. It is not unlikely, however, that we will need two kinds of prosodic group boundary for a more complete treatment of prosodic grouping in Swedish. In their proposal for a phonetic transcription of French sentences, Autessere et al. 1989 decided to limit the number of prosodic categories to three, all three related to boundary: minor boundary, continuative major boundary and conclusive major boundary. In accordance with the IPA recommendations 1989, we shall assume two boundaries for French: minor group boundary [|] and major group boundary [‖]. The notation of prosodic grouping can be considered as basically phonological.

Voice and pitch range

We have devoted special attention to the variation in the general pitch of successive prosodic phrases in a dialogue. What we refer to as the overall pitch range of a prosodic phrase has been called by others more or less synonymously key, (relative) pitch level, or register, to name a few. We believe that the phonetic cuing of pitch range is at least duplex in the normal case. Thus an increase in perceived pitch range probably involves an increase in both F0 range (mainly by raising the F0 peaks) and voice level (intensity) in combination. Although these variables are to be kept apart in theory and can be varied separately, we will assume that they often work together in the practical situation of an ordinary conversation. Furthermore, we assume an auditory normalization for inter-speaker variation depending on speaker size (age, sex, voice register).

A frequently used system for denoting pitch range assumes three different ranges or keys; cf. for example the model of discourse intonation in English proposed by Brazil et al. 1980. We have felt, however, that an analysis using three degrees is not enough for a fairly detailed phonetic notation of pitch range.

Therefore, in Swedish, overall pitch range for a prosodic phrase was analyzed syntagmatically in relation to the neighbouring phrases and may assume five different values: [→] = same, [↗] = slightly raised, [⇑] = markedly raised, [↘] = slightly lowered, [⇓] = markedly lowered. The location of an arrow is at the beginning of each prosodic phrase. Mertens 1987 proposed three registers for French: a middle register, a low register and a high register. The middle register is placed in the central part of the speaker's tonal range; change from this central tonal register (the speaker's usual register) toward a lower tonal register is coded as ⇓, and toward a higher tonal register as ⇑. The coding arrows are inserted at the register's changing points. Here we assume for French, as for Swedish, five different range values.

Our notation of pitch range thus represents a fairly narrow phonetic transcription, as this has been in the focus of our attention. For a broader transcription and one that is basically phonological, we would suggest instead of five different values only two values: raised versus non-raised pitch range. Such a notation seems to be derivable from our present transcription, where [↗] and [⇑] are to be interpreted as raised and the remaining [→], [↘] and [⇓] as non-raised. A notation of pitch range with only two values would probably also capture the most salient pitch breaks in a Swedish and French dialogue. A phonologically oriented notation of pitch range using two values would also be more in agreement with the other prosodic categories that we are transcribing. A two-valued system for key (pitch range) is also part of the model of English prosody presented in Brown et al. 1980. As there is no symbolization of pitch range variation available within the current IPA system, our use of arrows for denoting pitch range - although iconic and transparent - is our own choice.

Boundary tones

Within a prosodic phrase and for a given pitch range, initial and final boundary tones are judged to be either raised (marked value = [˩]) or non-raised (unmarked). This means that the range of, for example, a final pitch rise, notated as a high boundary tone, can vary considerably but still be transcribed as the same category. The notation of boundary tones is thus basically phonological. For Swedish both initial and final boundary tones are transcribed. No specific coding judgment will, however, be made about the pitch realization of final boundary tone in French, as default accent and final boundary tone are prosodic events which co-occur at the end of a prosodic phrase. Our notation of boundary tones does not follow the IPA standard, as there is no specific symbolization of this category available.

Pausing

In our transcription system we think of prosodic grouping (phrasing) and pausing as distinct categories. Boundaries between prosodic phrases can of course be marked without any real pause, although a pause may also be present there and accompany the prosodic grouping. In the transcription of Swedish dialogues, we have assumed that where a real pause is perceived, two degrees of pause length are noted: short [(.)], and long [(..)]. Autessere et al. 1989 proposed for French that silent pause when perceived should also be noted, but they made no difference between short and long pauses. Here we assume for French as for Swedish two degrees of pause length. The symbolization of pausing is in accordance with the IPA recommendations.

Exemplification

The prosodic transcription used here for Swedish is exemplified in (1). The particular Swedish dialogue that we have chosen for our study is a recording from a popular radio program "Ring så spelar vi", which is a radio listener's conversation over the telephone with the program leader.

