The process of initiating speech and the search for good analysis tools

Pertti Palo

10 Mar 2025

Outline

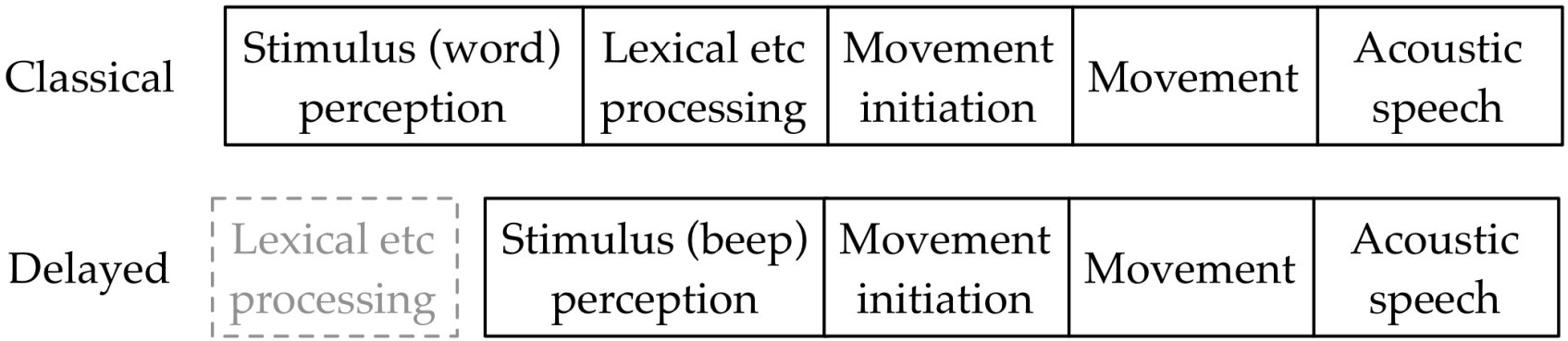
- Introductions - Me and the topic
- Some naming experiments
- First tool: Pixel Difference (PD)
- Second tool: Mean Squared Error (MSE)
- Conclusion

Who's this Pertti Palo?

- I have a couple of degrees in engineering and a PhD in Phonetics.
- I also have a life in folk music, folk dancing, wandering (hiking and long distance skiing), role-playing games, crafts (knitting, terrain crafting, and other things).
- I grew up in Finland, became me in Scotland and have since worked in Indiana and here in Edmonton.

Introduction

- In my thesis I concentrated on timing of utterance onset in both acoustics and articulation (Palo 2019).
- The data was high-speed tongue ultrasound from a delayed naming experiment – specifically one using the Rastle instructions (Rastle et al. 2005).



Demo of classical naming

- Let's try some versions of naming experiments.
- After the slide changes read out loud the word on the next slide as soon as you can.

red

Demo of classical naming

- Let's try that a second time.

green

Demo of delayed naming

- After the slide changes wait for me to snap my fingers.
- After you hear the finger snap, read the word out loud as soon as you can.

orange

Demo of delayed naming

- Rinse and repeat

purple

Demo of delayed naming with Rastle instructions

- After the slide changes wait **at rest** for me to snap my fingers.
- After you hear the finger snap, read the word out loud as soon as you can.

