Fitts' law in tongue and lip movements of repetitive speech

Tanner Sorensen,[1] Mark Tiede,[2] Adam Lammert,[3] Louis Goldstein,[1] Shrikanth Narayanan[1]
[1]University of Southern California, Los Angeles, CA
[2]Haskins Laboratories, New Haven, CT
[3]Worcester Polytechnic Institute, Worcester, MA

Background

- Prosody refers to the rhythmic grouping of supra-segmental linguistic units into prosodic domains (e.g., grouping pairs of syllables into feet).
- In speech production, modulating the rate of articulator movement marks the temporal extent of a prosodic domain.
- Specifically, movement slows down at the edge of the prosodic domain and speeds up inside the prosodic domain.
- In the theory of prosody developed within Articulatory Phonology [3, 2], this rate modulation is the result of varying a parameter of gestural stiffness (a result of local clock-slowing at the edges of prosodic domains; see [5]).
- At the edge of the prosodic domain, low gestural stiffness produces slow movements. Inside the prosodic domain, high gestural stiffness produces fast movements [4, 6, 8, 1].

Hypothesis

Our hypothesis is based on two premises:

1. Gestural stiffness k is low at the edge of prosodic domains and high inside prosodic domains [4, 6, 8, 1].
2. As shown by [11, 13], gestural stiffness k is equal to the inverse square of the slope of Fitts' law [9, 11, 10], a linear equation relating movement time T and an index of movement difficulty I_d.

$$T = a + \left(\frac{1}{\sqrt{k}}\right) I_d \qquad \text{(Fitts' law)}$$

The slope is used as a measure of information throughput [9] in human-computer interaction and ergonomics.

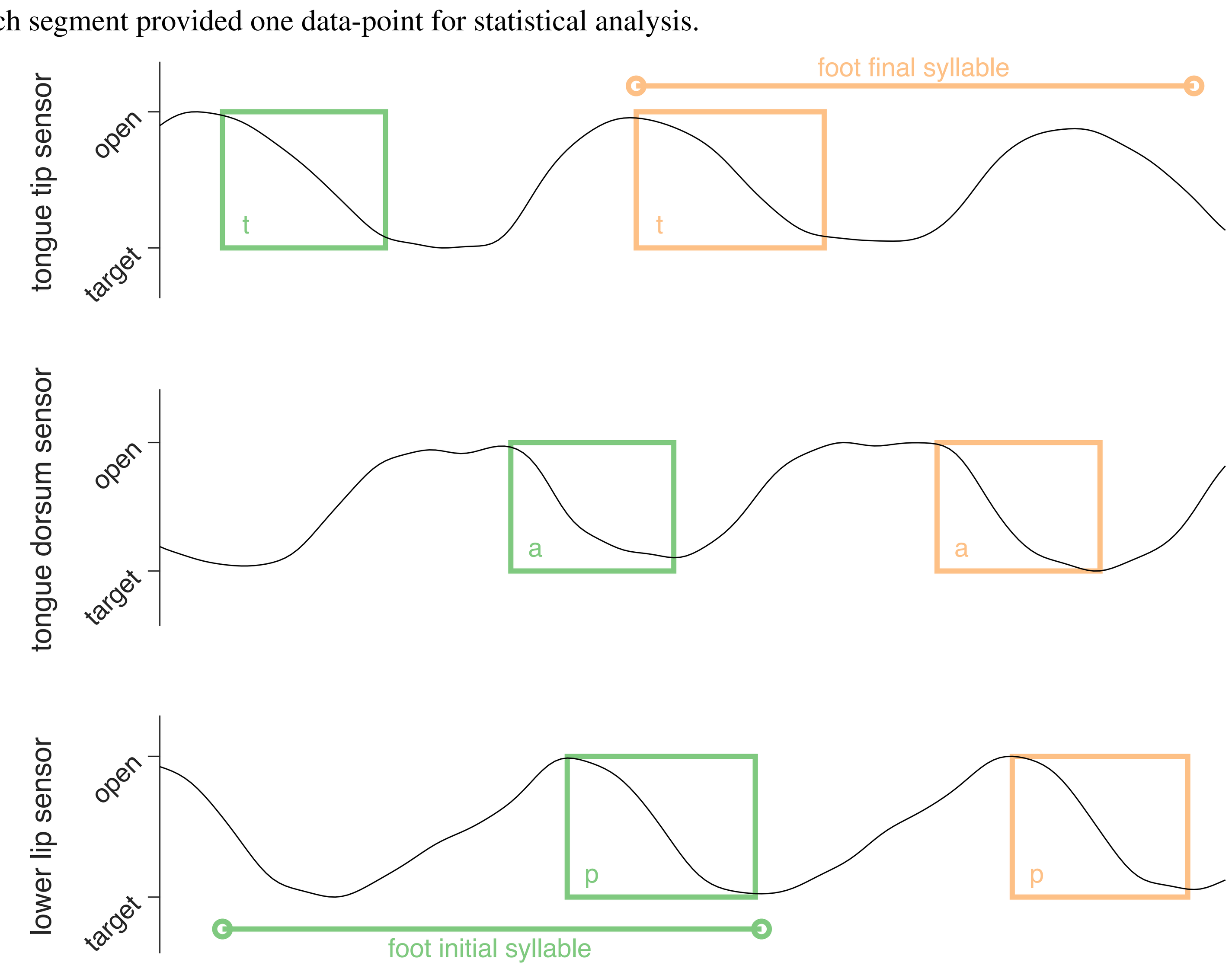
Premises 1 and 2 lead to the following hypothesis:

> At the edge of the prosodic domain, low gestural stiffness makes the slope of Fitts' law steep.
> Inside the prosodic domain, high gestural stiffness reduces the slope of Fitts' law.

Methods

Data

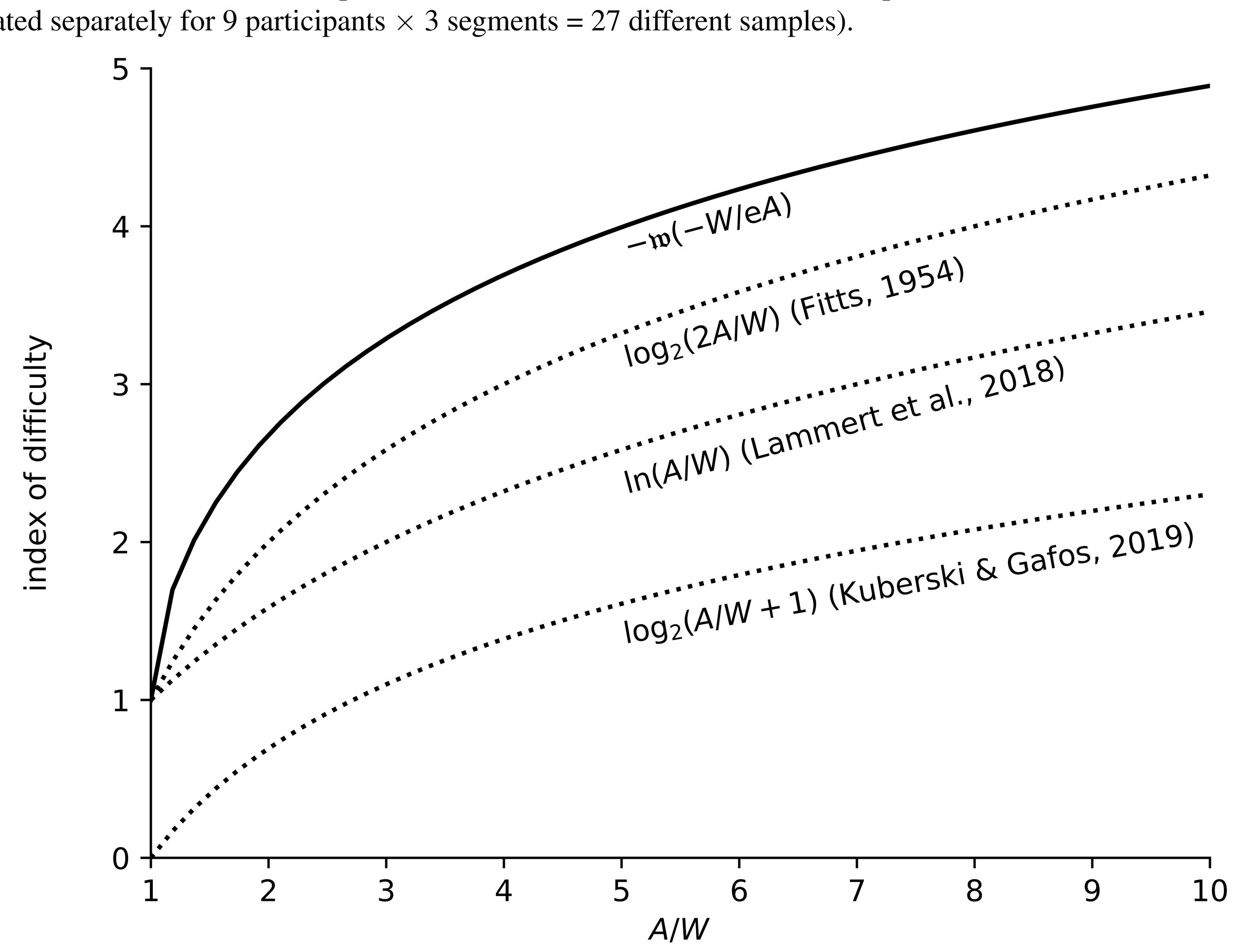
- 54 electromagnetic articulography recordings (9 participants × 6 trials/participant × 1 recording/trial) [14].
- Participants repeated a pair of words (e.g., "top-top") in time to a metronome (1 word/beat; 170 bpm from time 0-7.5 s, increasing to 230 bpm from time 7.5-15 s).
- The pair of words consisted of two identical syllables that constituted a prosodic foot.
- Each syllable consisted of a bilabial, coronal, or velar onset stop consonant, a vowel nucleus, and a bilabial, coronal, or velar coda stop consonant. Onset and coda were never identical.
- Each segment provided one data-point for statistical analysis.