(1)
B ↗ do dom 'fråga om du "hàde ci'vila 'kläder ||
A → jo dom frå "frågade mej 'de ||
B ↘ å du sv be"jakade 'frågan ||
A → därför att
A → 'jo ja (.) ja va 'nåmligen hade se"mester den 'här 'vèckan 'nu å ↗ || (.)
↘ och 'därför så va ja ci"vìl,klädd på "ròtarysamman,trädet i'går ||
B → 'ja så de 'blev en 'smärre "chock för 'dèl,tagarna ||
A → "ja de 'blev de || (..)
B ↑ "ja 'Bert ||(..) → "så 'e de ||
A → mm ↗ ||
B ↘ nu ska du 'vìnna en 'skìva || (.)
A → nja de e 'inte 'säkert ||

The prosodic transcription exemplified for French represents in (2) a monologue section of a political debate and in (3) a small section of an interview between an adult male reporter and a ten-year-old young actress.

(2)
→ J'es'time | qu'il devrait y avoir un "re'fus | (.) du "cumul des man'dats ||
→et je serais heu'reux | que vous propo'siez | cette mo'tion ||
→une telle "loi à l'Assem'blée Natio'nale ||
↘ mais je 'vois très 'bien | ↓ que vous ne le ferez 'pas ||
↑ je "crois que nous sommes le "seul pays au "monde ||
→où le "maire (.) est égale'ment (.) "dépu'té ||(.)
→où le "maire "partage son "temps |↘ entre "Pa'ris | (.) ↗et sa pro'vince ||

(3)
A →tu 'joues le 'rôle de euh |(..) ↗ d'Hélène Ke"ller ||
B → 'oui
A → et tu "sais qu' Hélène Ke"ller | a exis'té | et qu'elle e'xiste encore d'ai'lleurs ||
B →ah woui (..) woui | woui | ↗ e(lle) vit en Amé'rique ||
A → 'oui (..) | elle a quatre-vingts 'ans ||
B ↗ quatre-vingt un 'ans ||

Simultaneous speech in (1) and (3) is marked by underlines.

ACOUSTIC-PHONETIC ANALYSIS

We consider the auditory analysis in terms of a prosody oriented transcription as illustrated above to be a useful basis and starting point for the acoustic-phonetic analysis of dialogue prosody: the qualitative and quantitative study of prosodic patterns from acoustic recordings of F0 and speech waveform.

The acoustic-phonetic analysis that we have undertaken so far has been centered around pitch. The standard procedure for us has been to have the recorded material digitalized 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. Based on the segmentation of acoustic recordings into relevant prosodic domains from syllable size to larger units through visual inspection and interactive listening, a first part of the acoustic-phonetic analysis consists in isolating relevant pitch patterns for accentuation, phrasing, boundary signalling and pitch range. In the present section we will focus specifically on variation in accentual prominence and changes in overall pitch range.

Accentual prominence

Starting from the abstract phonological notation of accentual prominence as symbols for different prominence levels as part of the prosody oriented transcription, the idea now is to try to relate this representation to an intermediary phonological (or abstract phonetic) pitch representation (Gårding 1981) in terms of pitch patterns expressed as tonal turning points (cf. Gårding 1976 for the concept of turning point). What will concern us here is how the two higher levels of prominence ['x] and ["x] in the transcription can be transformed into another structure of accentual prominence in terms of H(igh) and L(ow) tonal points as well as their synchronization with the segmental structure.

Swedish

According to our model of prosody based on studies of laboratory speech (cf. Bruce 1986) we have characterized the two higher levels of prominence in Swedish - non-focal accent ['x] and focal accent ["x] - in terms of their pitch characteristics as:

non-focal	<	accent I	['x]	=	H L*
		accent II	['x̀]	=	H* L
focal	<	accent I	["x]	=	H L* H
		accent II	["x̀]	=	H* L H

Thus, in our analysis the word accents appear to share the same H L pitch gesture, with a distinctively different timing of the pitch gesture relative to the stress, accent I being timed earlier (H L*) than accent II (H* L). Focal accent is marked by an extra pitch gesture to a H tonal point after the word accent H L, usually in combination with increased duration of the actual

stress group. An illustration of the intermediary phonological analysis of accentual prominence in terms of H's and L's of a small section of a Swedish dialogue is given in Figure 1.