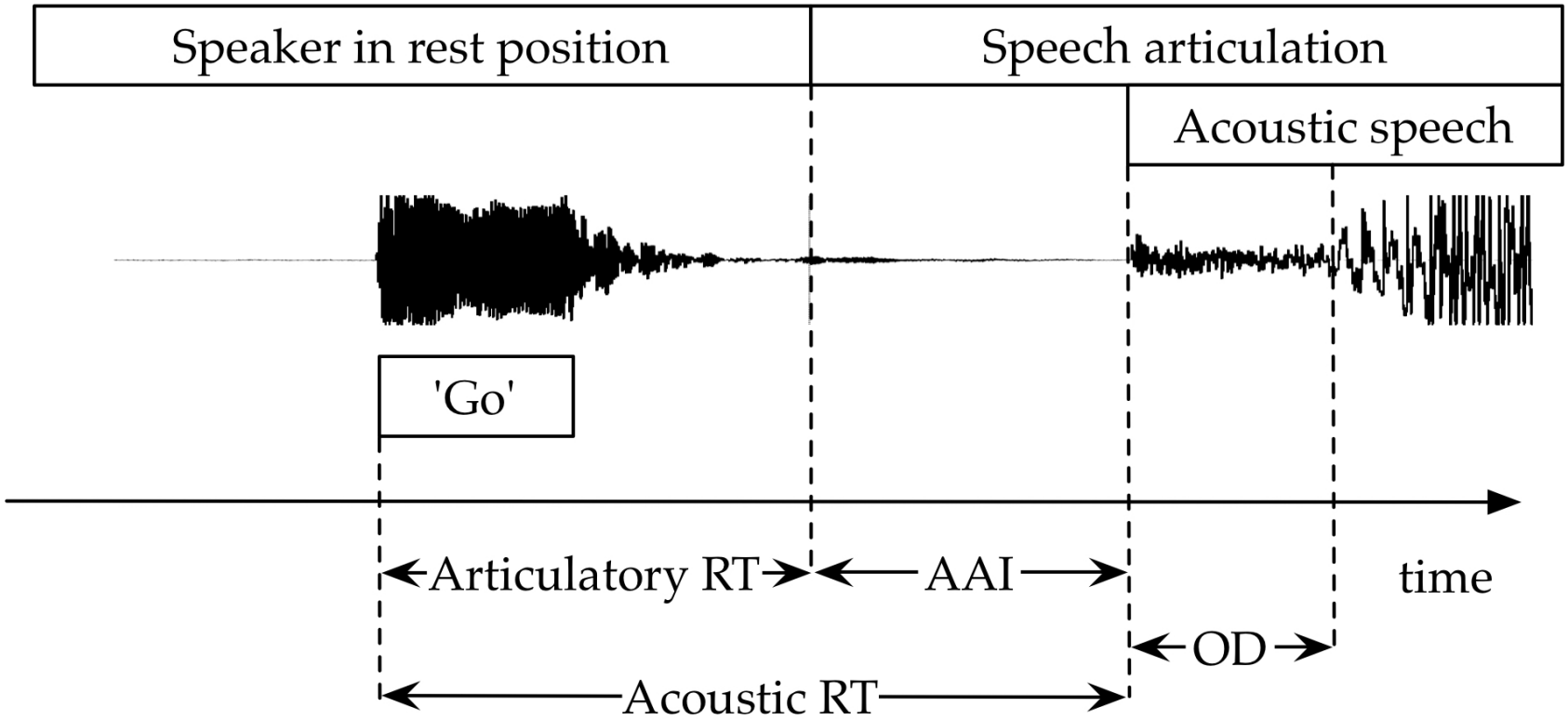
yellow

Demo of delayed naming with Rastle instructions

- Once more

blue

What happened in the last one?



- AAI = Articulatory onset to Acoustic onset Interval
- OD = Onset (or obstruent) Duration

Introduction

- Now imagine that final experiment was recorded with tongue ultrasound imaging.
- And your job was to find the articulatory onset in the resulting greyscale videos.
- Here, let's try it. [external slide set coming up]