- The dependent variable was *movement time* (start- and end-points were defined as speed rising above and falling below 10% peak velocity).
- The independent variable was an *index of difficulty* that [13] derived from the Task Dynamics model [12].

$$I_d = -\mathfrak{w}_{-1}\left(-\frac{W}{eA}\right) \qquad \text{(index of difficulty)}$$

- $\mathfrak{w}_{-1}$ is the lower real branch of the Lambert W function [7].
- e is Euler's number.
- A is movement amplitude: the path length from the start- to end-point of an individual movement.
- W is error tolerance: the sample standard deviation of the movement end-points relative to the centroid (calculated separately for 9 participants × 3 segments = 27 different samples).



Linear Mixed Effects Model

- The fixed factors were *foot position* (levels: foot-initial, foot-final) and *syllable position* (levels: onset, coda).
- The random factors were *participant* and *articulator* (levels: lips, tongue tip, and tongue dorsum).
- Linear mixed effects model included all possible interactions of fixed effects and included random intercepts and slopes for *index of difficulty* for each *participant-articulator* combination.

Results

Slope of Fitts' law is steeper at prosodic boundaries than inside the prosodic domain

Statistical finding Factors *syllable position* and *foot position* interact with *index of difficulty* ($F(1, 4589.4) = 14.77$, $p = 1.23 \times 10^{-4}$).

Implication *Syllable position* and *foot position* interacted to determine the steepness of the slope for Fitts' law.

Details At the edge of a prosodic domain, a 1 bit increase in difficulty results on average in a 32 ms increase in movement time. Inside the prosodic domain, a 1 bit increase in difficulty resulted on average in a 27 ms increase in movement time.

Slope of Fitts' law is steeper for onset than for coda position

Statistical finding Factor *syllable position* interacts with *index of difficulty* ($F(1, 4534.1) = 69.88$, $p = 8.28 \times 10^{-17}$).

Implication The slope for *index of difficulty* was steeper for movements in syllable onset position than for movements in syllable coda position.

