

p. 2894

2aSC27. Inclusion of a prosodic module in spoken language translation systems. Robert Eklund and Bertil Lyberg (Telia Res. AB, Systems Res., Spoken Language Processing, S-136 80 Haninge, Sweden)

Current speech recognition systems mainly work on statistical bases and make no use of information signaled by prosody, i.e. the segment duration and fundamental frequency contour of the speech signal. In more advanced applications for speech recognition, such as speech-to-speech translation systems, it is necessary to include the linguistic information conveyed by prosody. Earlier research has shown that prosody conveys information at syntactic, semantic, and pragmatic levels. The degree of linguistic information conveyed by prosody varies between languages, from languages such as English, with a relatively low degree of prosodic

p. 2895

disambiguation, via tone-accent languages such as Swedish, to pure-tone languages. The inclusion of a prosodic module in speech translation systems is not only vital in order to link the source language to the target language, but could also be used to enhance speech recognition proper. Besides syntactic and semantic information, properties such as dialect, sociolect, sex, and attitude, etc. is signaled by prosody. Speech-to-speech recognition systems that will not transfer this type of information will be of limited value for person-to-person communication. A tentative architecture for the inclusion of a prosodic module in a speech-to-speech translation system is presented.