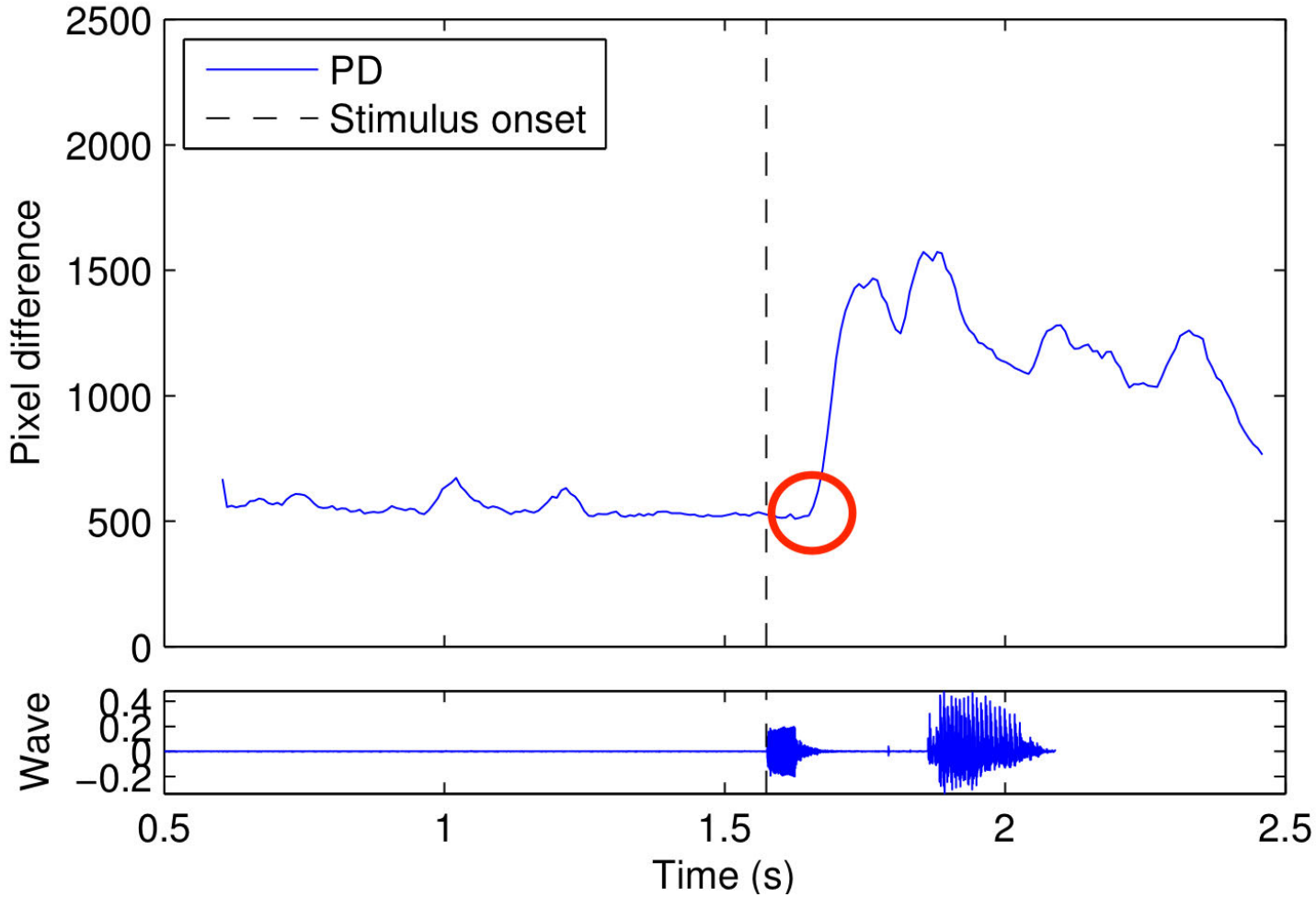
Introduction

- When trying to identify movement onset in greyscale videos with a lot of speckle 'noise', it doesn't take long to grow a desire for an easier way.
- The speckle 'noise' maybe caused by a number of factors including bubbles in the acoustic gel between the chin and the probe, noise sources in the equipment, and more interestingly changes in internal structures of tissues – such as muscle fibres tensing and relaxing.

Pixel Difference (PD)

Pixel Difference (PD)

- The first tool out of the box happened to work adequately – and so for my thesis I used Euclidean distance or *l*2-norm to identify articulatory onsets.

