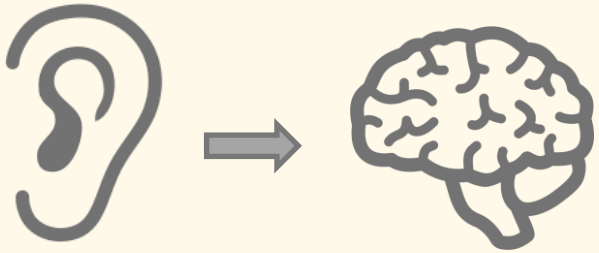
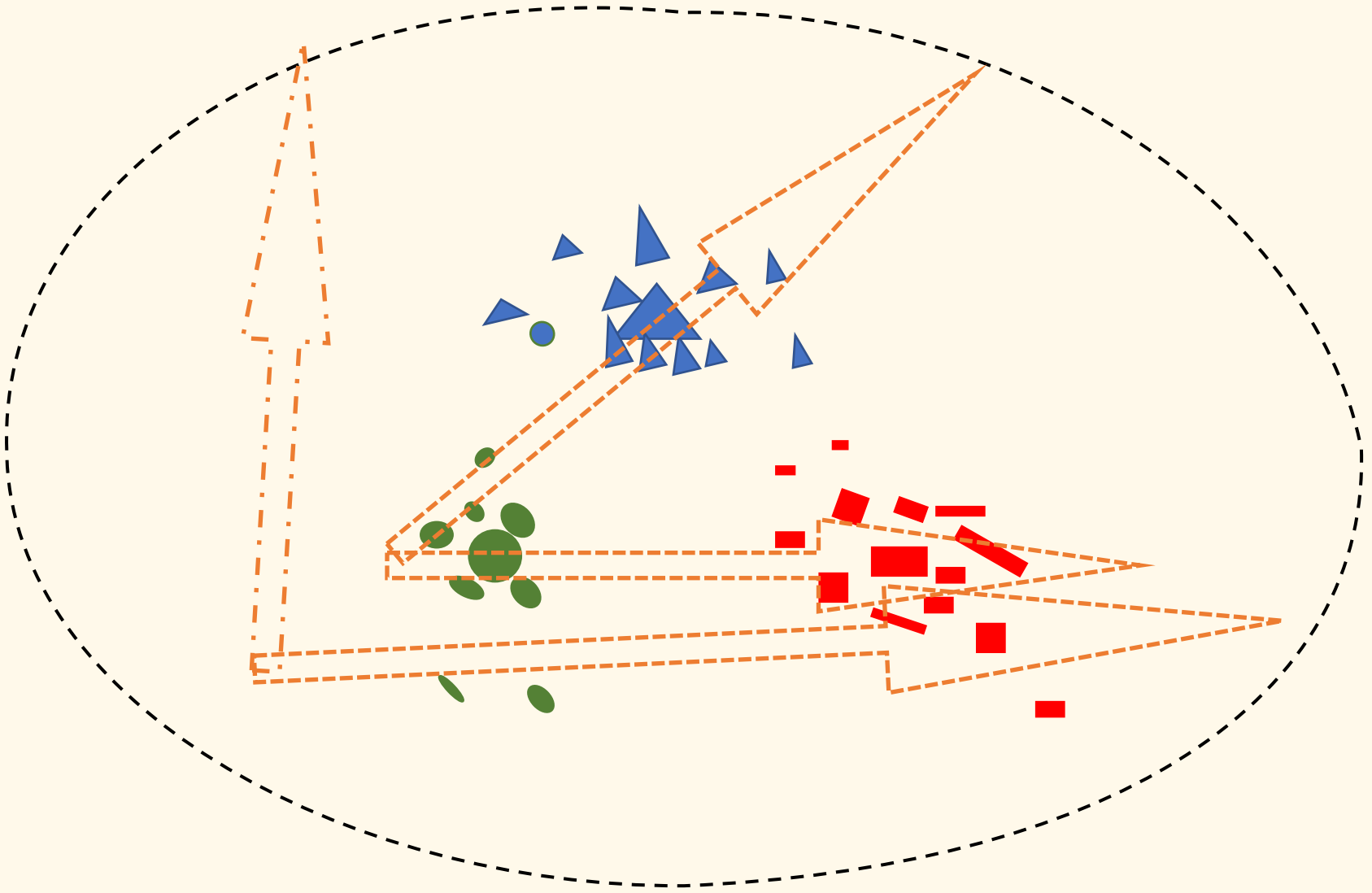
Incremental cue training: a study of lexical tone learning by non-tonal listeners

Yanyu Li, Laurence White & Ghada Khattab

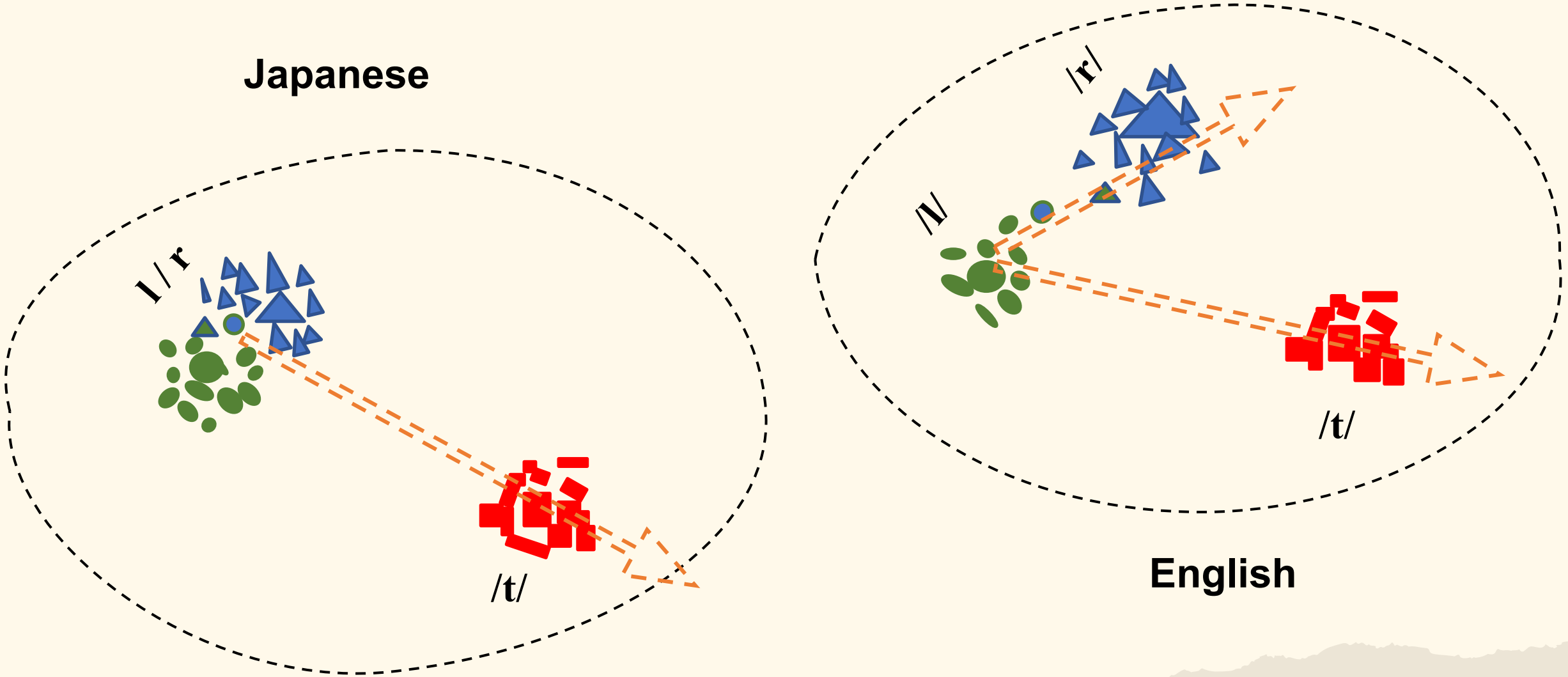
Newcastle University

y.li218@newcastle.ac.uk | laurence.white@newcastle.ac.uk | ghada.khattab@newcastle.ac.uk



