Chunking in the production of tap/flap sequences

Connor Mayer
University of California, Irvine

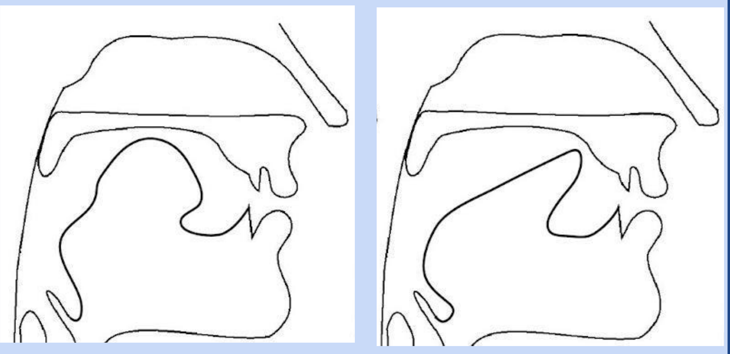
Donald Derrick
University of Canterbury

Bryan Gick
University of British Columbia, Haskins Labs

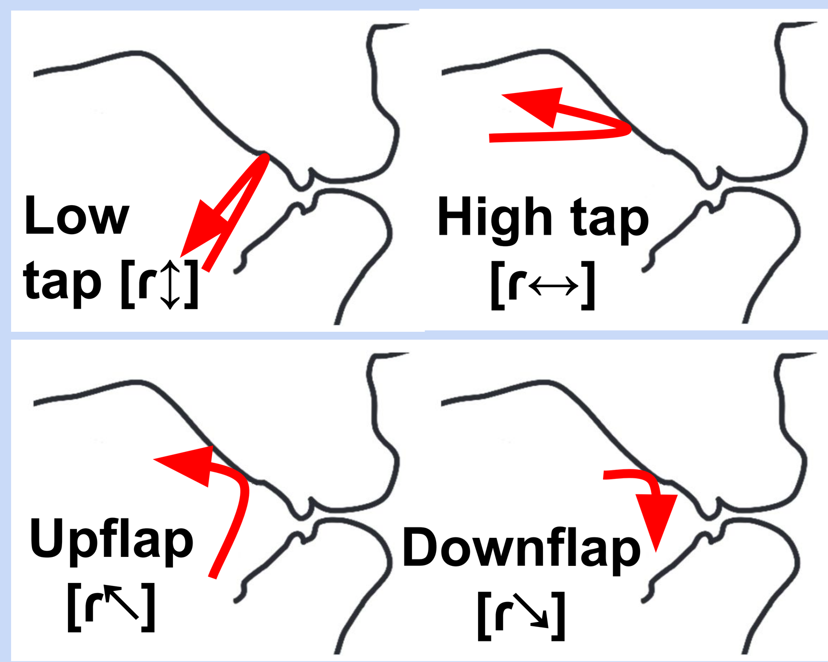
2024 Annual Meeting on Phonology

1. Categorical variation in two English sounds

Rhotics



Taps/flaps



2. Gait change in tap/flap sequences [1,2]

Input: /VTVTV/ (e.g. 'edit a')

North American English (NAE) - gait change

[V_Lɾ↕V_Lɾ↕V_L] → (Faster speech rate) [V_Lɾ↖V_Hɾ↘V_L]

New Zealand English (NZE) - no gait change

[V_Lɾ↕V_Lɾ↕V_L] → (Faster speech rate) [V_Lɾ↕V_Lɾ↕V_L]

3. Why doesn't NZE show gait change? [2]

Rhotic vowels induce variability in adjacent taps based on mechanical ease [8]

'Berta': [ɹɾ↘] 'Otter': [ɾ↖ɻ] 'Murder': [ɹɾ↔ɹ], [ɹɾ↖ɻ], …

Chunking: more efficient to reuse highly-practiced movement patterns than to compute new ones [3,4, 5, 6, 7]

NAE has rhotic vowels

- Speakers build a repertoire of diverse strategies (chunks)
- **Speakers draw from this repertoire in new contexts**

NZE does not

- Speakers learn fewer strategies (chunks)
- **Speakers' repertoire of strategies is limited**