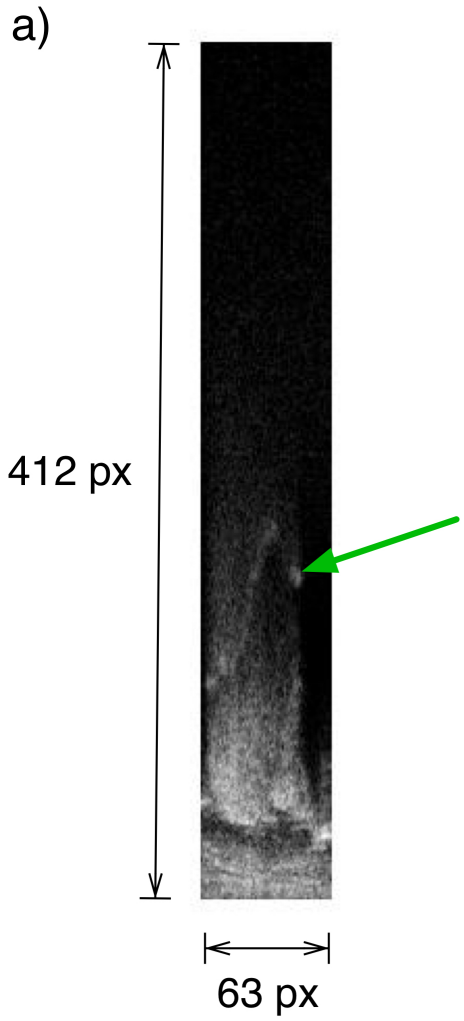


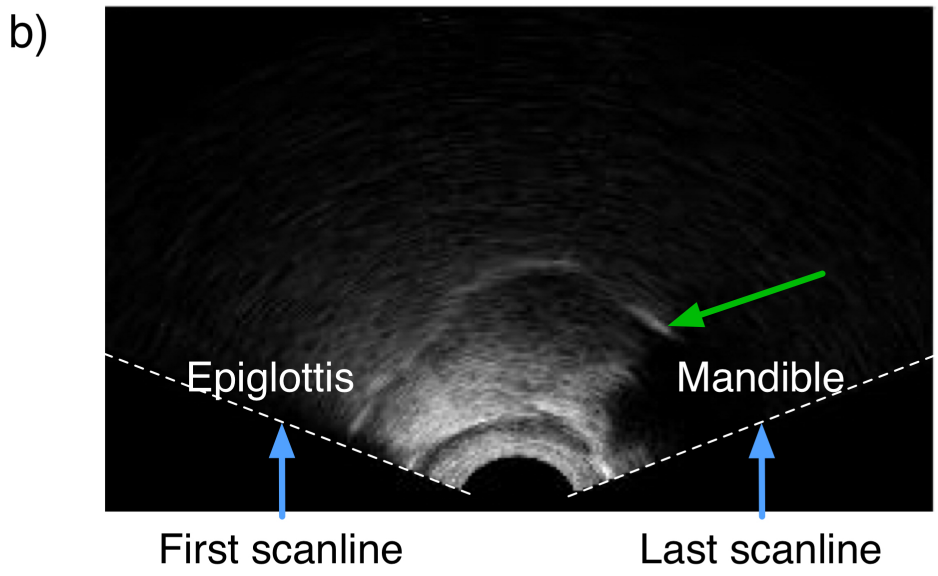
Pixel Difference (PD): Background

- The analysis methods presented here are similar to methods developed by
  - McMillan and Corley (2010) and Drake et al. (2013) who used Euclidean distance on ultrasound frames and
  - Raeesy et al. (2011) who used a similar method on MRI data.
- The way I have used it, it is actually just the Pythagorean theorem applied in a space with a lot more dimensions than 2.

Pixel Difference (PD): Raw vs. Interpolated

- PD is usually calculated on
  - (a) uninterpolated (probe-return) ultrasound data instead of
  - (b) interpolated (human-readable) data.