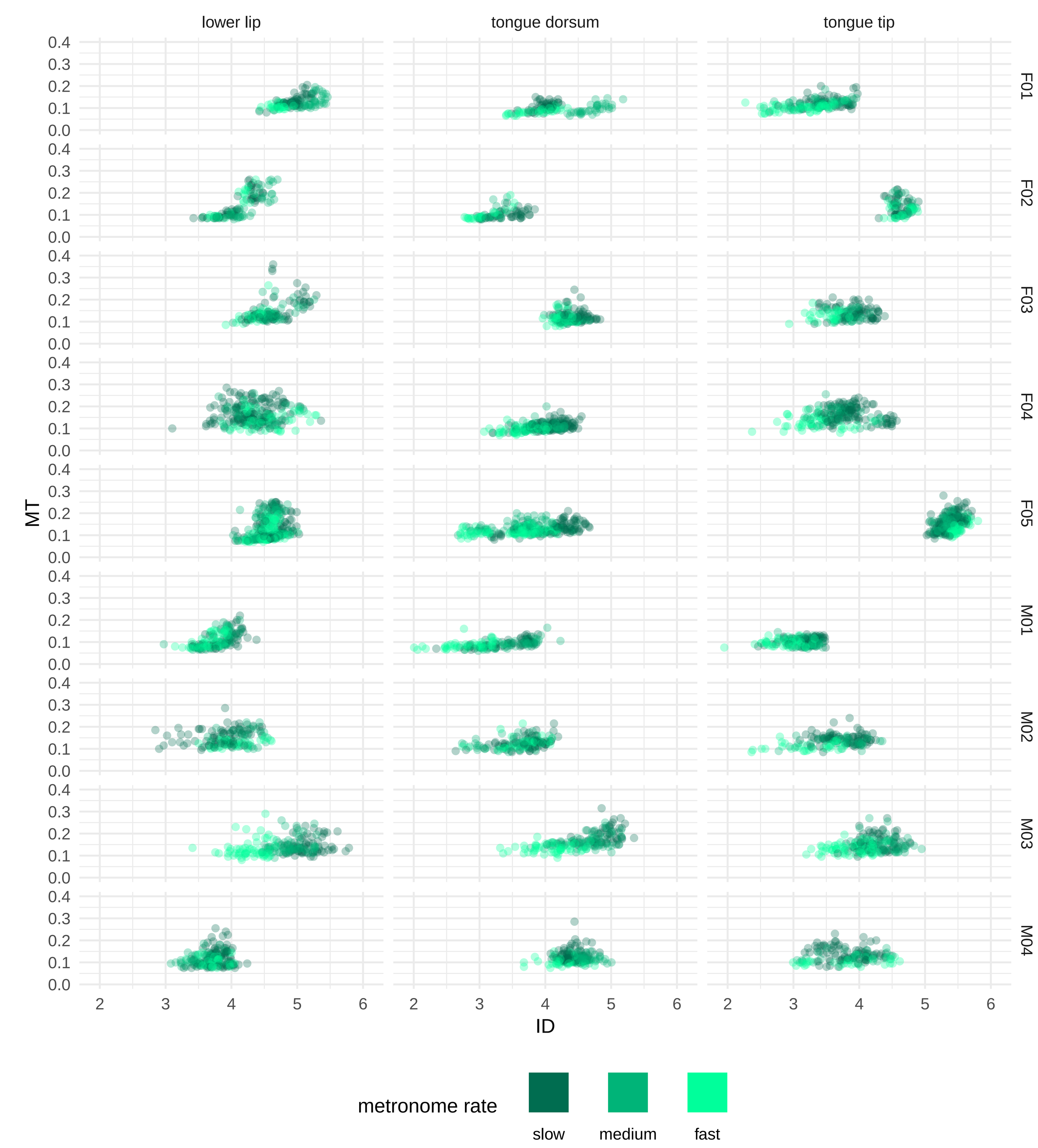
Details The slope was steeper for movements in syllable onset position than for movements in syllable coda position. Foot-initial syllable onsets had a steeper slope than foot-initial syllable codas ($z = 8.83$, $p = 1.04 \times 10^{-18}$), and foot-final syllable onsets had a steeper slope than foot-final syllable codas ($z = 3.66$, $p = 2.50 \times 10^{-4}$).