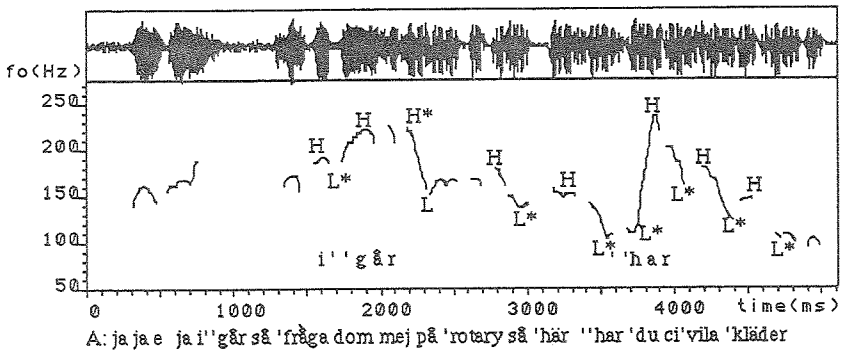


Figure 1:1. Downstepping in a Swedish dialogue, early focus locations; waveform (above), pitch contour (centre) and orthographic transcription with prosodic markers (below). Key word is aligned with important pitch event. See text for explanation of accent symbols (HL) used.

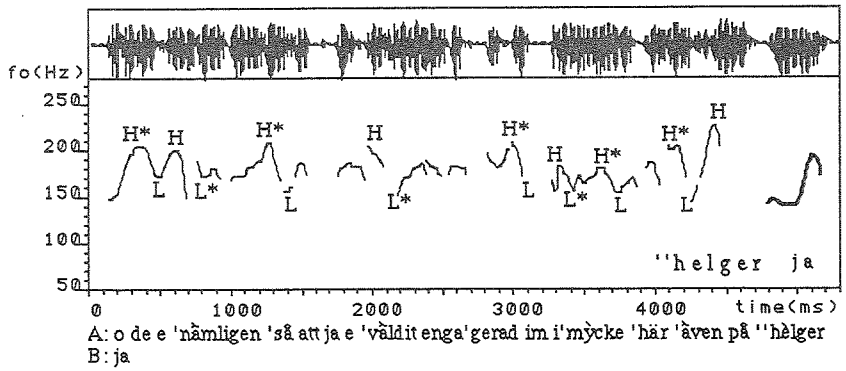


Figure 1:2. Non-downstepping in a Swedish dialogue, late focus location; waveform (above), pitch contour (centre) and orthographic transcription with prosodic markers (below). Key word is aligned with important pitch event. See text for explanation of accent symbols (HL) used.

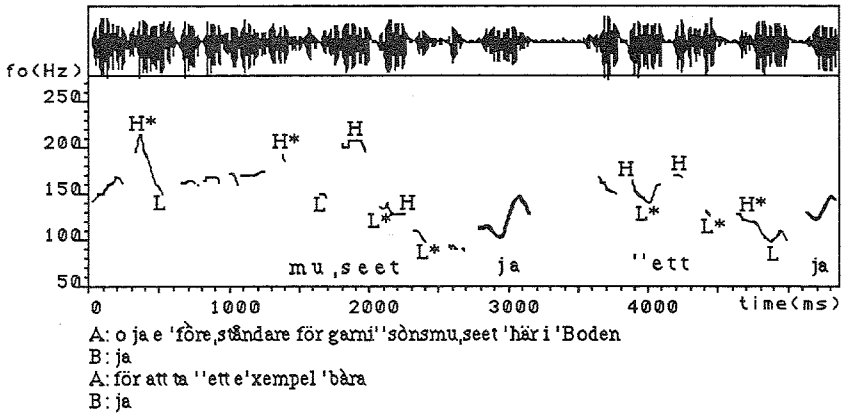


Figure 1:3. Non-downstepping and downstepping in a Swedish dialogue, medial focus locations; waveform (above), pitch contour (centre) and orthographic transcription with prosodic markers (below). Key word is aligned with important pitch event. See text for explanation of accent symbols (HL) used.