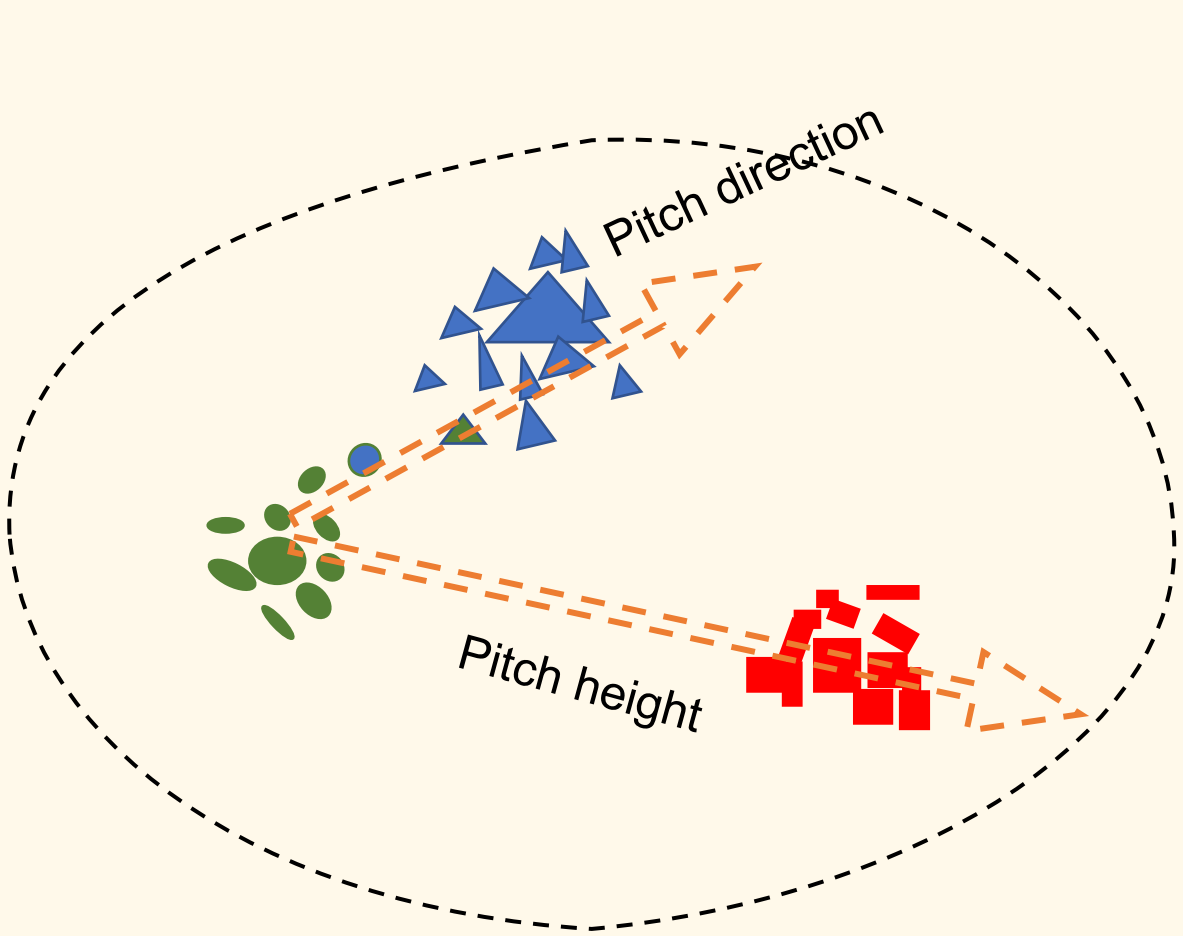


Iverson and Kuhl (1995)

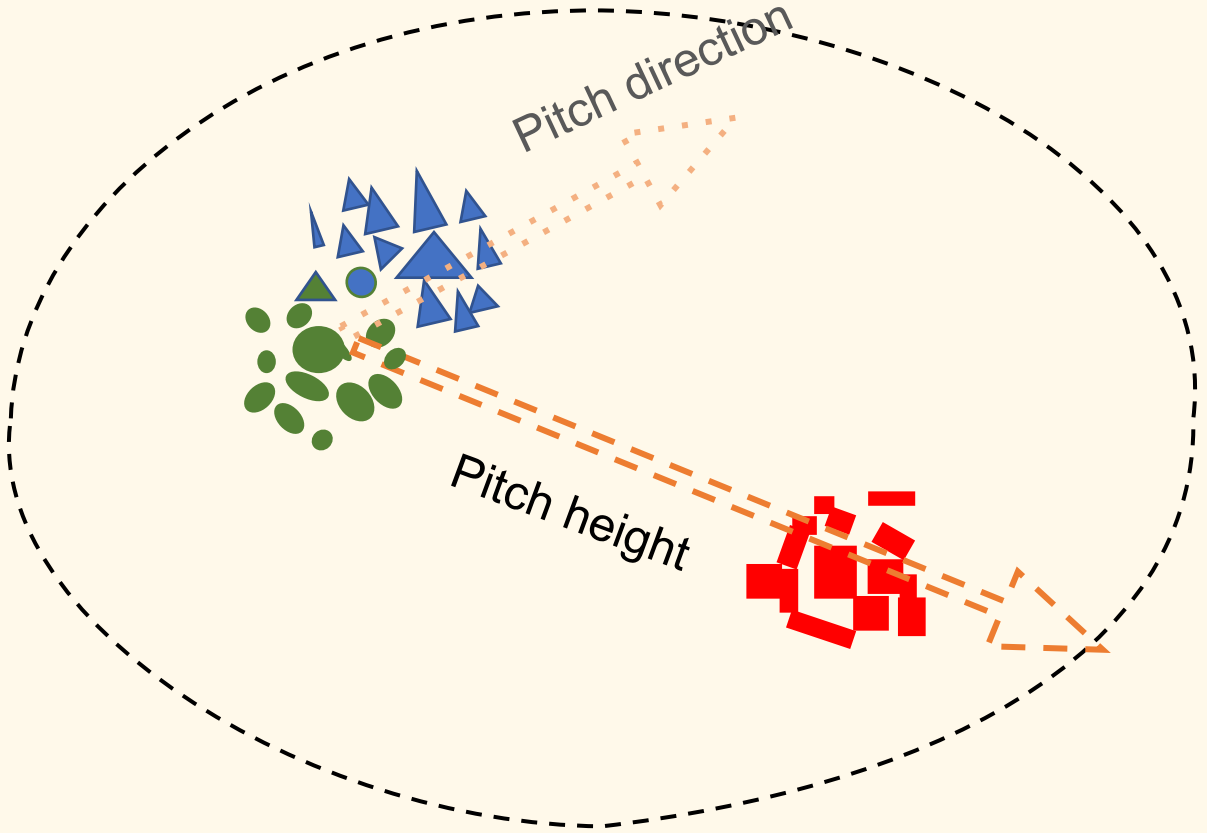


Iverson and Kuhl (1995)

Lexical tone?



