

SELECTION OF GLOTTAL EXCITATION PARAMETERS OPTIMIZING
THE NATURALNESS OF SYNTHETIC SPEECH

Janusz Zalewski and Wojciech Myślecki, Institute of Telecommunication and Acoustics, Technical University of Wrocław, Wrocław, Poland

The shape and periodicity of source excitation influences the naturalness of synthetic speech (1,2). In the present study these problems were investigated for short phrases of Polish synthetic speech.

Procedure, Experiments and Results

Glottal pulses were shaped by means of time functions previously examined by Rosenberg (2). Amplitude and frequency of the glottal excitation were controlled by a set of simple rules. All synthetic utterances were generated by a digital series formant synthesizer and were subjectively evaluated by means of an A-B test. To obtain an interval preference scale, Thurstone's model V of comparative judgment was accepted and Mosteller's least squares solution (3) was used.

The goal of the experiments was to determine optimal pulse shape function, optimal relative opening t_o and closing t_c times, and optimal amplitude A_D and frequency F_D of fine pitch deviation. In the first experiment optimal t_o, t_c combinations for five examined pulse shape functions f_A, \dots, f_E^2 were established. In the next experiment these functions were compared for naturalness and it was found that the best was the trigonometric function f_C with $t_o = 0.41$, $t_c = 0.2$. The optimal deviation parameters were $A_D = 1.7$ Hz and $F_D = 6$ Hz.

Conclusions

The results of the experiments have shown that pulse shape and t_o, t_c values strongly influence the naturalness of synthetic phrases. The obtained preference scores for investigated phrases, synthesized applying various pulse shape functions, are quite similar to Rosenberg's results (2). It was also found that, for each pulse shape, besides the optimal t_o, t_c pair, there was a distinctively different second pair of t_o, t_c which provided comparable naturalness.

References

1. Zalewski, J. and W. Myślecki (1975): "Research on the selection of F_o for the optimal synthesis of vowels", 8 Int. Cong. Phon. Sc., Leeds.
2. Rosenberg, E.A. (1971): "Effect of glottal pulse shape on the quality of speech from parallel formant synthesizer", JASA 49.
3. Mosteller, F. (1951): "Remarks on the method of paired comparisons - the least squares solution", Psychometrica 16.