USING STATE FEEDBACK TO CONTROL TO AN ARTICULATORY SYNTHESISER

Ian S. Howard[1] & Peter Birkholz[2]
[1]Centre for Robotics and Neural Systems, University of Plymouth, Plymouth, PL4 8AA UK
Email: ian.howard@plymouth.ac.uk

[2]Institute of Acoustics and Speech Communication, Technische Universität Dresden, 01062 Dresden, Germany.
Email: peter.birkholz@tu-dresden.de

ABSTRACT

In order to speak, we need to move our speech articulators in an appropriate fashion. Therefore, at its lowest mechanical level, speech production can be considered to be a motor task that leads to acoustic consequences. Of course, it is the latter which is of primary interest to a listener. Controlling any real physical system, including the human speech apparatus, involves not only dealing with the dynamics of the moving parts, but also with any unpredictable disturbances that may occur. The field of control engineering provides a useful means to understand such issues, and also offers computational solutions to these kinds of problems. State feedback control has been recently proposed as a good approach to understand observed phenomena in human speech production [1]. In particular, the observed behaviours seen in perturbations studies, which involve both proprioceptive and auditory perturbations, can be at least partially explained using this paradigm. Here we apply state feedback control to operate an articulatory speech synthesiser [2], which we use to generate some simple speech utterances. We first explain the architecture of the system from a signal flow control perspective. We then go on to show that the internal models needed for effective operation can be acquired during a babbling phase. In particular, the inverse mapping between the synthesizer's control parameters and their auditory consequences can be learned using a neural network. Such an inverse model provides a means to map errors that occur in the acoustic speech domain back to the articulatory domain, where they can be used to make compensatory adjustments. Finally, we showed this approach can lead to stable productions of speech.

REFERENCES

[1] J. F. Houde and S. S. Nagarajan, "Speech production as state feedback control.," *Front Hum Neurosci*, vol. 5, p. 82, 2011.

[2] P. Birkholz, D. Jackel, and B. J. Kröger, "Construction and Control of a Three-Dimensional Vocal Tract Model," presented at the 2006 IEEE International Conference on Acoustics Speed and Signal Processing, 2006, vol. 1.