NAE: [V_L ɾ↖] ɻ [ɾ↘ V_L] — Saturday → [V_L ɾ↖] V_H [ɾ↘ V_L] — edit a (**chunk reuse in new context**)

NZE: V_L ɾ↕ V_L ɾ↕ V_L — Saturday; V_L ɾ↕ V_L ɾ↕ V_L — edit a

4. Simulation study

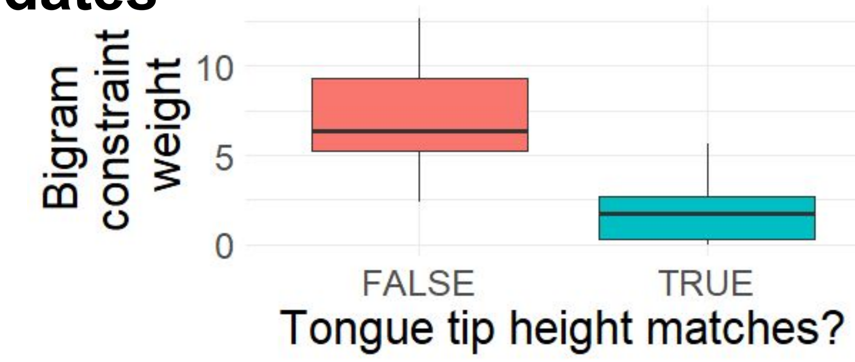
Integrating the concept of chunking into a model that learns to produce tap/flap sequences qualitatively replicates the differences between NAE and NZE speakers

Basic model

Implemented in maximum entropy optimality theory [9]

- Uses **weighted constraints** to generate **probability distributions over candidates**

/V/ → [V_L, V_H]
/R/ → [ɹ, ɻ]
/T/ → [ɾ↕, ɾ↖, ɾ↘, ɾ↔]



Markedness constraints:

- all unigram (e.g. *ɾ↔) and bigram (e.g. *ɾ↖V_L) constraints
- Weights fit to frequencies from ultrasound study [8]

Constraints for chunking - weights set by hand

- **UseListed:** Violation is sum of costs of all chunks in form
  - Chunk cost decreases as its usage frequency increases
- **Share**: One violation per chunk

NAE input: {VTV, VTVTV, VTVTVTV, VTR, VTVTR, RTV, …}
NZE input: {VTV, VTVTV, VTVTVTV}

Outputs: Chunked candidates ("|" = chunk boundary)
[V_Lɾ↕V_Lɾ↕V_L], [V_Lɾ↕|V_Lɾ↕V_L], [V_L|ɾ↕|V_L|ɾ↕|V_L], [V_Lɾ↖V_Hɾ↘V_L], …