Tonal listeners



Non-tonal listeners

Gandour, 1983; Chandrasekaran et al. 2010

Incremental training

Training 1
(**Most** exaggerated)

→

Training 2
(**Less** exaggerated)

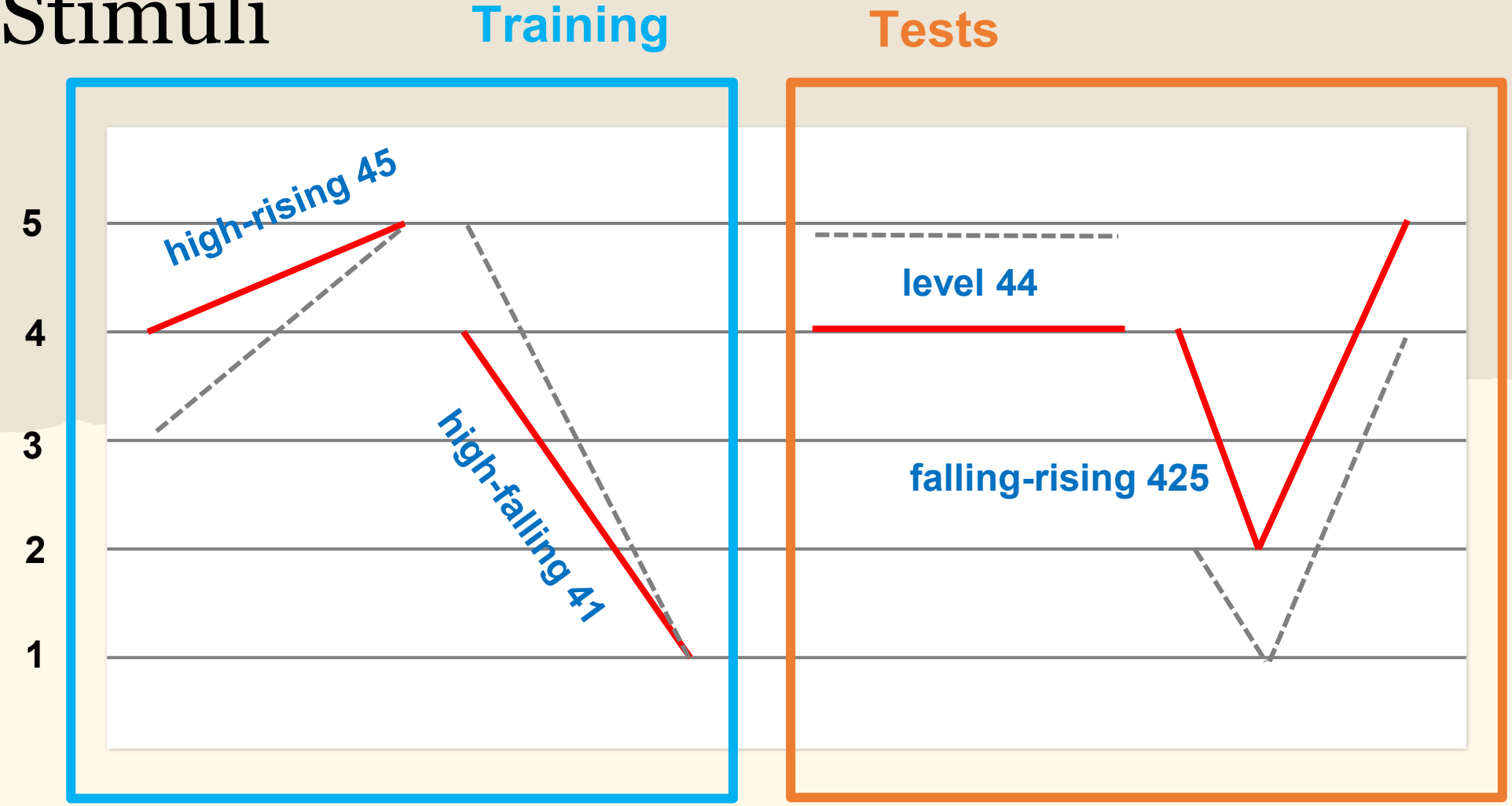
→

Training 3
(**No** exaggeration)

Is Incremental training method **effective** on non-tonal listeners?

Does **Incremental** group differ from the baseline (**Fixed**) group?

