

PERCEPTION OF VOWEL FORMANT TRANSITIONS

Iman H. Slis, Instituut voor Fonetiek, Katholieke Universiteit, Nijmegen, The Netherlands

In the majority of speech synthesizers the filter parameters are changed stepwise during formant transitions. Demonstrations with vocoders show that continuous speech which is synthesized this way can be of excellent quality. It is not known, however, whether a steplike approximation will also suffice for shorter stimuli. Therefore, we started a series of experiments concerning the detectability of vowel formant transitions within one F1 period.

In a pilot experiment, a formant filter was excited with one pulse. The filter parameters in the reference stimulus were kept constant. In the synthesis of the test stimuli the coefficients of the formant filter were linearly interpolated on a per-sample basis; bandwidth was kept constant. Only rising formant transitions were synthesized. The starting points of the test stimuli were arranged around the frequency of the reference stimulus. This experiment was repeated with four filters in series. These experiments were then extended using stimuli of more than one excitation pulse.

We hypothesized that the auditory impression caused by the test stimuli would be dominated by the duration of the first cycle, the amplitude of which is much greater than that of the remaining ones. Therefore we expected to find almost no perceptual difference between reference and test stimuli which have a first cycle of approximately equal duration. However, the first results show that six subjects were capable of hearing substantial differences between reference and all test stimuli; and that the smallest difference was found when the duration of the third cycle of the test stimulus was equal to the cycle duration of the reference stimulus.