Starting chunk inventory: {V_L, V_H, ɹ, ɻ, ɾ↕, ɾ↖, ɾ↘, ɾ↔}

Part 1: Chunk learning

Sample input → /VTV/ → Compute probabilities using grammar →

[V_Lɾ↕V_L]	0.95
[V_L|ɾ↕V_L]	0.03
[V_L|ɾ↕|V_L]	0.01
…	…

→ Sample output → [V_Lɾ↕V_L] → Update chunk inventory, Including sub-chunks →

V_Lɾ↕V_L += 1,
V_Lɾ↕ += 1
ɾ↕V_L += 1
…

x10,000

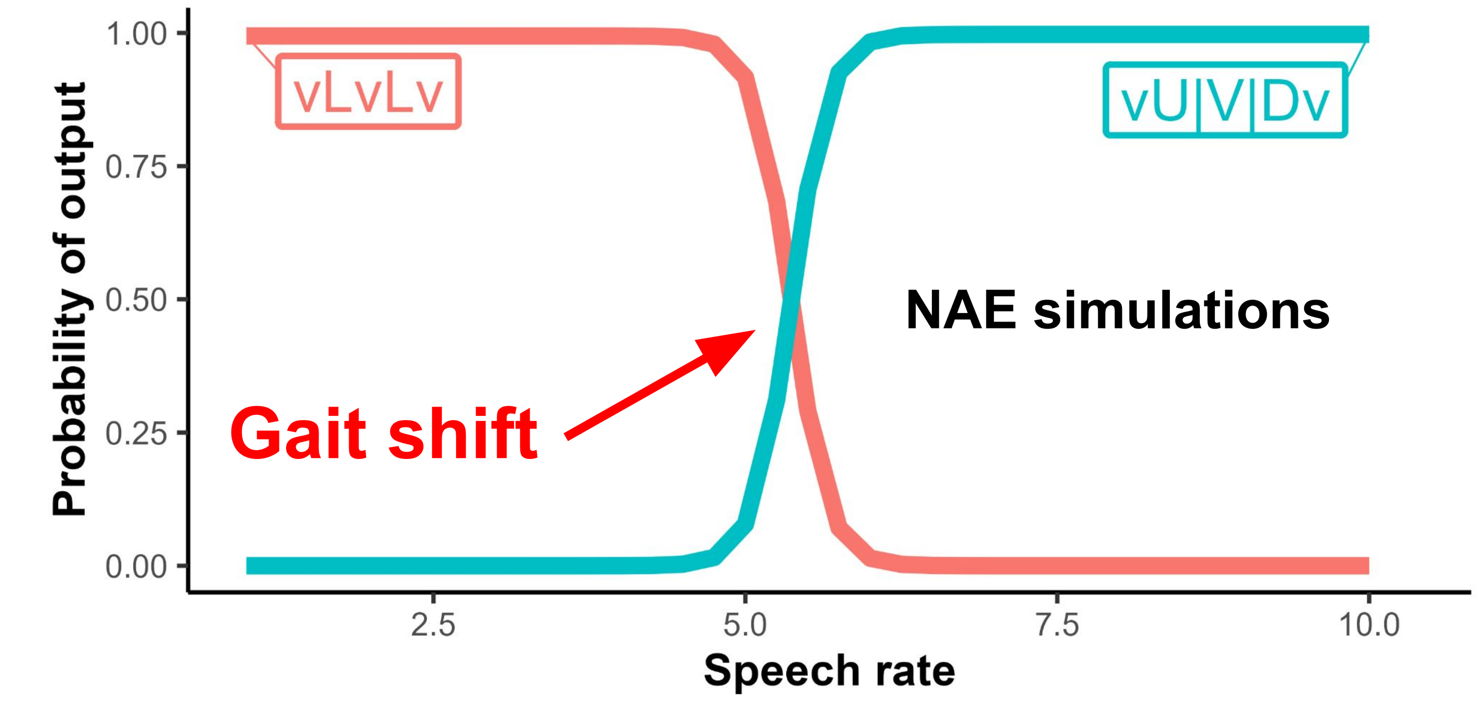
Part 2: Productions under different speech rates

