# multipulse LPC articulatory modeling in the Wisconsin X-ray microbeam speech production database

Athanasios Katsamanis, Erik Bresch, Louis Goldstein and Shrikanth Narayanan
http://sail.usc.edu

## so, what?

Our goal is to computationally model speech production in a way that it would allow us to:

- gain deeper understanding of the underlying speech system by exploiting rich articulatory data (e.g., as acquired by x-ray microbeam, rt-MRI)
- discover new or strengthen suggested links between articulatory observations and theoretical expectations

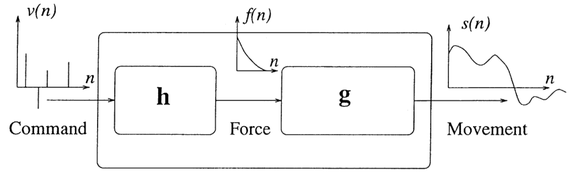
In this direction, we adopt the viewpoint of articulatory phonology [1] which is based on the description of an utterance "as an organized pattern of overlapping articulatory gestures".

Multipulse modeling of articulatory movements can provide a flexible and intuitive representation of the dynamic behavior of speech articulators [2].

Original multipulse LPC articulatory modeling studies only focused in a limited set of articulations by a single speaker as these were imaged using X-rays.

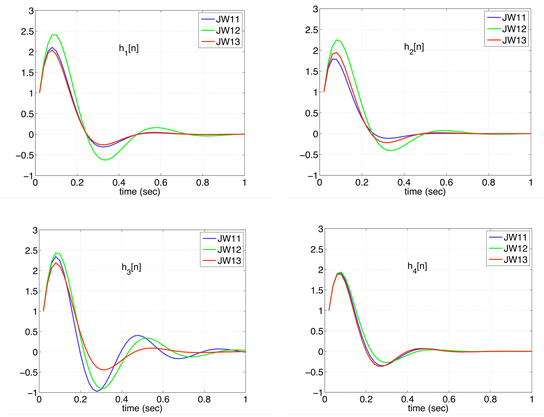
We investigate the application of the multipulse modeling framework in the Wisconsin X-ray microbeam speech production database [7].

## multipulse modeling



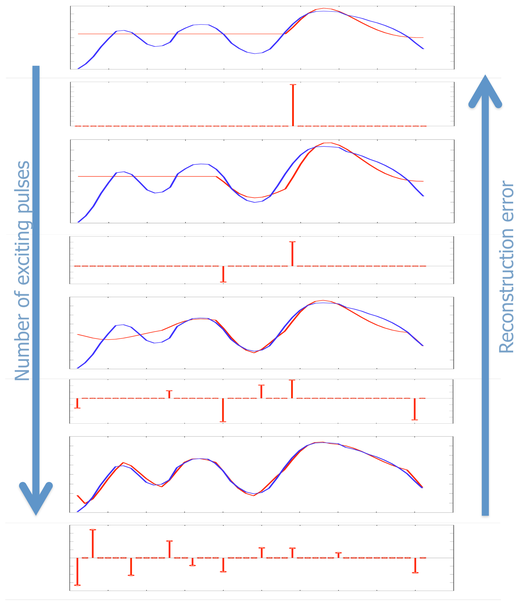

$$H(z) = \frac{1}{1 - \sum_{i=1}^{2} a_i z^{-i}} \quad G(z) = \frac{1}{1 - \alpha z^{-1}}$$

- **G(z) is a de-emphasis filter**
- **H(z) is estimated by Linear Prediction analysis of articulatory variable signals**