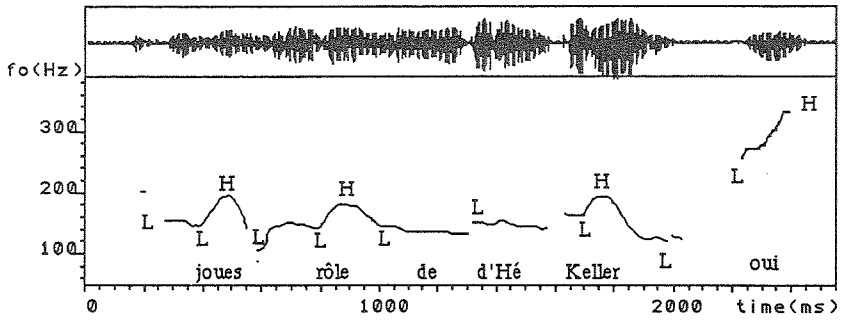
The location of a focal accent in (Standard) Swedish represents a pivot - to use Gårding's term (Gårding 1981) - of a prosodic phrase or utterance. The pivotal character of the focal accent in Swedish can be illustrated by its role in determining the presence or absence of a downstepping pitch contour (cf. Bruce 1982). 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 non-focal accents occur on more or less the same pitch level. However, after a focal accent, the downstepping of successive non-focal accents is a characteristic pitch pattern. This downstepping seems to be the expression of equal prominence of successive post-focal accent within the phrase.

It is interesting to note that in our spontaneous dialogue speech there are several, typical examples 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 describe above. Examples of downstepping after early and medial locations of focal accent are found in Figures 1:1 and 1:3, and examples of non-downstepping before late or medial locations of focal accent in Figures 1:2 and 1:3.

French

The main relevant categories of accentual prominence in French are phrase accent (*accent de groupe*) ['x] and focal accent (*accent d'insistance*) ["x]. 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 LH in utterance non-final groups and HL in utterance final groups and for focal accent LH. 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 lengthening of a syllable that is assigned focal accent (*accent d'insistance*).

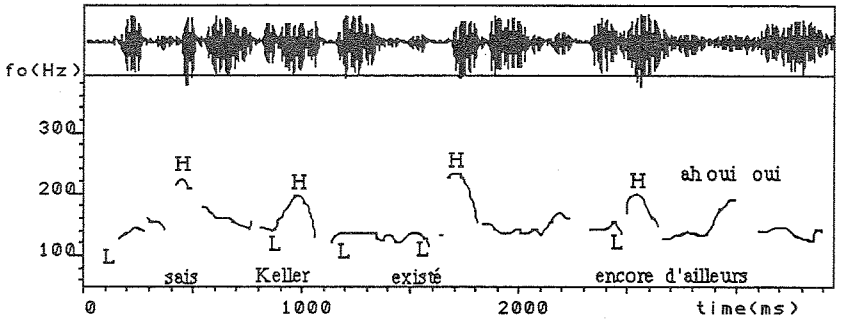


A: tu 'joues le 'rôle de euh |(..)| d´Hélène Ke"ller ||
B:'oui ||

Figure 2. Phrase accent, initial juncture and pivot in a French interview; waveform (above), pitch contour (centre) and orthographic transcription with prosodic markers (below). Key word is aligned with important pitch event. See text for explanation of accent symbols (HL) used.

As seen in Figures 2 and 3, group initial juncture is often manifested by a pitch minimum (L). The combination between phrase accent in utterance non-final groups and initial juncture produces co-lateral prosodic events, 'pivots' with a LHL representation (cf.'*joues le*','*rôle de*' in Fig. 2). Hesitation implies in some cases a tonal resetting of the initial juncture of the following prosodic group (cf.'*d'Hé*' in Figure 2).

It is also interesting to note that there are different types of interpolation between group initial juncture and the next accent in the form of rising, falling and level pitch patterns. Examples of rising pitch interpolation as in '*et tu sais*', falling pitch interpolation as in '*qu'Hélène Keller*' and level as in '*a existé*' are found in Figure 3.



A: et tu sais qu´Hélène Ke"ller | a exis'té | et qu'elle e'xiste encore d´ailleurs ||

B: ah woui (..)woui woui |

Figure 3. Rising, falling and level pitch interpolations in a French interview; waveform (above), pitch contour (centre) and orthographic transcription with prosodic markers (below). Key word is aligned with important pitch event. See text for explanation of accent symbols (HL) used.

Pitch range

In our auditory analysis of voice and pitch range as part of the prosodic transcription we made three basic assumptions: 1) that the phonetic cuing of overall pitch range often involves both variation in F0 and intensity in combination, 2) that pitch range is a syntagmatic relation among successive prosodic phrases, and 3) that pitch range may assume five different values in relation to the preceding prosodic phrase: same, slightly and markedly raised, and slightly and markedly lowered.

