AN ALGORITHM FOR RECOGNITION OF STRESS IN DANISH
AND ITS APPLICATION IN AN ASR SYSTEM

Michael Bundgaard

Institute of General and Applied Linguistics, University of Copenhagen

ABSTRACT

A three step algorithm for automatic detection of stressed syllables in Danish and its application in an Automatic Speech Recognition system is described in this paper.

The ASR system HEAD is an acoustic-phonetic expert system, incorporating a neural network which combined with a Greatest Likelihood Relation (GLR) segmentation-algorithm propose allophone hypothesis. This string of allophones is then used by the lexicon to produce word hypothesis. Finally the words are joined to form syntactically legal sentences.

The first step in the stress algorithm is the detection of the syllable nuclei, primarily based on band-passed filtered energy corresponding to sonorant energy, and secondarily on pitch movements.

In the second step the stress-marking pitch movements are detected, based on the model for Danish stress-group patterns proposed by N Thorsen (1). Here the pitch is the primary cue to stress, and all stresses are marked by a rise in the pitch.

The third step concerns the assigning of the feature stress from each of the found stress-marking pitch movements to a nearby located syllable nucleus. The stressed syllable is located close to the increase in the pitch curve, although the exact position is not predictable due to dialectical variations. Still the primary cue is the pitch signaling that in this area there is a stressed syllable. If the postioning of the syllable nuclei gives reason to any doubt, the secondary stress marking parameters are taken into account, ie duration and intensity. The durational aspects are derived from the

GLR-algorithm, whereas the filtered energy from the syllable detection normalized for intrinsic vowel intensity.

The prosodic feature stress is assigned to the vocalic hypothesis as well as the feature syllabic. Once a vowel has been categorised as stressed, it will play a central role in the HEAD recognition system - the 'islands of clarity' idea. The hypothesis containing the stress feature will have higher weight, and thus normally the vowels are the first to trigger off a lexicon search. This very central role attributed to the stress implicates that a false alarm is not tolerated. In contrary it is not essential if one stressed syllable is missed, because still the system allows a match of those words in lexicon.

Besides this strategic role the prosodic features stress is used as a constraint on lexical access and in finding word boundaries.

(1)
Thorsen, Nina: Standard Danish Intonation
ARIPUC vol 22, 1988