Fitts' law holds in all cells of the experiment

Statistical finding Fitts' law applied to constriction movements all cells of the experiment.

Implication Fitts' law applied to broadly to movements elicited in the experiment.

Details The slope of Fitts' law differed significantly from zero for foot-initial syllable onsets ($z = 7.88$; $p = 1.7 \times 10^{-15}$), foot-final syllable onsets ($z = 7.03$, $p = 1.1 \times 10^{-12}$), foot-initial syllable codas ($z = 4.49$, $p = 3.6 \times 10^{-6}$), and foot-final syllable codas ($z = 5.56$, $p = 1.3 \times 10^{-8}$).



Significance

Confirming our hypothesis indicates a connection between linguistic prosody and the information throughput:

> Information throughput is low at the edge of a prosodic domain (32 bits/s on average) and high inside the prosodic domain (39 bits/s on average).

This finding provides an information theoretic basis for linguistic prosody:

> The consequence of rhythmically grouping supra-segmental linguistic units into prosodic domains is to modulate information throughput such that time intervals of low information throughput separate time intervals of high information throughput.

Acknowledgments

The authors acknowledge funding through National Institutes of Health (NIH) Grant Nos. R01DC007124 and T32DC009975, and National Science Foundation (NSF) Grant No. 1514544. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or NSF.

References

[1] M. E. Beckman and J. Edwards. "Intonational categories and the articulatory control of duration". In: *Speech perception, production, and linguistic structure*. Ed. by Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka. 1st ed. Tokyo, Japan: Ohmsha, Ltd, 1992, pp. 359–375. [2] C. P. Browman and L. Goldstein. "Articulatory gestures as phonological units". In: *Phonology* 6.2 (1989), pp. 201–251. [3] C. P. Browman and L. M. Goldstein. "Towards an articulatory phonology". In: *Phonology Yearbook* 3 (1986), pp. 219–252. [4] D. Byrd and E. L. Saltzman. "Intragestural dynamics of multiple prosodic boundaries". In: *Journal of Phonetics* 26.2 (1998), pp. 173–199. [5] D. Byrd and E. L. Saltzman. "The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening". In: *Journal of Phonetics* 31.2 (2003), pp. 149–180. [6] D. Byrd et al. "Phrasal signatures in articulation". In: *Papers in Laboratory Phonology V*. Ed. by M. Broe and J. Pierrehumbert. Cambridge, UK: Citeseer, 2000, pp. 70–87. [7] R. M. Corless et al. "On the Lambert W function". In: *Advances in Computational Mathematics* 5.1 (1996), pp. 329–359. [8] J. Edwards, M. E. Beckman, and J. Fletcher. "The articulatory kinematics of final lengthening". In: *the Journal of the Acoustical Society of America* 89.1 (1991), pp. 369–382. [9] P. M. Fitts. "The information capacity of the human motor system in controlling the amplitude of movement". In: *Journal of Experimental Psychology* 47.6 (1954), pp. 381–391. [10] S. R. Kuberski and A. I. Gafos. "Fitts' law in tongue movements of repetitive speech". In: *Phonetica* (2019). [11] A. C. Lammert et al. "Speed-accuracy tradeoffs in human speech production". In: *PloS One* 13.9 (2018), e0202180. [12] E. L. Saltzman and K. G. Munhall. "A dynamical approach to gestural patterning in speech production". In: *Ecological Psychology* 1.4 (1989), pp. 333–382. [13] T. Sorensen et al. *Derivation of Fitts' law from the Task Dynamics model of speech production*. arXiv:2001.05044. 2020. arXiv: `2001.05044 [q-bio.NC]`. [14] M. Tiede, C. Mooshammer, and L. Goldstein. "Noggin Nodding: Head Movement Correlates With Increased Effort in Accelerating Speech Production Tasks". In: *Frontiers in Psychology* 10 (2019), p. 2459.