In our present acoustic-phonetic treatment of pitch range we will only consider variation in F0 and thus disregard intensity variation. Although in our auditory analysis of pitch range we have found it convenient to work with five different values for the notation of pitch range, we hypothesize

that pitch range is in fact regulated more or less continuously. Thus we do not believe that there is a fixed repertoire of pitch ranges forming a paradigmatic class from which the speaker can choose. Instead the pure syntagmatic regulation of pitch range, the increase or decrease of a suitable size by the speaker in relation to the preceding phrase depending on the estimated needs of the situation, seems to us more likely.

Another question concerns how the actual variation in F0 range is accomplished. In a study of variation in F0 range as an expression of speaker involvement in read, laboratory speech (Bruce 1982), it was shown that the 'floor' of the speaker's voice range is basically constant, while the peak values of the F0 contour are highly variable with speaker involvement (involved versus detached). That F0 peaks vary, while valleys do so much less, with variation in pitch range thus constitutes our expectation also for spontaneous dialogue.

Furthermore, variation in overall pitch range (for the domain of a phrase) can be assumed to be exploited for interactive purposes. Differing degrees of attention generally seem to correlate with variation in range. Another more specific hypothesis has been to ascribe variation in pitch range a possible connection with boundaries in the dialogue structure, for example to speech paragraphs or to the introduction of a new conversation topic (cf. Brown et al. 1980, Hirschberg and Pierrehumbert 1986).

Swedish

For Swedish we have observed the exploitation of variation in pitch range, specifically a more or less successive decrease in range, for the expression of textual coherence within a larger speech paragraph consisting of several successive prosodic phrases and a shift in pitch range (usually a widening in range) for the marking of boundaries in the dialogue structure. An apparent example of this use of variation in pitch range - a successive decrease in range during a speech paragraph produced in interaction between the two speakers and then an extra wide range for rounding off the paragraph followed by a new decrease in range for the new interactive paragraph - in a section of the Swedish dialogue transcribed above (1) is found in the consecutive Figures 4 and 5.