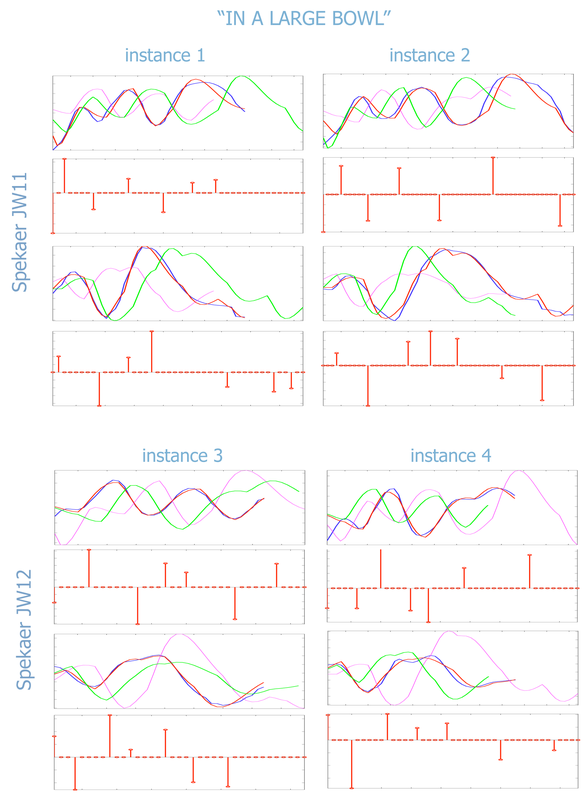



## matching pursuit

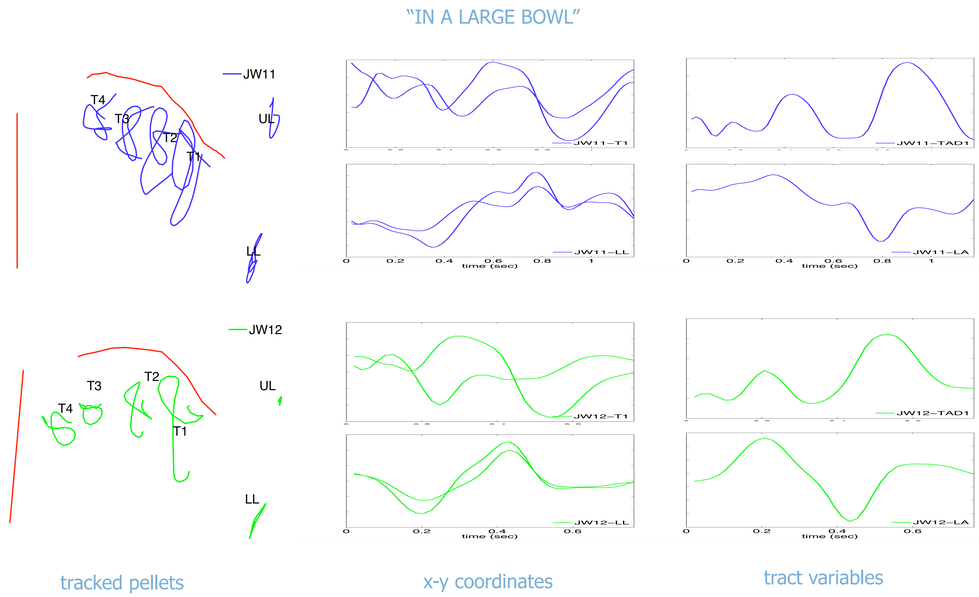
- **Greedy search to estimate the pulse amplitudes and locations**
- **Pulse amplitudes are re-optimized each time a new pulse is introduced**
- **Minimize the squared reconstruction error**



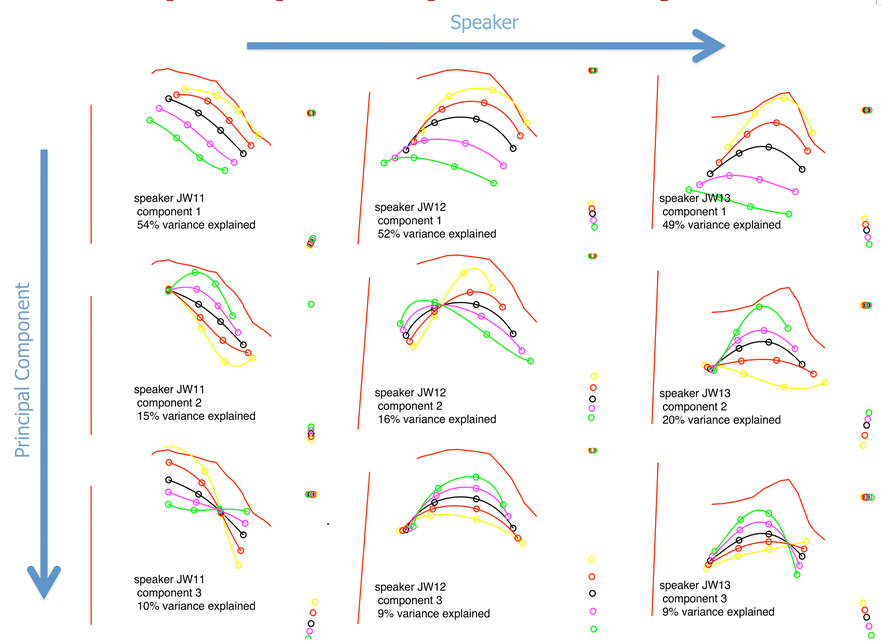

- **Blue: Original articulatory signals**
- **Red: Reconstructed using pulses from same instance**
- **Green: Reconstructed using pulses from other instance**
- **Magenta: Reconstructed using pulses from other speaker**




## extraction of tract variables




## principal component analysis




## references

[1] Catherine Browman and Louis Goldstein, "Towards an articulatory phonology," Phonology Yearbook, vol. 3, pp. 219–252, 1986.
[2] Soumya Bouabana and Shinji Maeda, "Multi-pulse LPC modeling of articulatory movements," Speech Communication, vol. 24, pp. 227–248, 1998.
[3] Catherine R. Browman and Louis Goldstein, "Articulatory gestures as phonological units:," Phonology, vol. 6, pp. 201–251, 1989.
[4] B.S. Atal and Suzanne L. Hanauer, "Speech analysis and synthesis by linear prediction of the speech wave," J. of the Acous. Soc. Am., vol. 50, pp. 637–655, 1971.
[5] B. S. Atal and J. Remde, "A new model of LPC excitation for producing natural-sounding speech at low bit rates," in Proc. IEEE Int'l Conf. Acous., Speech, and Signal Processing, 1982.
[6] Sharad Singhal and S. Bishnu Atal, "Amplitude optimization and pitch prediction in multipulse coders," IEEE Trans. Acoust., Speech, Signal Process., vol. 37, pp. 317–326, 1989.
[7] Westbury, J.R. X-Ray Microbeam Speech Production Database User's Handbook. Madison, WI: University of Wisconsin Waisman Center, 199

**acknowledgments :** This work is supported by NIH Grant DC007124.