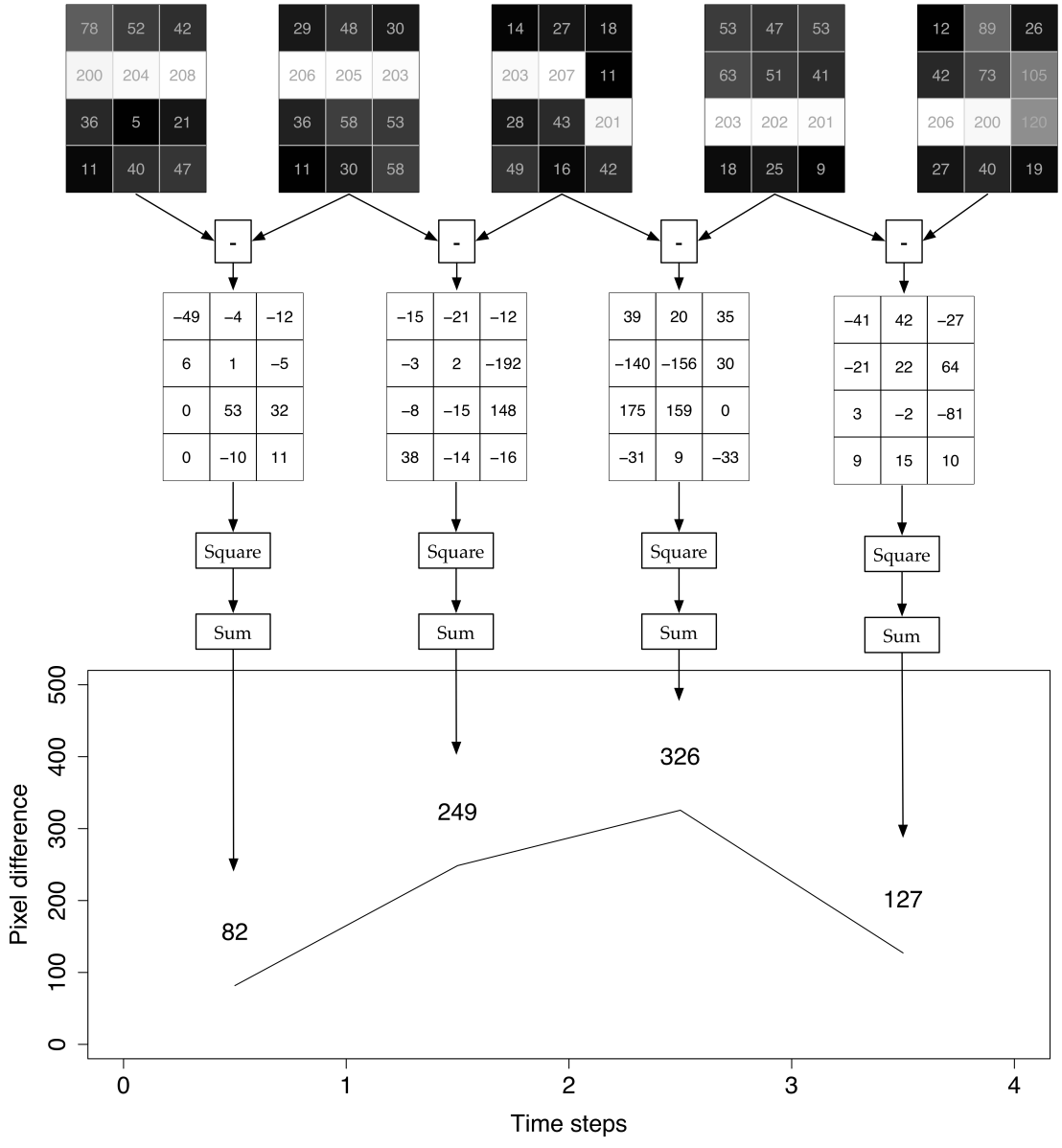


Pixel Difference (PD): The maths

$$I2(t+0.5) = \sqrt{\sum_{i,j}(x(i,j,t+1) - x(i,j,t))^2}$$

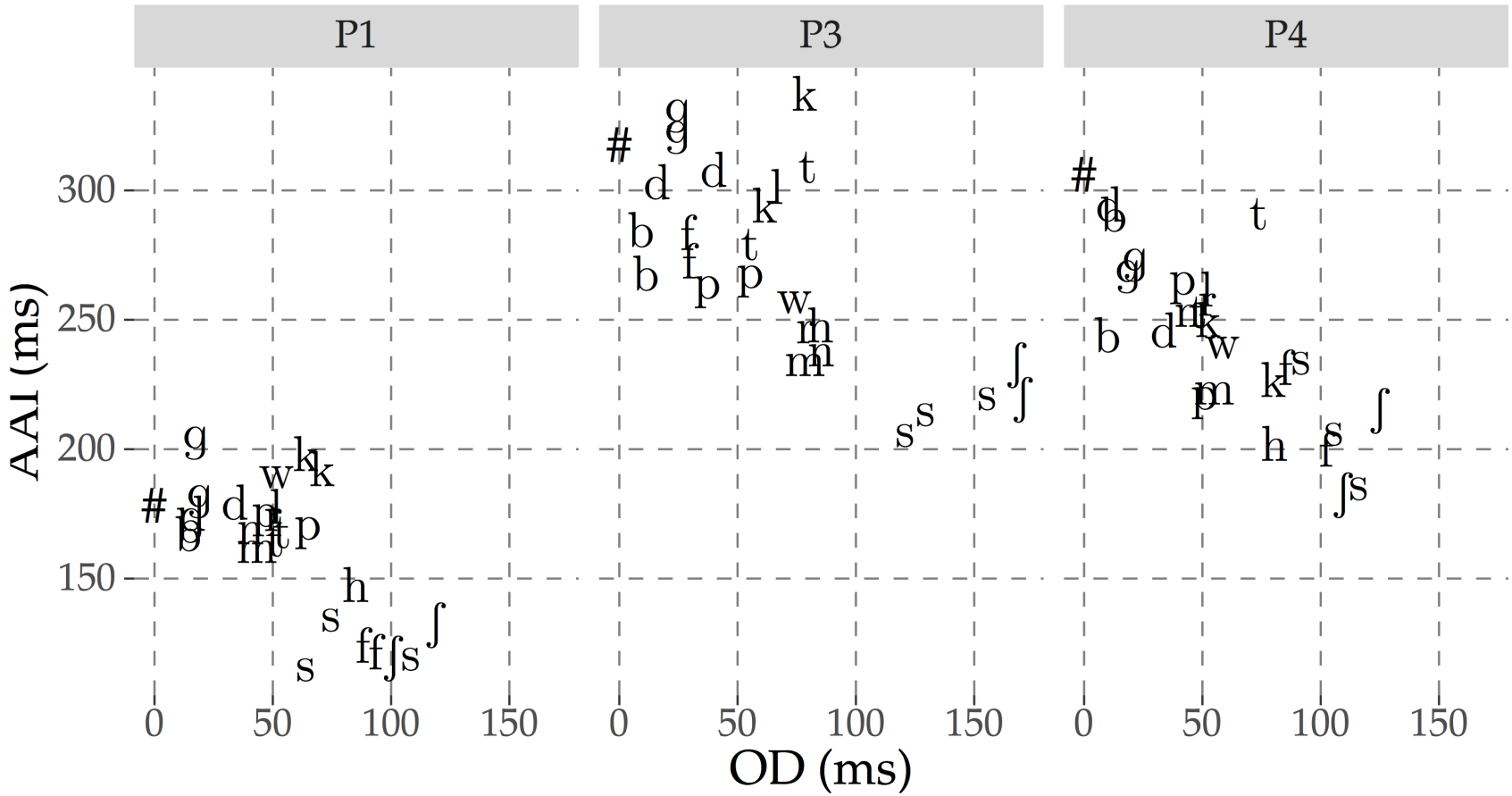
- i and j are indices that span the width and height of the image, t is the time index.
- Like said, this is actually just the Pythagorean theorem applied in a space with a lot more dimensions than 2.

