Report from the KIPROS project

Touati, P. and Bruce, G.

journal: STL-QPSR
volume: 30
number: 1
year: 1989
pages: 051-054

http://www.speech.kth.se/qpsr
REPORT FROM THE KIPROS PROJECT

Paul Touati and Gösta Bruce
Dept. of Linguistics and Phonetics, Lund University

Abstract

The purpose of the KIPROS project is to investigate dialogue prosody in French, Greek, and Swedish. One goal of our research is the development of a model for prosody in an interactive perspective. In the present paper, the so-called Lund model of prosody is applied to the analysis of a short French dialogue.

INTRODUCTION

The research reported on in the present paper was done within the KIPROS project. The full title of this project is Contrastive Interactive Prosody with the acronym KIPROS based on the project title in Swedish. The project is supported by the Bank of Sweden Tercentenary Foundation. 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). The ultimate goal of our research will be the development of a model for French, Greek, and Swedish prosody having interactive perspective.

The present paper is centered around French and gives a short report of the second research step in the French part of the project. After a preliminary analysis of different types of dialogue as a first approach to the prosodic structure of spontaneous speech in French (see Bruce, Touati, Botinis, & Willstedt, 1988), the second step aims at integrating the dialogue analysis in a model perspective. We believe that a model like the so-called Lund model of prosody (cf. Bruce, 1977; Bruce & Gårding, 1978; Gårding & Bruce, 1981; Gårding, 1982; Touati, 1987) which has already provided method and tools for the prosodic analysis of read sentences, in so-called laboratory speech, should also serve as useful reference framework for the analysis of prosody within a larger and different domain, i.e., interactive prosody and dialogues.

OUTLINE OF THE MODEL

The main components of the model are presented in Fig. 1a (cf. Touati, 1987). The model is operating as an input-output system. The input is here a dialogue marked with labels corresponding to the main prosodic categories and information about grammatical boundaries (Fig. 1a). The prosodic categories and the labels used in the analysis of the dialogue are: for the phrase accent "Accent Démarcatif or AD", for the accent d'insistance or focal accent "Accent Focal or AF" and for the initial juncture "Jointure Initiale or JI". The output is a string of syllables structured as a well-formed temporal profile and a tonal configuration.

The syllable structure rules component (SSR) (cf. Fig. 1a) is designed to account for the basic process of syllabification; words within phrases all along the dialogue are syllabified by applying rules like liaison, elision; this stage is an important one because it is achieving syllabification between the boundaries for the functional domain of French prosody, i.e., between a phrase composed of at least one lexical word plus possible preceding grammatical words (marked with "\(\neq\)"). The syllable duration rules component
(SDR) accounts for the temporal structure of prosody. The pitch rule component (PR) is concerned with generation of pitch configurations. The intermediary pitch representations of prosodic categories are Low and High turning-points (AD=LH in utterance non-final groups and HL in an utterance final group, AF=LH and JI=L). The turning-points are then transformed into a concrete string of Fo minima and maxima by the algorithm.

**INPUT**

```
JI AD JI AD JI AD JI AD AD
[# tu joue le rôle d'Hélène Keller #] [oui #]
```

**OUTPUT**

TEMPORAL PROFILE  
TONAL CONFIGURATION

*Fig. 1a. The main components of the model.*

**SPEECH MATERIAL AND ANALYSIS**

For the analysis performed here, the recording was taken from a broadcast interview (see the D2 dialogue in Bruce & al, 1988, pp. 31-32 and 45-46). The recorded material has been digitized on the VAX 11/730 of the Lund Phonetics laboratory and analyzed through the API program of the ILS package, where pitch extraction is done with a modified cepstral technique.

Following the first stages of the model, syllabic segmentation and label attribution of the dialogue were made on computer-generated oscillogram, intensity and Fo traces with auditive feed-back using a program specially designed for this purpose by Lars Eriksson. The analysis of temporal profiles and tonal configurations was derived from measurements obtained from the labelling and segmentation processes.

**PRELIMINARY OBSERVATIONS**

The whole dialogue has being analysed within the model framework. In this report our comments are restricted to the two first lines of the analysed part of the dialogue as an illustration of the method. The input representation is presented in Fig. 1a. Figs. 1b and 1c show the concrete temporal and tonal realisations of these two lines.

Now listening to the lines and looking at the output we can make a first observation: the first line, the journalist’s question, is broken down into several minor prosodic groups by hesitations. It is clear that in contrast to reading style where the first line would have been pronounced as a whole group, the temporal profile and the tonal con-
Figuration of the line in the dialogue is typically patterned by hesitations. The second observation is a very trivial one: there is a pause between the two speakers' lines.

Fig. 1b. The concrete temporal profile for two lines of the dialogue ("tu joues le rôle d'Hélène Keller?" - "oui").

Fig. 1b illustrates the temporal profile for the two lines. It shows that the hesitations have variable temporal values. As a result the line is rhythmically patterned into two major groups with a major temporal boundary between "rôôle de" and "d'Hélène"; in its turn, the first major group is divided into two minor ones: "tu joues" and "le rôle de". Looking at the AD labels in the temporal profile (Fig. 1b), it seems that like in reading style the longest syllable is located at the end of the sentence in a preboundary position. Apparently the AD labels on "joues" and on "lène" do not result in a noticeable lengthening as compared with the AD on "ller" and in some extent on "rôôle". Maybe in this last case the lengthening is due to the following hesitation. It would therefore be interesting to test if there is any temporal correlation between an AD and a following hesitation.

Fig. 1c. The concrete tonal configuration for the two lines of the dialogue ("tu joues le rôle d'Hélène Keller?" - "oui").

Fig. 1c illustrates the tonal configuration for the two lines. Even if the journalist's line is semantically a question, it shows a global pitch configuration consisting of a downdrifting Fo as expected in a declarative. It seems as if the modality of the question is locally carried by an LHL configuration on the AD in the "ller" syllable of Keller. The impression given by the tonal configuration (rise-fall) is that it is not a question asking...
for information but rather asking for confirmation ("tu joues le rôle d'Hélène Keller, n'est-ce pas?"). The major hesitation also divides the tonal configuration into two major tonal phrases causing an Fo resetting at the beginning of the second phrase. All the other AD labels are correlated with a rising Fo during the stressed vowel except the one on "lène" which is deleted as part of the prosodic parsing of "d'Hélène Keller". The first AD on "joues" also seems to correspond to a focal accent; it has a strong and rapid Fo rising. It may be argued that this feature which has been found to be typical of a word characterized by accent d'insistance (cf. Touati, 1987) may potentially also attributed to the corresponding syllable of a phrase.

The answer by the other speaker (the child) -the "oui"- would by default expectation be a one word declarative sentence marked with JI/AD labels. The model would predict for such a sentence a LHL configuration as a result of the L for the JI and of the HL for the AD. The data shows instead a rising Fo (LH). Another interpretation would be to say that the "oui" as a back channel item is instead to be regarded as a discourse unit internal item, presupposing a continuation, normally by the other speaker. Probably a "oui" with a falling Fo should have sounded too conclusive.

The SDR component of the model has to be developed in the light of the dialogue studies in order to relate for example the distribution and duration of hesitations and pauses to some dialogue categories. Concerning the PR component, an interesting observation is the manifestation and the domain for focal accentuation. It seems that the domain for focal accent is larger than earlier assumed.

References
Bruce, G. & Gårding, E. (1978): "A prosodic typology for Swedish dialects," pp. 219 - 228 in (Gårding et al., eds.) Nordic Prosody (Dept. of Linguistics, Lund University.)