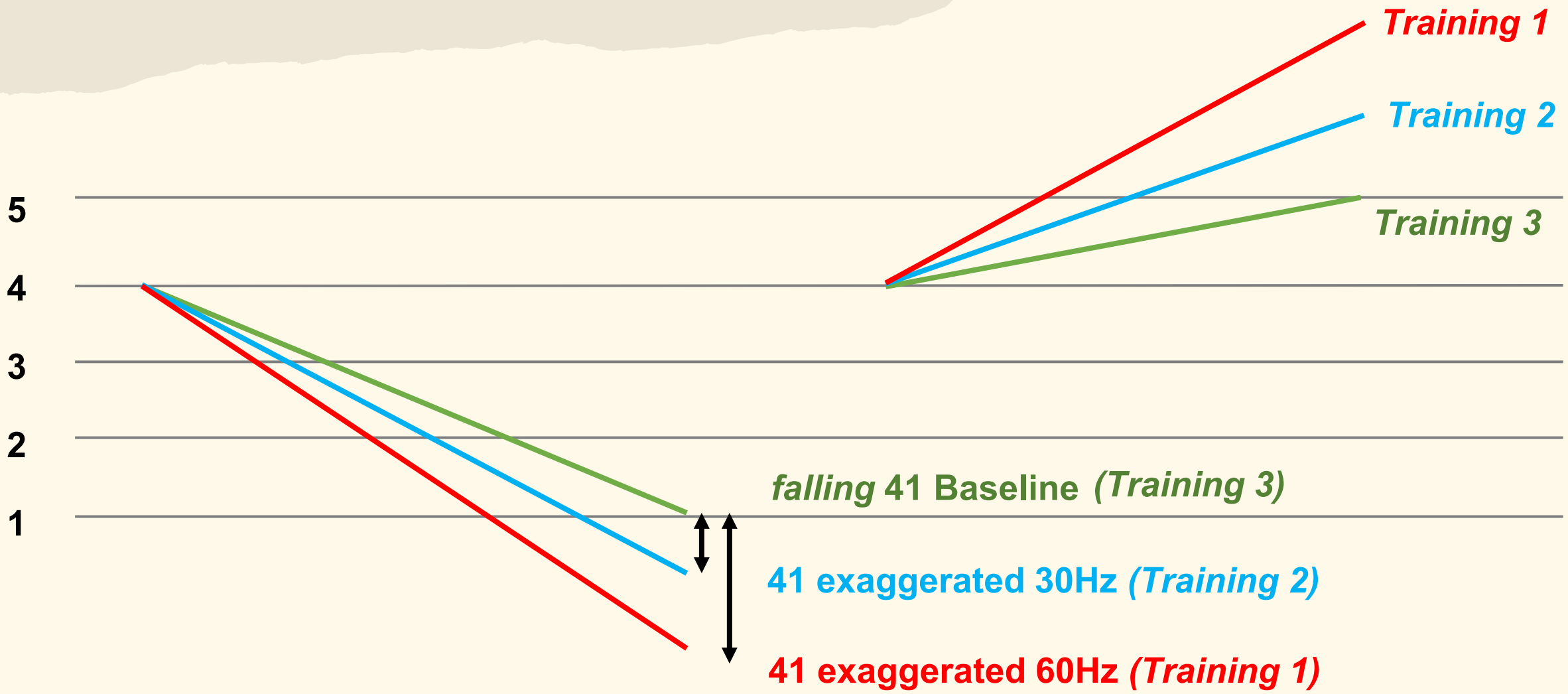
Stimuli



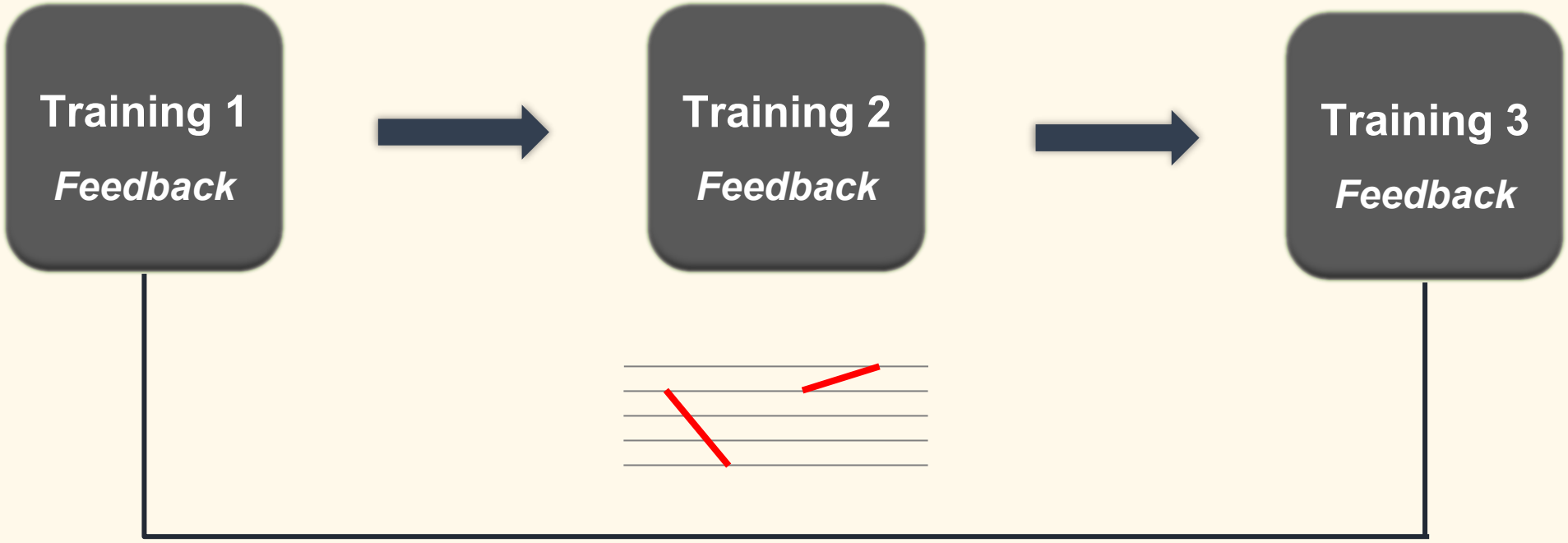
9 Syllables: /pa:/, /ta:/, /ka:/, /pi:/, /ti:/, /ki:/, /pu:/, /tu:/, and /ku:/

Chao (1930)

Exaggerated stimuli



ABX tasks



Exaggeration 60Hz

Exaggeration 30Hz

Baseline

Baseline

Baseline

Baseline

Pre-test
~~Feedback~~

Training 1
Feedback

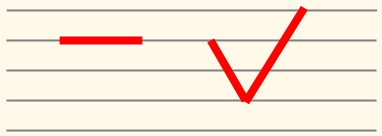
Training 2
Feedback

Training 3
Feedback

Post-test
~~Feedback~~

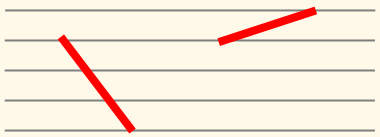
44 vs 425

level vs. falling-rising

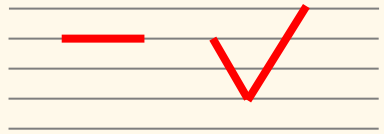


41 vs 45

high-falling vs high-rising



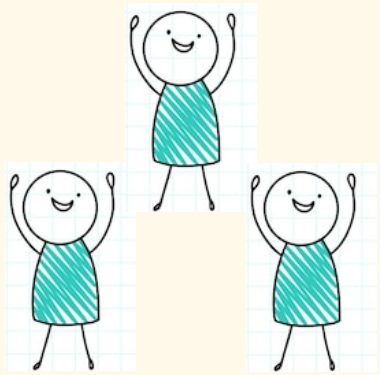
Same as pre-test



Participants

Incremental group: **N=40**

*With **exaggerated** stimuli*



Fixed group: **N=46**

*With **baseline** stimuli*

Native English speakers without lexical tone experience.

Data Analysis

- **Accuracy** data
- **Reaction time** data
- **Separated** Test & Training blocks

