

Wavelet parameters for speech synthesis – Brian C. Tuttle and Claus P. Janota

A standard method of analyzing human speech is to divide it into its constituent phonemes. Synthesizing speech from these phonemes without including the effects of coarticulation, however, can result in discontinuous and sometimes unintelligible sounds. A recently proposed model for the analysis of speech coarticulation uses wavelet system characterization where wavelet transforms describe the time-frequency behavior of a signal's transmission channel. The objective of this research is to verify the proposed model by synthesizing the original speech from its analysis. The coarticulated speech in question is the consonant–vowel–consonant combination that occurs in words such as “deed” or “bib” where the two consonants are the same. These words along with their vowels spoken in isolation are recorded digitally and processed using the wavelet system model. The result is an analysis of the coarticulated word with respect to its vowel. The analysis shows in time and frequency how the vowel changes when spoken in context. Synthesizing the speech using the inverse process serves to verify the model. Comparing the result to the original speech reveals that the accuracy is determined mainly by the scale resolution of the wavelet transforms.

138th Meeting of the Acoustical Society of America, Columbus, OH. J. Acoust. Soc. Am. **106**, 2249 (1999)