UCLA
UCLA Previously Published Works

Title
Perceptual importance of time-domain features of the voice source.

Permalink
https://escholarship.org/uc/item/9jc8k8b4

Journal
The Journal of the Acoustical Society of America, 135(4)

ISSN
0001-4966

Authors
Garellek, M
Chen, G
Gerratt, BR
et al.

Publication Date
2014-04-01

DOI
10.1121/1.4878046

Peer reviewed

Our previous study examined the perceptual adequacy of different source models. We found that perceived similarity between modeled and natural voice samples was best predicted (in the time dimension) by thematch between waveforms at the negative peak of the flow derivative (R^2 = 0.34). The extent of fit during the opening phase of the source pulses added only 2% to perceived match. However, in that study model, fitting was unweighted, and results might differ if another approach were used. In this study, we constrained the models to fit the negative peak of the flow derivative precisely. We fit 6 different source models to 40 natural voice sources, and then generated synthetic copies of the voices using each modeled source pulse, with all other synthesizer parameters held constant. We then conducted a visual sort-and-rate task in which listeners assessed the extent of perceived match between the original natural voice samples and each copy. Discussion will focus on the specific strengths and weaknesses of each modeling approach for characterizing differences in vocal quality, and on the importance of matches to specific time-domain events versus spectral features in determining voice quality.