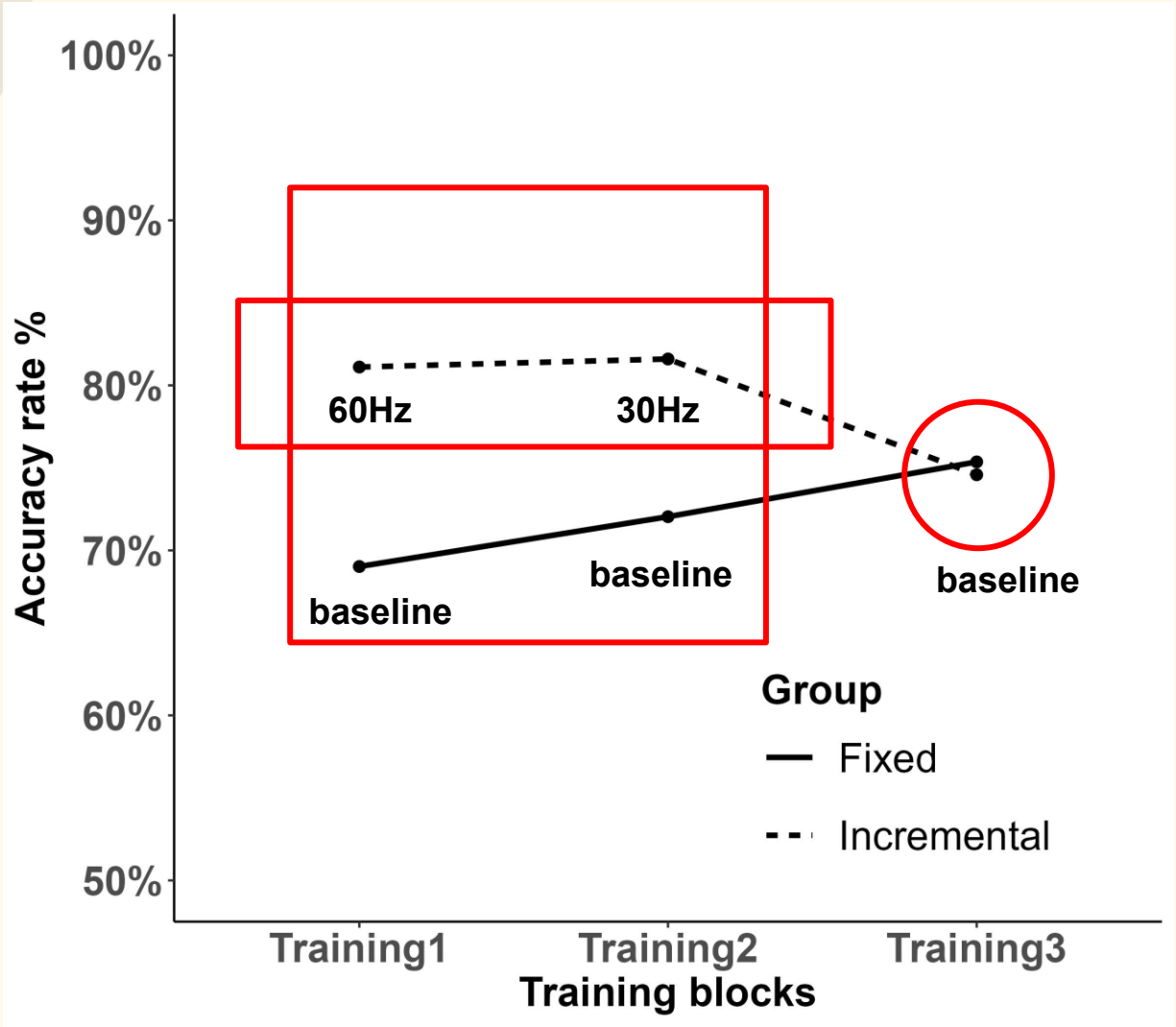
Accuracy data

***Logistic** mixed-effect regression models*

- **Fixed effects:**
  - Group: Incremental & Fixed
  - Test: Pre-test & Post-test
  - Training: Training1, Training2, & Training3
- **Random effects:**
  - Random intercepts: participant, syllable
  - Block-by-participant random slope

Accuracy training blocks

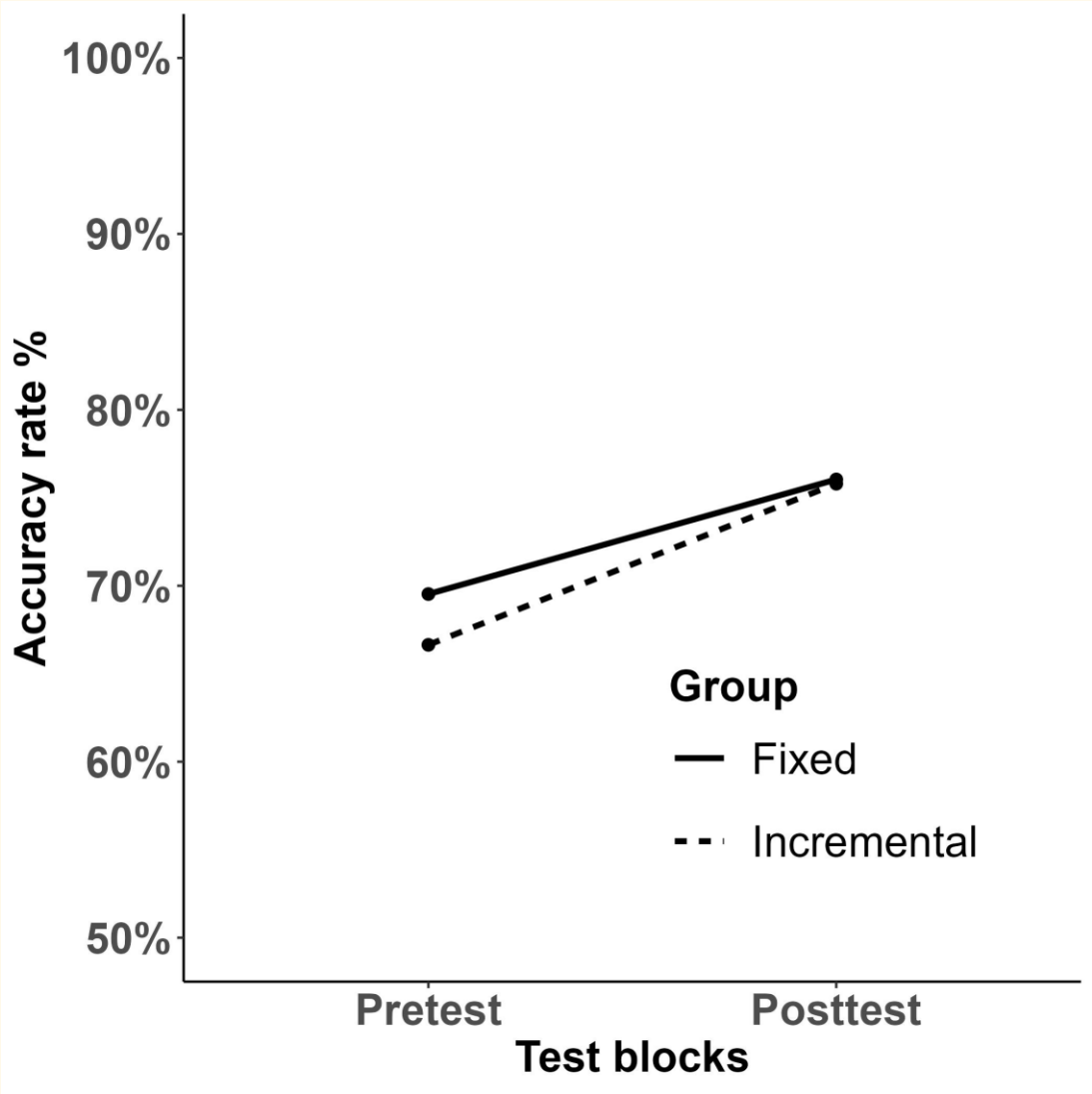
- **Group**: $\chi 2(1) = 9.54$, p = .002
- **Block**: $\chi 2(2) = 10.56$, p = .005
- **Group x Block**: $\chi 2(2) = 20.10$, p < .001



*glmer(accuracy ~ group * block + (1 + block | participant))*

Accuracy test blocks

- **Block**: $\chi 2(1) = 31.96$, $p < .001$



glmer(accuracy ~ block + (1 + block | participant))

