Intonation in Spoken Language Generation

Spoken Language Generation (SLG) is a combination of Natural Language Generation (NLG) and Speech Synthesis. The central question is how one can program computers to produce high-quality speech from some computer-internal representation of information (rather than from text, as in Text-to-Speech synthesis). The research on SLG described in this thesis focuses on generating Dutch speech with an optimal intonation. The motivation is primarily linguistic in the sense that SLG is used as a tool to test and improve linguistic theories about the relation between the intonation of an utterance and other aspects of its structure.

The research objective in this work is to design, build and evaluate a linguistically-bases SLG system that generates speech with pitch accents and intonational boundaries that are most acceptable to experts. It is driven by research questions regarding the adequacy of the linguistic theories employed, as well as questions regarding the best architecture for an SLG system. While focusing on intonation (phonological phrasing, intonational phrasing, and pitch accent placement in Dutch) and system architecture, it provides an overview of all stages of the SLG process and the underlying syntactic, phonological and phonetic theory, and includes contributions to text planning, surface realization, tune choice, and speech synthesis. In addition, it reports on an evaluation experiment, comparing output of the SLG system with that of human experts and other systems at the symbolic level.

This study is of interest to theoretical and computational linguists, as well as researchers in the field of Speech Synthesis and Natural Language Generation.