Pixel Difference (PD): The maths visually



Developments with PD

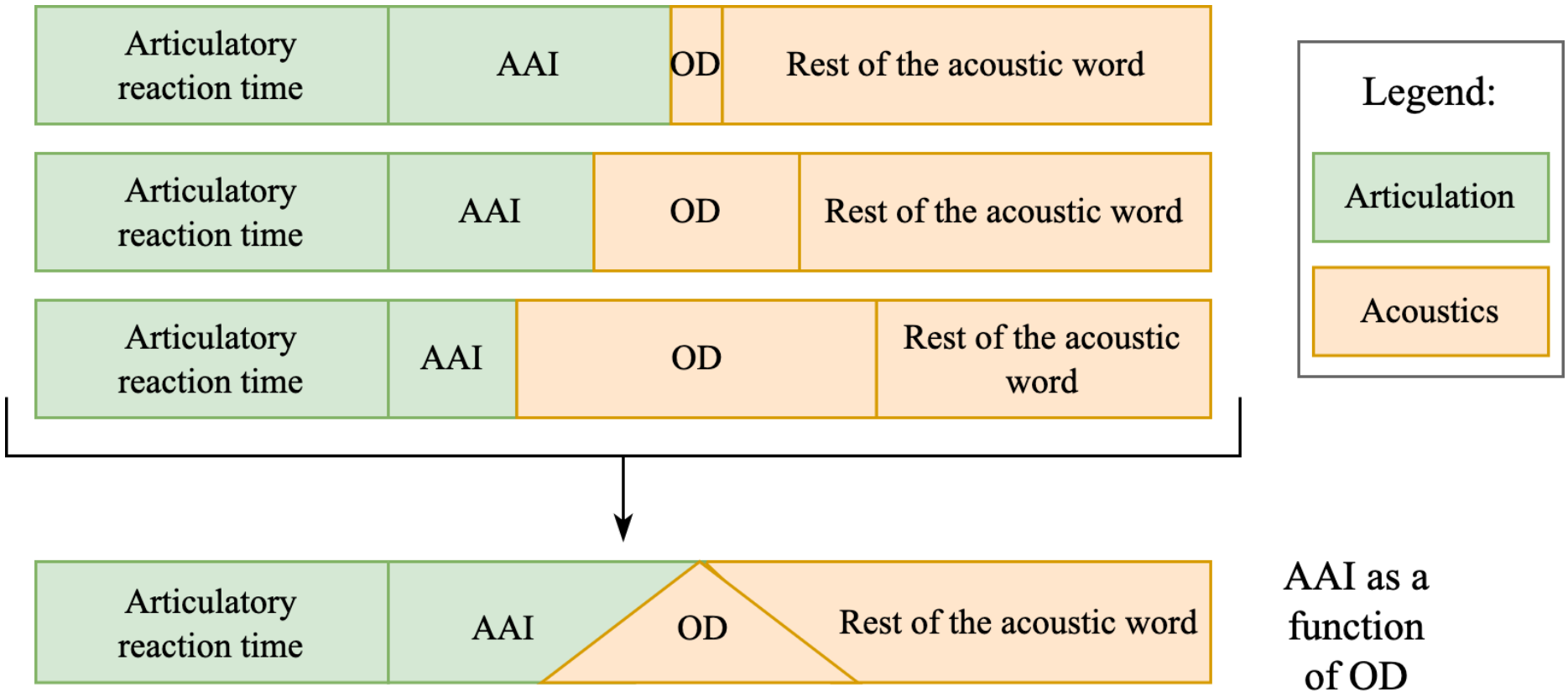
Delayed naming results: Articulatory to Acoustic Interval