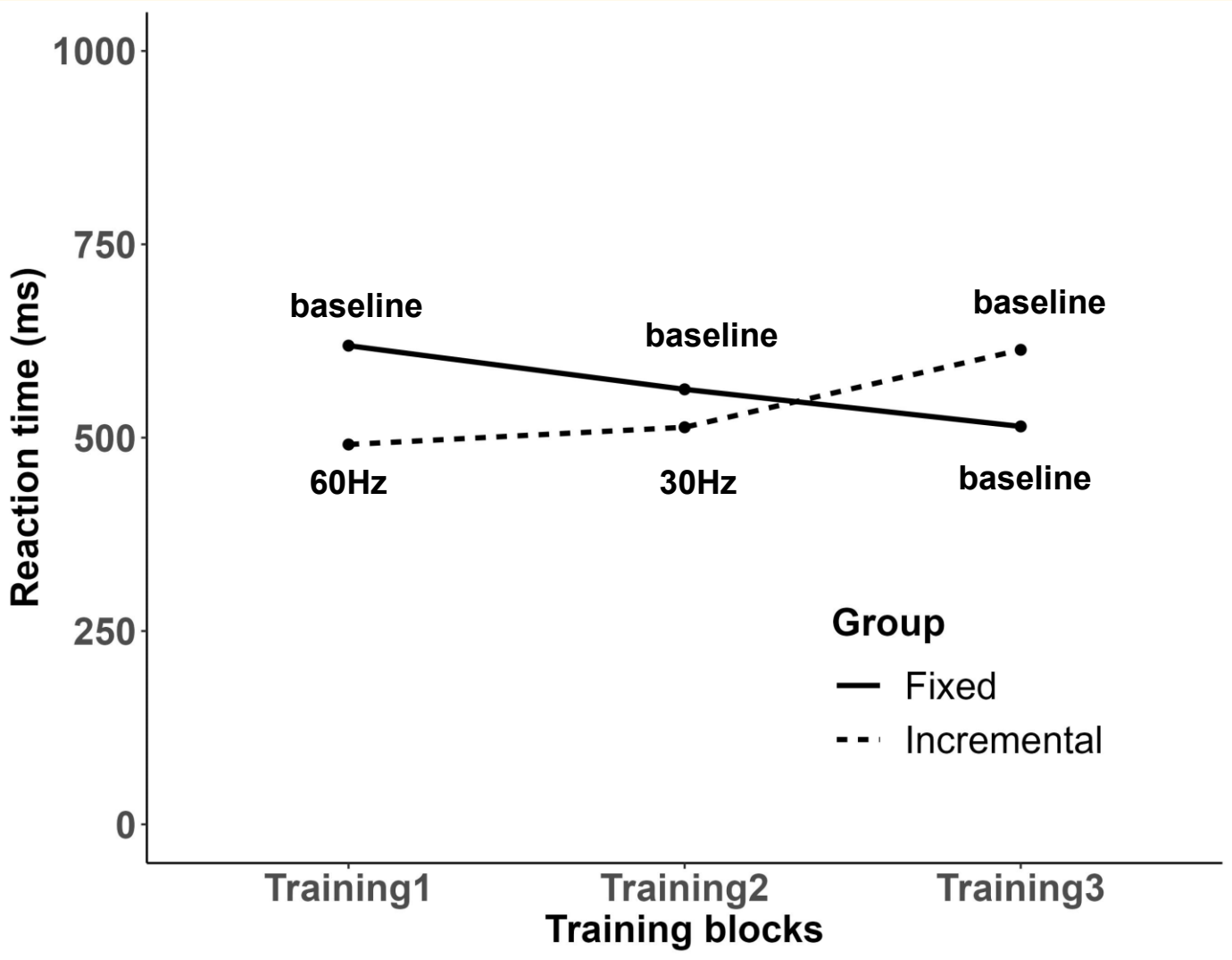
Reaction time (RT) data

Linear *mixed-effect regression models*

- Same model structure as accuracy data

RT training blocks

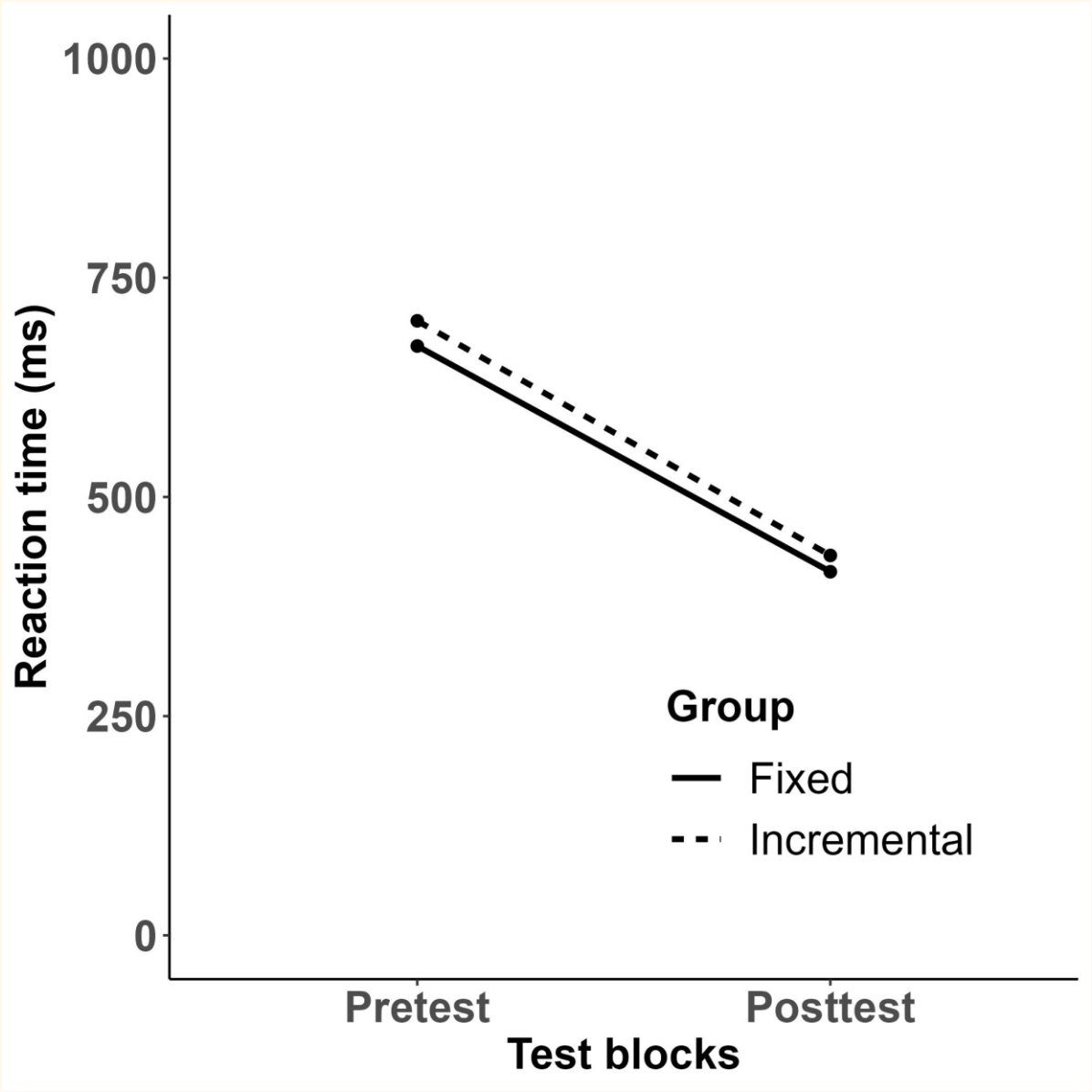
- **Numeric**
  - Block: p = .45
  - Group: p = .71
  - Block x Group: p = .37



RT test blocks

- **Block**: $\chi 2(1) = 53.80$, $p < .001$

lmer(logreact ~ block + (1 + block | participant))



Summary

- Exaggerated stimuli **boosted discrimination** *(Training 1 & 2).*
- The boosting effect **did not sustain** when exaggeration was removed (*Training 3*).
- Participants **improved discrimination** on the **untrained** tonal contrast (*Tests*).