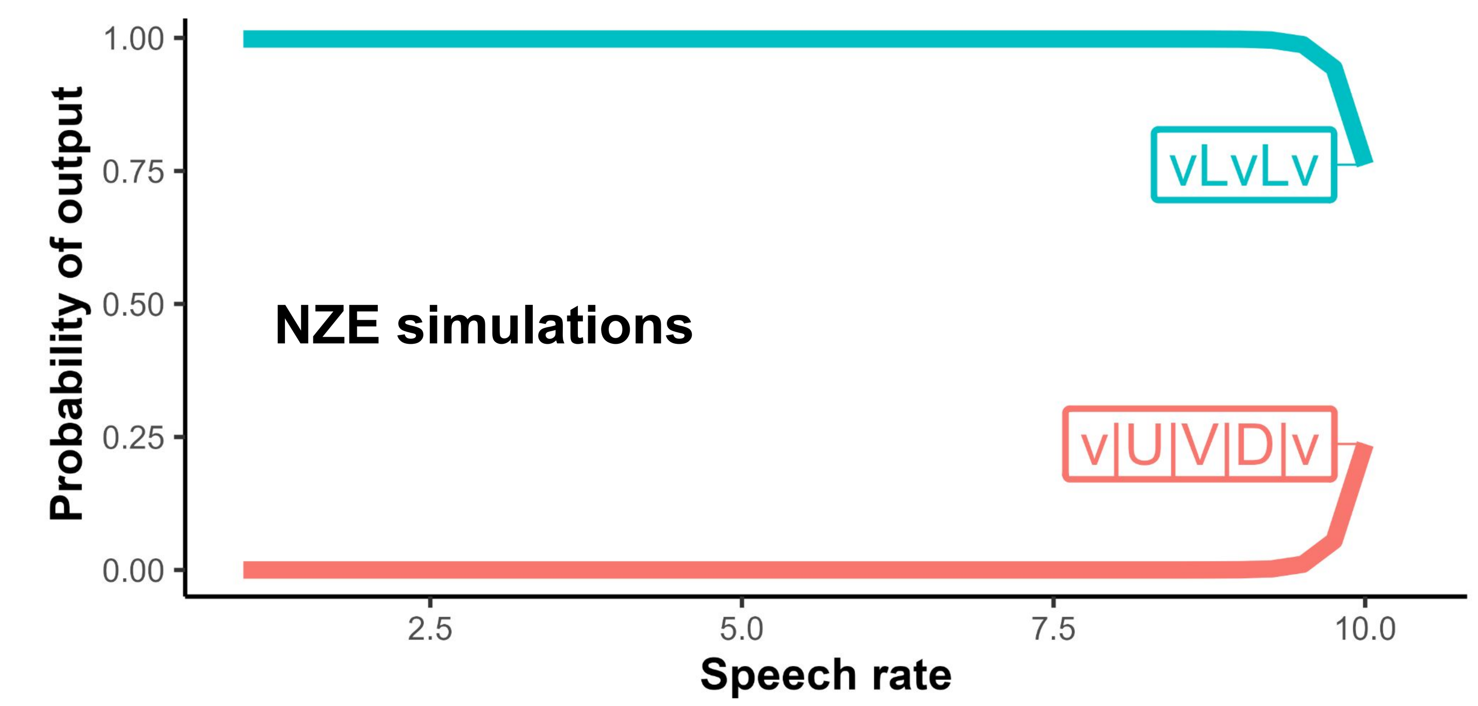
/VTVTV/ → Compute probabilities using grammar and chunk inventory; Scale bigram constraint weights up as speech rate increases (cf. [10]) →

[V_Lɾ↕V_Lɾ↕V_L]	0.75
[V_Lɾ↕|V_Lɾ↕V_L]	0.004
[V_Lɾ↖V_Hɾ↘V_L]	0.20
…	…

Rates: [1, 1.25, 1.5, …, 9.5, 9.75, 10]

5. Simulation results





6. Discussion

NAE speakers undergo gait shift because they can repurpose existing chunks in new contexts

- **Chunks learned in rhotic contexts can be deployed in non-rhotic contexts**

NZE speaks have a more limited chunk repertoire

- **Not as flexible in their production strategies**

Considering phonology from an **embodied perspective** provides substantive insights into phonological patterns

- **Speech production is not optimal**, but includes the reuse of frequently produced movements
- May provide substantive account of some frequency and paradigm uniformity effects

References. [1] Derrick, D., & Gick, B. (2021). Gait change in tongue movement. Scientific Reports, 11(1), 16565. [2] Derrick, D., Mayer, C., & Gick, B. (2024). Uniformity in speech: The economy of reuse and adaptation across contexts. Glossa: a journal of general linguistics, 9(1). [3] Verwey, W. B. (2001). Concatenating familiar movement sequences: The versatile cognitive processor. Acta psychologica, 106(1-2), 69-95. [4] Abrahamse, E. L., et al. (2013). Control of automated behavior: insights from the discrete sequence production task. Frontiers in human neuroscience, 7, 82. [5] Keating, P. (2003, August). Phonetic and other influences on voicing contrasts. ICPhS 15 (pp. 20-23). [6] Chodroff, E., & Wilson, C. (2017). Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. Journal of Phonetics, 61, 30-47. [7] Faytak, M. D. (2018). Articulatory uniformity through articulatory reuse: insights from an ultrasound study of Sūzhōu Chinese. University of California, Berkeley. [8] Derrick, D., & Gick, B. (2011). Individual variation in English flaps and taps: A case of categorical phonetics. Canadian Journal of Linguistics/Revue canadienne de linguistique, 56(3), 307-319. [9] Goldwater, S., & Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In Proceedings of the workshop on variation within Optimality Theory (pp. 111-120). [10] Coetzee, A. W. (2016). A comprehensive model of phonological variation: Grammatical and non-grammatical factors in variable nasal place assimilation. Phonology, 33(2), 211-246.