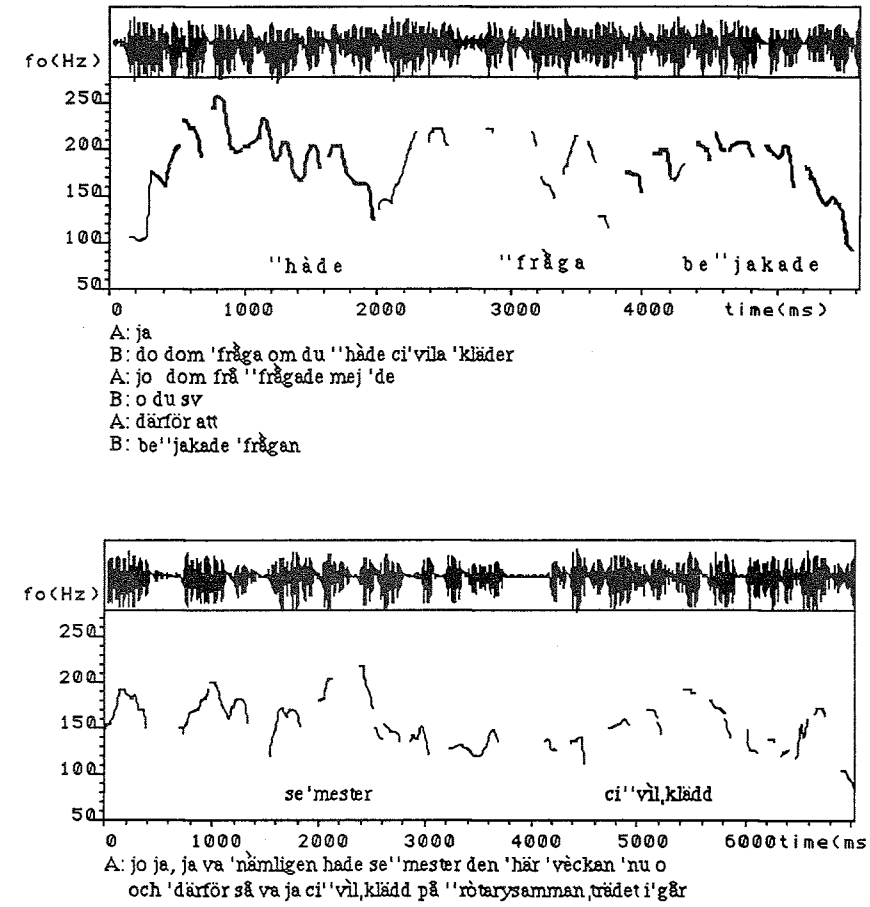


Figure 4. Range variation: successive decrease in range in a Swedish dialogue; waveform (above), pitch contour (centre) and orthographic transcription with prosodic markers (below). Key words are aligned with important pitch events.

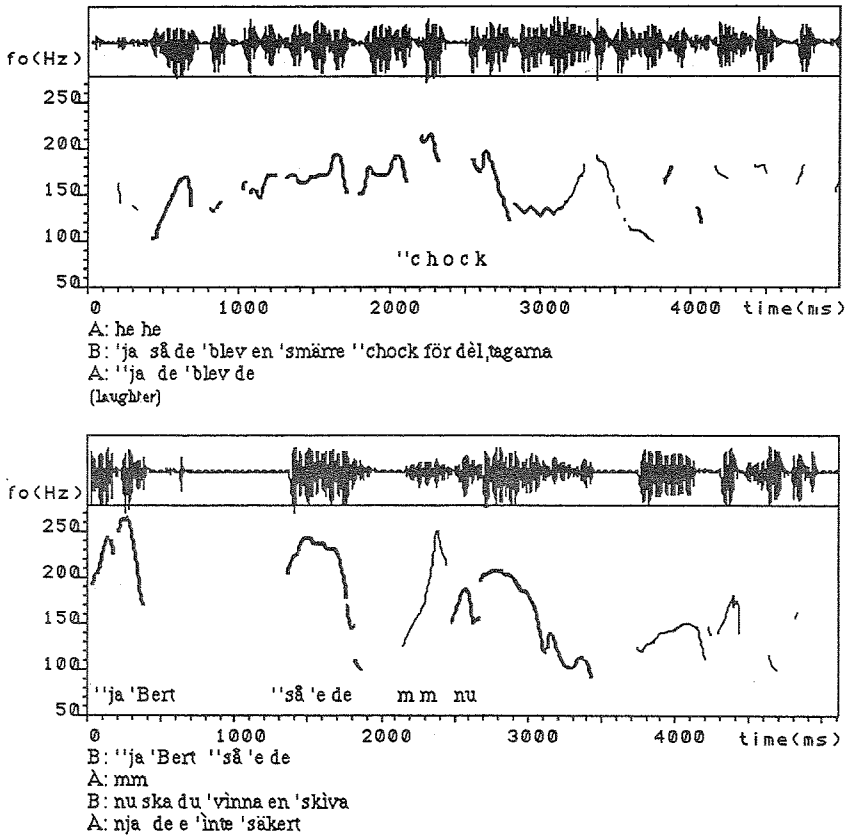


Figure 5. Range variation: extra wide range associated with rounding off a topic in a Swedish dialogue; waveform (above), pitch contour (centre) and orthographic transcription with prosodic markers (below). Key words are aligned with important pitch events.

French

Variation in pitch range can apparently also be used by a speaker in French in order to highlight or downplay consecutive prosodic phrases in a speech paragraph. To illustrate this point, acoustic displays of a monologue section of the French dialogue (2) are presented in Figure 6.

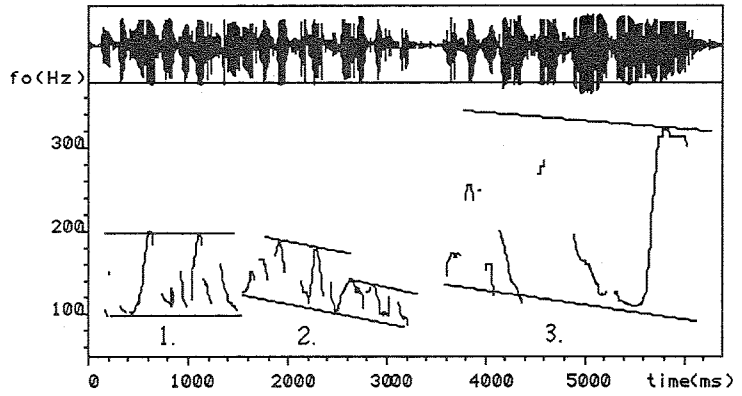


Figure 6. Range variation associated with highlight versus downplay of consecutive prosodic phrases in a turn of a French political debate; waveform (above), pitch contour (centre) and orthographic transcription with pitch range markers (below)

However, our experience from studying dialogue prosody so far indicates that the connection between variation in pitch range and sectioning of a dialogue is probably best regarded as a tendency rather than as an obligatory feature.