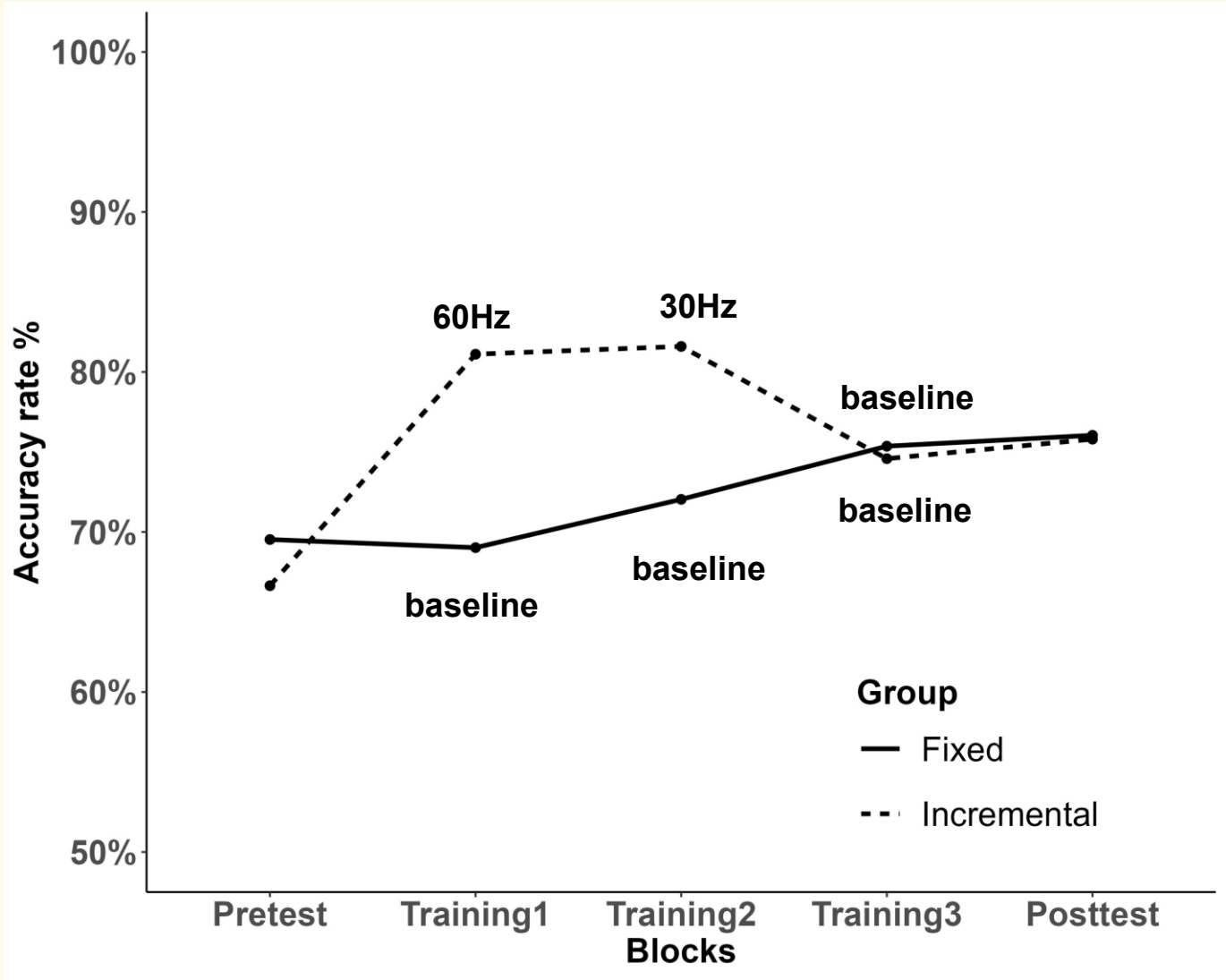
Two follow-ups

1. Reduced exaggeration
2. Trained Mandarin listeners

@yanyuniverseli

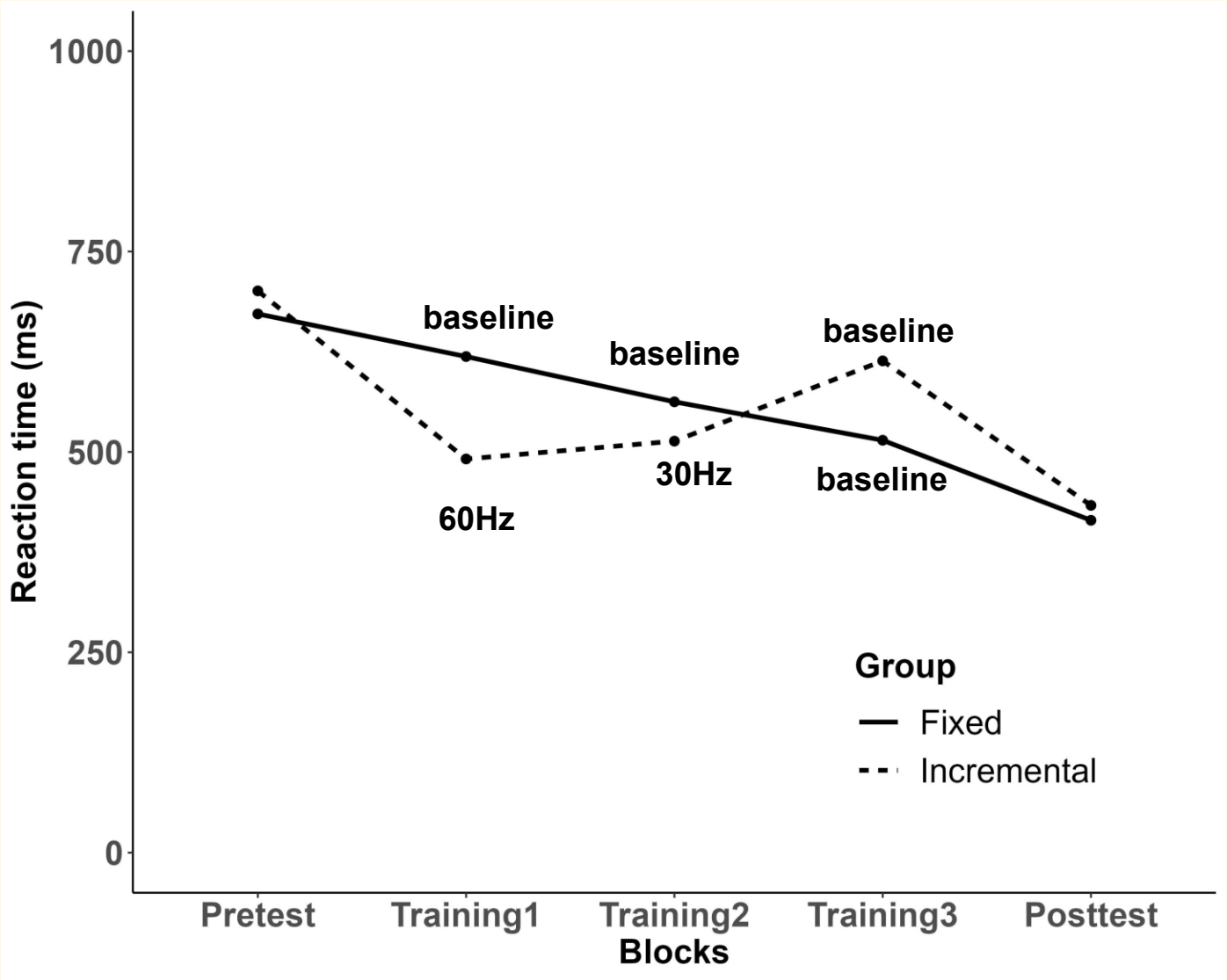
Accuracy overview

- **Pretest vs Training1**
- Block: p = .74
- **Posttest vs Training3**
- Block: p = .50