Medianised within participant, over several repetitions and over the vowels /a,i,ɔ/. Over all analysable n = 1386: 439 from P1, 672 from P3, and 275 from P4.

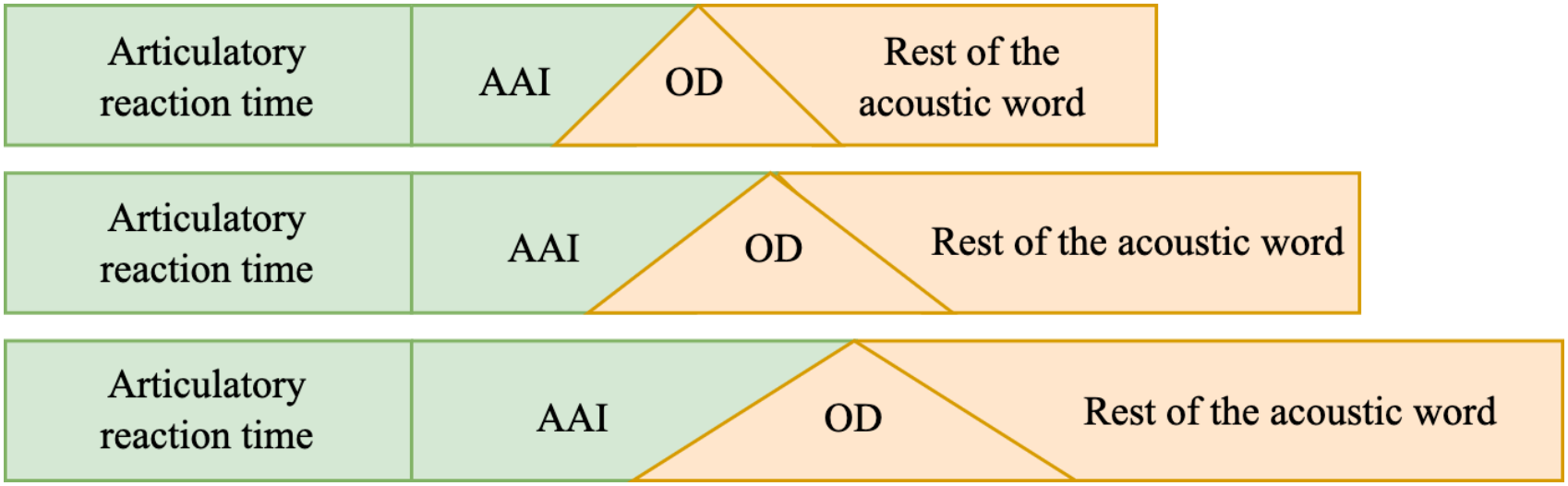
Theory: Effect of OD on AAI

- As the Onset Duration (OD) gets longer, Articulatory to Acoustic Interval (AAI) shortens.
- First three lines represent individual utterances, final line is a conceptual model of the effect of continuously lengthening OD.6



