

DISTINCTION BETWEEN 'NORMAL' AND 'CONTRASTIVE/EMPHATIC' FOCUS

Anja (Petzold) Elsner
e-mail: ape@ikp.uni-bonn.de

Institut für Kommunikationsforschung und Phonetik (IKP), University of Bonn,
Poppelsdorfer Allee 47, 53115 Bonn, Germany

ABSTRACT

A method to extract and classify focus accents has been developed. It works for German spontaneous speech. The method tries to distinguish 'normal' and 'contrastive/emphatic' focus accents using phrase boundaries. It was found that contrastive/emphatic accents tend to have greater distances to phrase boundaries than normal focus accents. Moreover, for contrastive/emphatic accents there was found a much steeper F_0 rise for accents with a rather high distance from the next phrase boundary.

1. INTRODUCTION

In the German research project VERBMOBIL several modules work together in recognizing speech from spontaneous dialogues. Further processing includes translation and synthesis of these dialogues in this application. It is important for translation and synthesis to have additional prosodic information which otherwise would have to be derived from the linguistic context. In many cases it is even impossible to recognize special emphasis intended by speakers in spoken dialogues without prosody.

Let us consider the following examples (focussed parts emphasized):

(1a) *In **der** Woche kann ich nicht.*

(In this week it's impossible for me. [but perhaps in another week])

(1b) *In der **Woche** kann ich nicht.*

(During the week it's impossible for me. [but perhaps on the weekend])

In this case linguistic analysis has no chance to find the correct meaning without prosodic information. Nevertheless it is likely that a speech processing system takes the second version (1b) as standard interpretation because 'Woche' (week) is a content word. The other reason is that we normally would expect a phrase boundary after 'Woche' - we will come back to this later (see Section 3).

(2a) *Ich könnte um elf noch einen **Termin** reinschieben*

(I could insert another **date** at eleven o'clock.)

(2b) *Ich könnte um **elf** noch einen Termin reinschieben*

(I could insert another date at **eleven** o'clock.)

At first sight we have no meaning difference in this example. However, in (2b) we could suppose that the *only* possible time for the speaker would be eleven o'clock, while this question is left open in (2a). Thus, for translation or synthesis of this sentence we need prosodic information to reproduce the original intention of the speaker even if it implies only a small shift in meaning.

2. FOCUS RECOGNITION

Starting point for this investigation is an already existing algorithm for focus recognition. Focus is defined here as the semantically most important part of an utterance, which is in general marked by prosodic means. The focus accents reflect the intention of the speaker to mark those parts of a sentence which he feels to be important. Normally these are content words. Nevertheless, in special contrastive/emphatic aspects it is also possible to put a focus accent on a function word (see example (1a)).

Our method for focus recognition is as follows [1]: The algorithm tries to solve focus recognition by global description of the utterance contour, in a first approach represented by the fundamental frequency F_0 . Investigations of Swedish spontaneous speech [2] have shown that declination can be controlled by the focal accent: It was found that in pre-focal position there is no downstepping, but after a focal accent downstepping is significant and characteristic.

To examine this feature in German spontaneous speech, the reference line was computed as follows: First the F_0 contour was postprocessed by a special smoothing algorithm described in [3]. (Without smoothing results get worse by about 7 %.)