Running or speed-walking?

Simulations of speech production at different rates

Joe Rodd, Hans Rutger Bosker, Mirjam Ernestus, Antje S. Meyer, Louis ten Bosch

1 BACKGROUND

Speakers can vary their speaking rate at will. How do they do this?

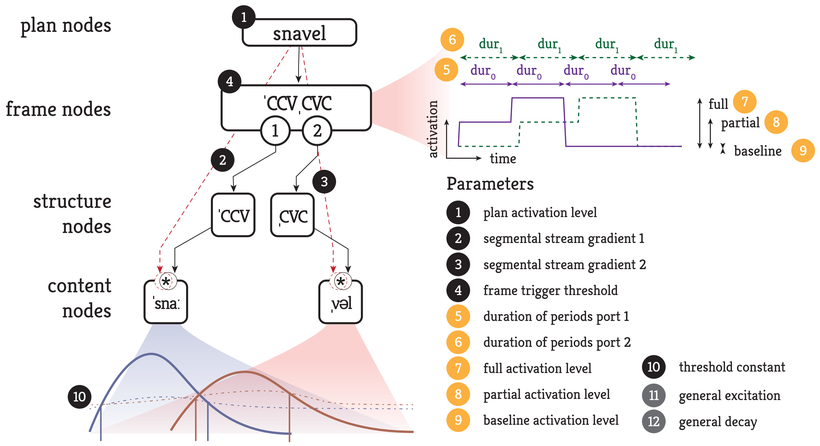
Presumably, different cognitive 'regimes' are engaged to produce different speaking rates.

How do these cognitive regimes relate to each other? Possibilities:

- Regimes entirely discrete (and differ qualitatively)
- Regimes qualitatively similar (and differ quantitatively)
- mixture ('gaits' of speech planning)

Our approach: construct a computationally implemented model of speech planning, model parameters ≈ cognitive dimensions

2 CONNECTIONIST MODEL



Parameters

1. plan activation level
2. segmental stream gradient 1
3. segmental stream gradient 2
4. frame trigger threshold
5. duration of periods port 1
6. duration of periods port 2
7. full activation level
8. partial activation level
9. baseline activation level
10. threshold constant
11. general excitation
12. general decay

Derived from Dell, Burger and Svec model (1997)

Joe.Rodd@mpi.nl

Max Planck Institute for Psycholinguistics

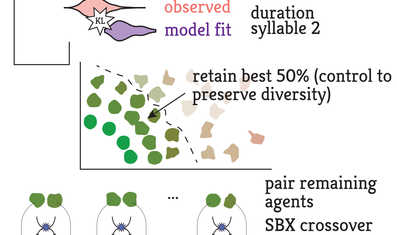
CLS | Centre for Language Studies
Radboud University

NWO
Netherlands Organisation for Scientific Research

LANGUAGE in INTERACTION

3 SPEECH CORPUS

Multiple picture naming (Dutch disyllables) to elicit controlled productions at three speaking rates.

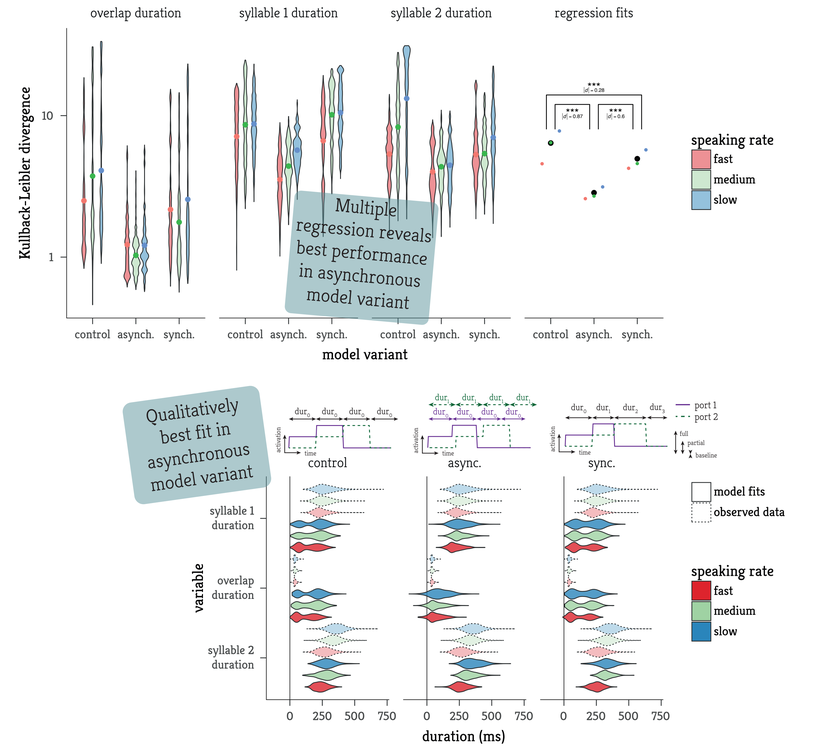


PiNCeR corpus; Rodd, Bosker, Ernestus, ten Bosch, and Meyer. Under review. 'How we regulate speech rate: phonetic evidence for a "gain strategy" in speech planning'.

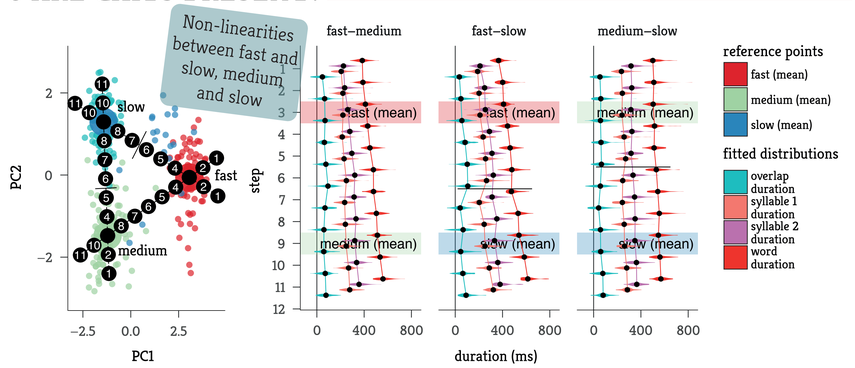
4 PARAMETER LEARNING

NSGA-III; Deb and Jain. 2014. 'An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems With Box Constraints'. IEEE Transactions on Evolutionary Computation 18 (4):577–601.

5 HOW DID MODEL VARIANTS PERFORM?



6 ARE GAITS PRESENT?



7 CONCLUSIONS

- Model learns to fit temporal properties of speech.
- Evidence for two gaits: a 'dawdle' gait for slow speech, and a 'walk' gait for medium and fast speaking rates.