Analysis-by-synthesis

An important and powerful method in our modelling of dialogue prosody and particularly the exploitation of pitch range is analysis-by-synthesis. The research tool which we have been using so far only for Swedish is the multilingual text-to-speech system developed by Carlson and Granström 1986. The prosody rules of the Swedish text-to-speech system have recently been modified by Bruce & Granström 1989, 1990. In our use of rule synthesis, the starting point is a phonetic transcription of prosodic features, basically the same features as described above in the section Prosodic Categories. Our preliminary testing of dialogue synthesis clearly shows that variation in overall pitch range may be considered a potentially important means for use in the development of a dialogue and its division into speech paragraphs (see further Bruce et al. 1990).

REFERENCES

Autesserre, D., Pérennou, G. & Rossi, M. 1989. 'Methodology for the transcription and labeling of a speech corpus'. *Journal of the International Phonetic Association 19 (1)*, 2-15.

Botinis, A. 1989. 'Discourse intonation in Greek'. *Working Papers 35*, 5-25. Lund: Dept. of Linguistics.

Brazil, D., Coulthard, M. and Johns, C. 1980. *Discourse intonation and language teaching*. London: Longman.

Brown, G., Currie, K. and Kenworthy, J. 1980. *Questions of intonation*. London: Croom Helm.

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

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

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

Bruce, G. 1986.'How floating is focal accent?'. *Nordic Prosody IV*, eds.K. Gregersen & H. Basbøll, 41-49. Odense University Press.

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

Bruce, G. and Granström, B. 1989. 'Modelling Swedish intonation in a text-to-speech system'. *STL-QPSR 1*, 17-21, Stockholm:KTH, Speech Transmission Laboratory.

Bruce, G. and Granström, B. 1990. 'Modelling Swedish prosody in text-to-speech: phrasing'. To appear in *Nordic Prosody V*, Åbo University: Finland.

Bruce, G. and Touati, P. 1990. 'Auditory and acoustic analysis of dialogue prosody in Swedish and French'. *Phonum 1*, 27-30. Department of Phonetics, Umeå University.

Bruce, G., Willstedt, U., Touati, P. and Botinis, A. 1988. 'Dialogue prosody'. *Working Papers 34*, 21-24. Lund: Dept. of Linguistics.

Bruce, G., Willstedt, U. & Touati, P. 1990. 'On Swedish interactive prosody: analysis and synthesis'. To appear in *Nordic Prosody V* , Åbo University: Finland.

Carlsson, R and Granström, B. 1986. 'Linguistic processing in the KTH multilingual text-to-speech system'. *In Proc. ICASSP 86, Vol.4*, 2403-2406. Tokyo.

Grønnum-Thorsen, N. 1987.'Suprasegmental transcription'. *Aripuc 21*, 1-27. Copenhagen.

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

Gårding, E. 1976. 'The importance of turning points for the pitch patterns of Swedish accents'. *SCOPIL 4* , ed. L. M. Hyman, 27-35. Los Angeles.

Gårding, E. 1981.'Contrastive prosody: a model and its application'. *Studia Linguistica 35*, 146-166.

Gårding, E. 1982. 'Swedish prosody'. *Phonetica 39*, 288-301.

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

I.P.A. 1989. 'Report on the 1989 Kiel Convention'. *Journal of the International Phonetic Association 19 (2)*, 67-80.

Mertens, P. 1987. *L'intonation du français. De la description linguistique à la reconnaissance automatique*. Katholieke Universiteit Leuven.

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

Touati, P. 1989. 'De la prosodie française du dialogue. Rapport du projet KIPROS'. *Working Papers 35*, 203-214. Lund: Dept. of Linguistics.

Touati, P. and Bruce, G. 1989. 'Report from the KIPROS project'. *STL-QPSR 1*, 51-54, Stockholm: KTH, Speech Transmission Laboratory.