Automatic identification of stable modes and fluctuations in a repetitive task using real-time MRI

Adam Lammert[1], *Michael Proctor*[2], *Louis Goldstein*[1,2], *Marianne Pouplier*[3], *and Shrikanth S. Narayanan*[1]
University of Southern California[1], Haskins Laboratories[2], LMU München[3]

web: sail.usc.edu/span
email: lammert@usc.edu

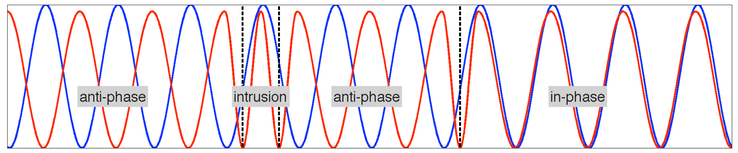
Stability in Articulator Coordination

Repetiton Tasks: Subjects repeat alternating syllables at increasing rates (e.g., cop-top, skip-hip, bag-ban, flee-free, kip-cop).

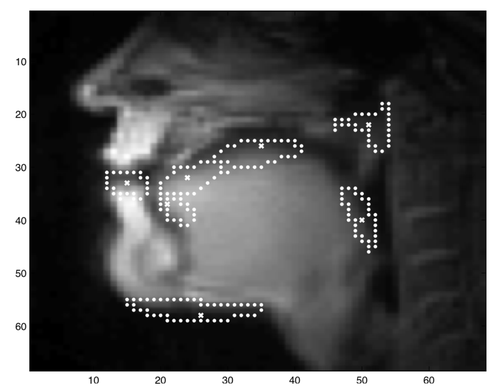
Transients Gestural intrusions and reductions are most common errors. In the extreme, they result in mode shifts [1].

Mode Shifts Two primary modes of coordination are common: in-phase and anti-phase [2].

Cartoon Example



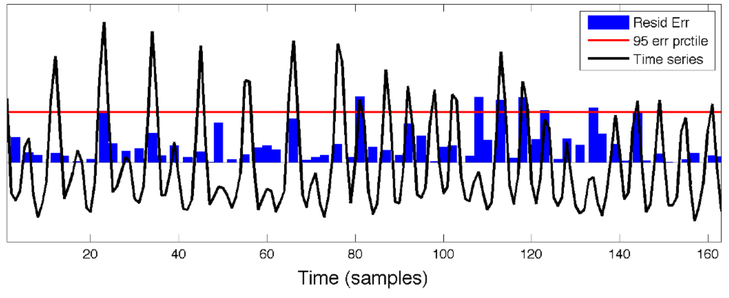
Data from rtMRI and EMA



- EMA: 2 subjects (1 male, 1 female), 200Hz acquisition rate, midsagittal coordinates of flesh points on the lips and tongue.
- RT-MRI: 4 subjects (2 male, 2 female), 33Hz acquisition rate, vocal tract constrictions at any point along the vocal tract [3].

Detection by Linear Prediction of Extrema

Alveolar Time Series (token: 'cop-top')



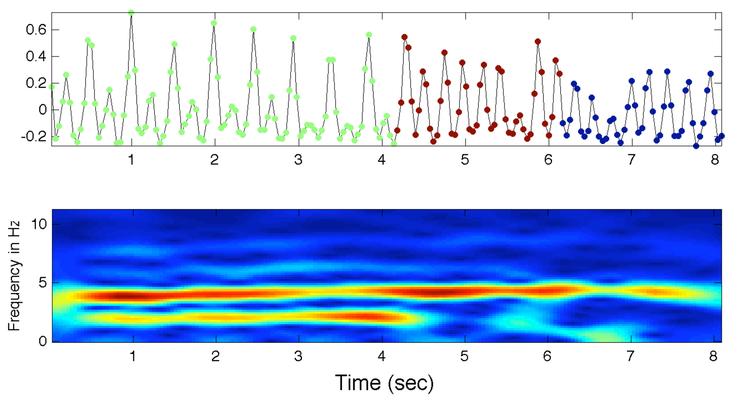
Amplitude of the m^{th} extremum is y_m. The estimate of that amplitude is:

$$\hat{y_m} = -a_2 y_{m-1} - a_3 y_{m-2} - \cdots - a_{p+1} y_{m-p} \quad (1)$$

The coefficients $[a_2, \cdots a_{p+1}]$ are trained on the entire time series.

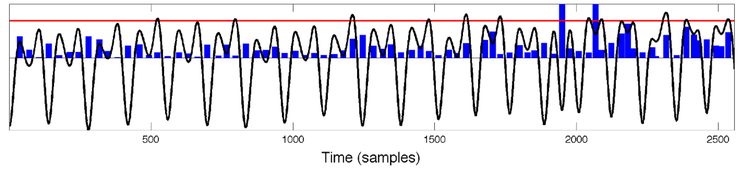
Discrimination by Frequency Domain Analysis

Alveolar Time Series (token: 'cop-top')

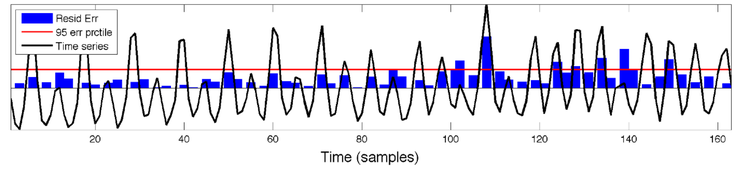


- 128-point spectrogram
- Windows of 1.64s width with 0.045s hop size
- GMM clustering with 3 components

Additional Examples



- EMA: Lip aperture during token 'cape-Kate'
- Single intrusion and return to anti-phase mode



- rtMRI: Velar constriction during token 'cop-top'
- Complex! Reduction followed by a mode shift

Conclusions:

1) Linear prediction modeling of time series affords detection of transients and mode shifts.

1) Frequency domain analysis affords discrimination between stable modes.

2) Detecting an error depends heavily on one's concept and definition of errant behavior.

Future Work:

1) Using multiple articulators could allow for detection of more subtle or new types of errors.

2) Utility of physical models for detecting errors.

References

[1] Goldstein, L., Pouplier, M., Chen, L., Saltzman, E. and Byrd, D. (2007). "Dynamic action units slip in speech production errors", Cognition 103 (3): 387–412.

[2] Schöner, G. and Kelso, J.A.S. (1988). "Dynamic Pattern Generation in Behavioral and Neural Systems", Science 239 (4847): 1513–1520.

[3] Narayanan, S., et al. (2004). "An approach to real-time magnetic resonance imaging for speech production", JASA, 115:1771.

Acknowledgements

This work was supported by NIH Grant DC007124 and DFG PO 1269/1-1.