

## SPEECH SYNTHESIS FROM INTERPOLATED LOG-AREA CODED TRANSITIONS

Hans Werner Strube, Drittes Physikalisches Institut, Universität Göttingen, Bürgerstrasse 42-44, D-3400 Göttingen, F.R. Germany

A speech synthesis system is described, based on concatenation of elements from a pool of 48 stationary sound segments together with many (up to 1322) sound transitions. The speech signals are coded by 13 log-area-ratios,  $\log(\text{pitch})$ ,  $\log(\text{power})$  and two binary parameters switching the noise and pulse generators. Output is generated by a computer-controlled hardware synthesizer (Strube 1977). The stationary sounds are coded by a single parameter-frame and the transitions by two frames only, taken from real speech. Intermediate frames are restored by interpolation during synthesis. Direct linear interpolation of the above parameters is in most cases a fairly good choice compared to other possibilities. The transition boundaries to be stored were determined using the spectra, the parameter curves and a subjectively optimized fit of straight-line trains to the curves.

The transition table is addressed through a (37 x 37)-matrix with first and second phoneme as row and column index, also containing the transition length. Thus the same frame pairs may be used for different transitions, also in opposite time sequence. When a transition is not yet in the table, single phonemes are concatenated; for many pairs, this is even good enough. The quasi-stationary part of a sound is either also interpolated or inserted discontinuously as a constant portion. Treatment of different phonemes, excitation, timing, and intrinsic pitch are controlled by a sound-property table.

Synthesis occurs in real time during input-text evaluation. Input is in ASCII characters, closely matching the IPA transcription. Pitch is controlled by numbers in the input text, whereas duration and intensity are given by the program. Pitch changes are smoothed and intrinsic pitch is added by the program. Investigations in automatic intonation rules (J. Kretschmar 1978) and intrinsic pitch in German are in progress. Results and examples will be presented.

#### References

- Kretschmar, J. (1978): "Untersuchungen zum Tonhöhenverlauf deutscher Sätze für die Sprachsynthese", Fortschritte der Akustik-DAGA '78, Berlin: VDE-Verlag, 455-458.
- Strube, H.W. (1977): "Synthesis part of a 'log area ratio' vocoder in analog hardware", IEEE Trans. ASSP-25, 387-391.