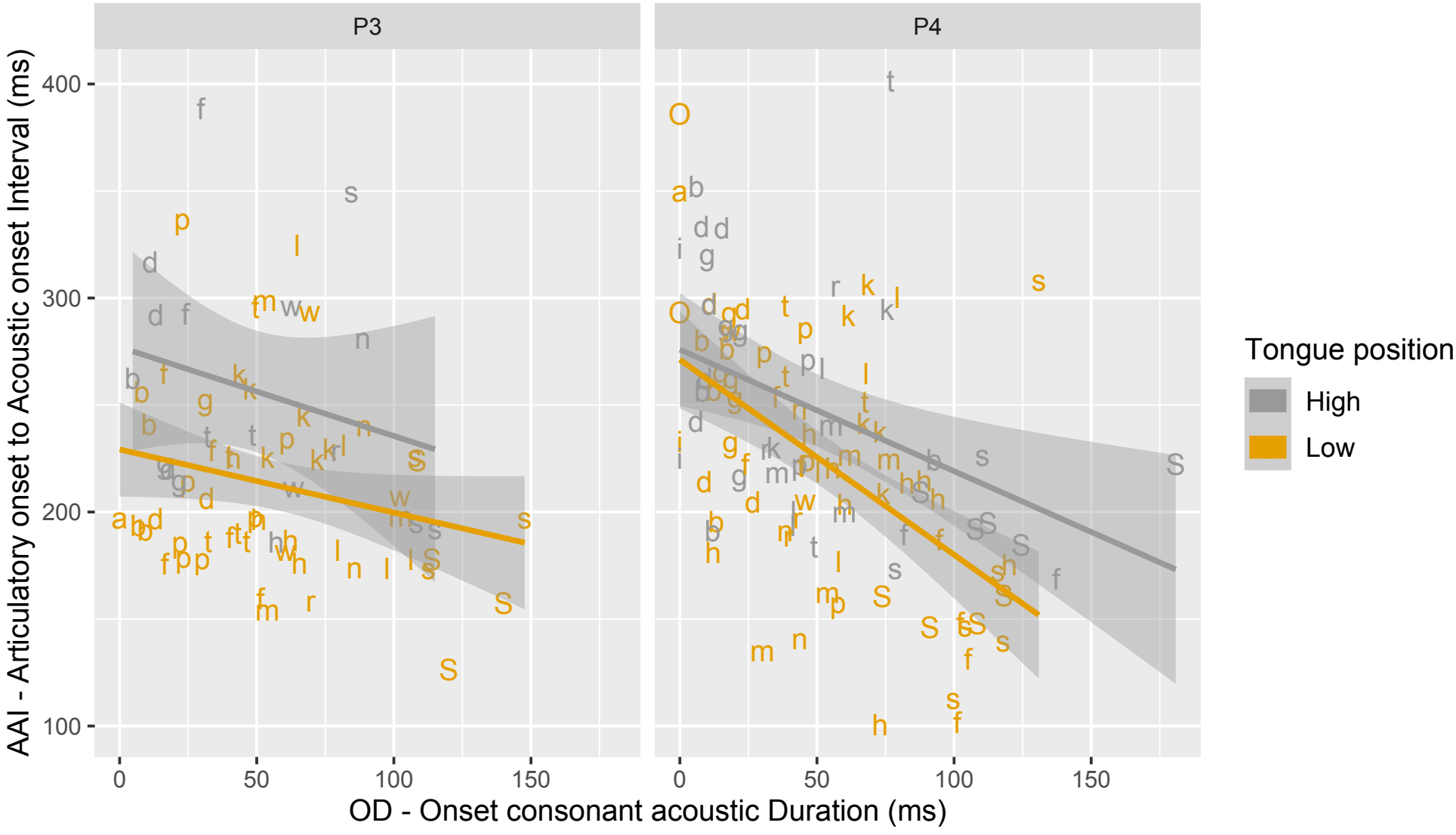
Theory: Effect of articulatory rate on AAI

- If we keep the utterance content constant but vary articulation rate, all parts (AAI, OD, and acoustic word) get longer as articulation rate goes down.



Starting position

- 'Remain at rest' does not define what 'rest' means.



References

Drake, E., Schaeffler, S., and Corley, M. (2013). ARTICULATORY EVIDENCE FOR THE INVOLVEMENT OF THE SPEECH PRODUCTION SYSTEM IN THE GENERATION OF PREDICTIONS DURING COMPREHENSION. In *Architectures and Mechanisms for Language Processing (AMLaP)*, Marseille.

McMillan, C. T. and Corley, M. (2010). Cascading influences on the production of speech: Evidence from articulation. *Cognition*, 117(3):243–260.

Palo, P. (2019). *Measuring Pre-Speech Articulation*. PhD thesis, Queen Margaret University, Edinburgh.

Raeesy, Z., Baghai-Ravary, L., and Coleman, J. (2011). Parametrising Degree of Articulator Movement from Dynamic MRI Data. In *12th Interspeech*, pages 2853–2856.

Rastle, K., Harrington, J. M., Croot, K. P., and Coltheart, M. (2005). Characterizing the Motor Execution Stage of Speech Production: Consonantal Effects on Delayed Naming

Something else i.e. section title