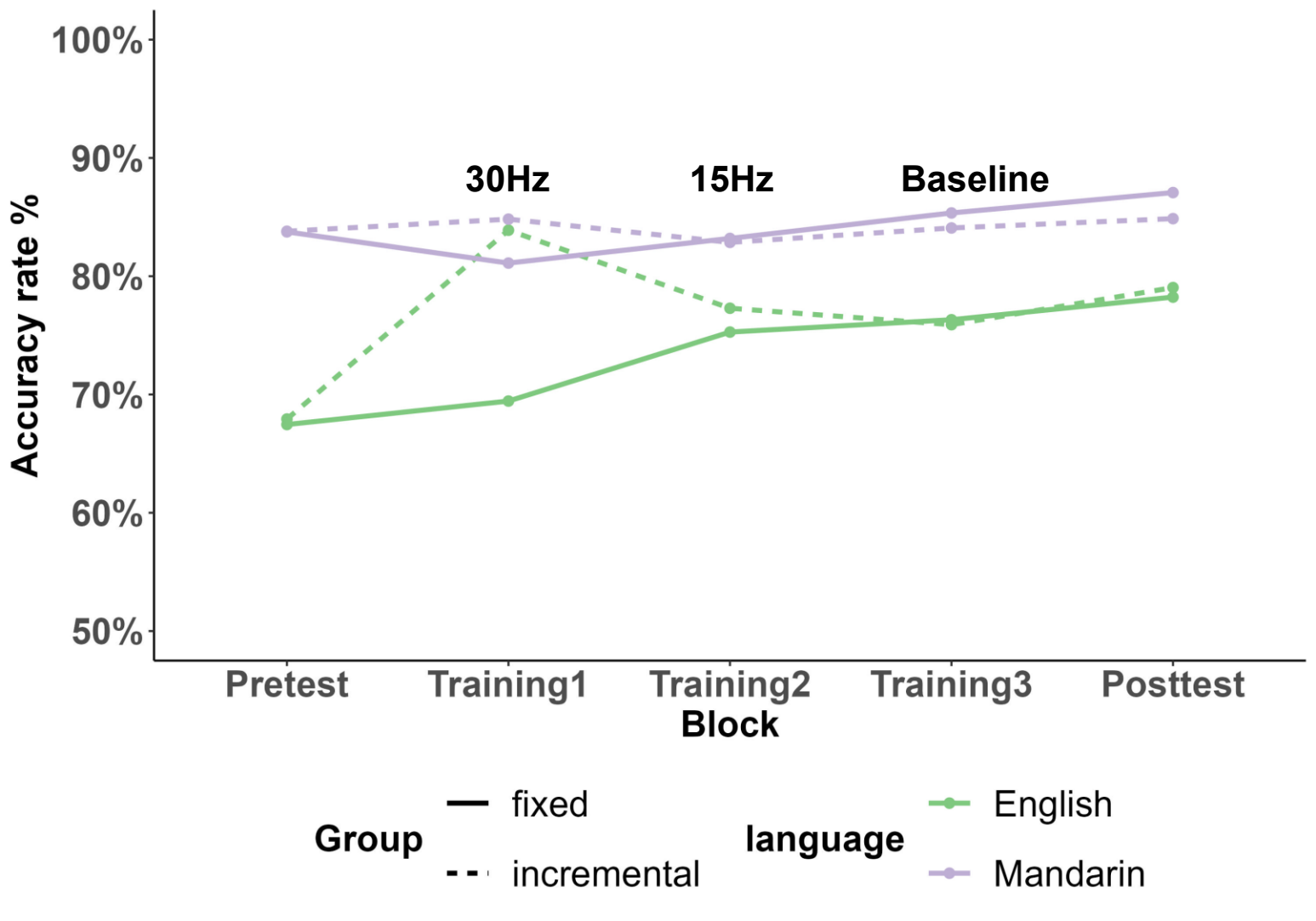


RT overview

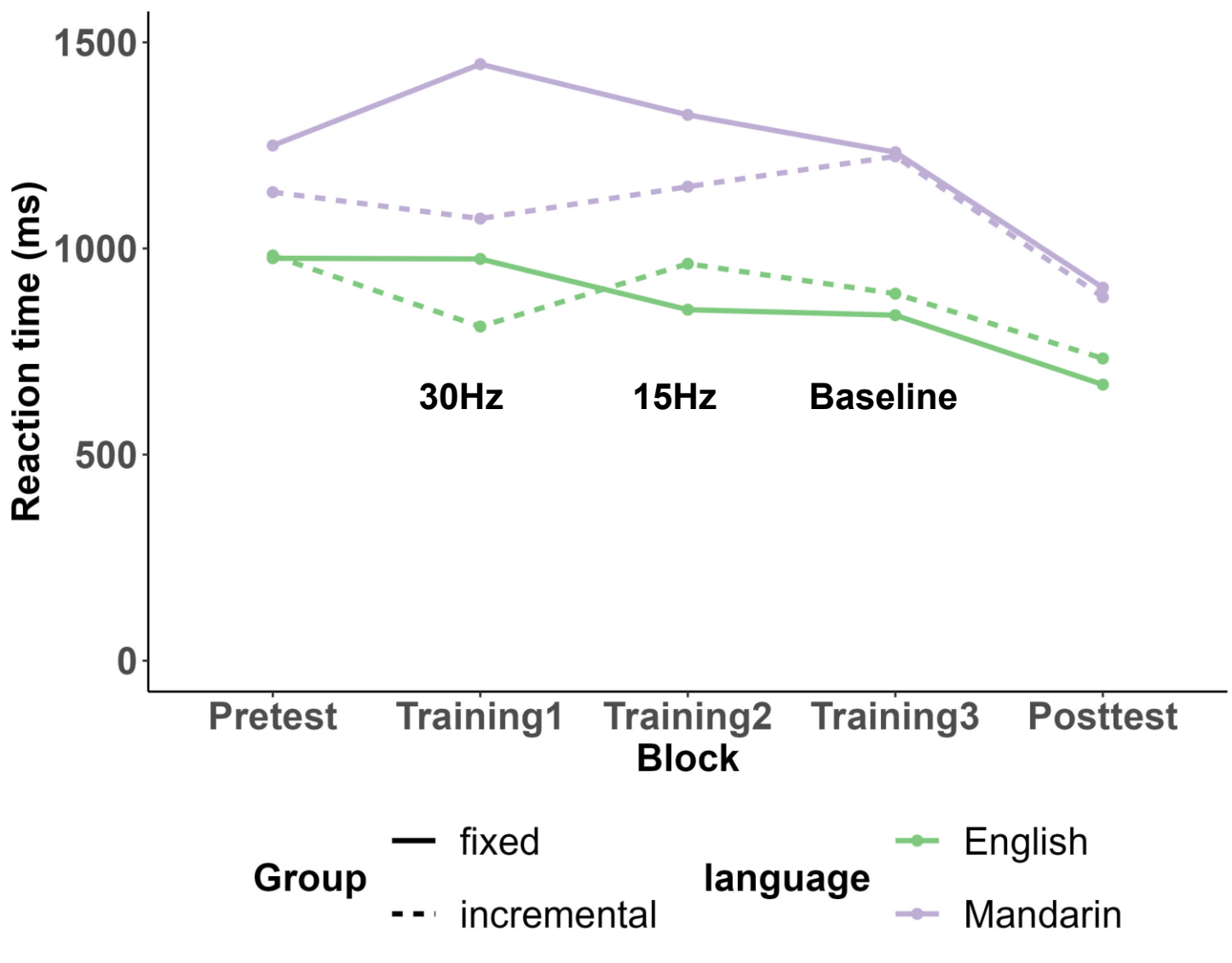
- **Pretest vs Training1**
- Block: p = .11
- **Posttest vs Training3**
- Block: p < .001 *



Follow-up study results (Accuracy)



Follow-up study results (RT)



Stimuli – Recording

- 2 female native Mandarin speakers. Recording done in Mandarin tones.
- E.g., “我读了’爬’字三次”, meaning “I read ‘**/paa/ 35**’ three times”.

我读了“pá”字三次。	我读了“爬”字三次。	我读了“pá”字三次。
我读了“pà”字三次。	我读了“怕”字三次。	我读了“pà”字三次。
我读了“pā”字三次。	我读了“趴”字三次。	我读了“pā”字三次。
我读了“pǎ”字三次。	我读了“pǎ”字三次。	我读了“pǎ”字三次。

Conversion

- 1) convert the tones represented in Chao-digits into Hertz;

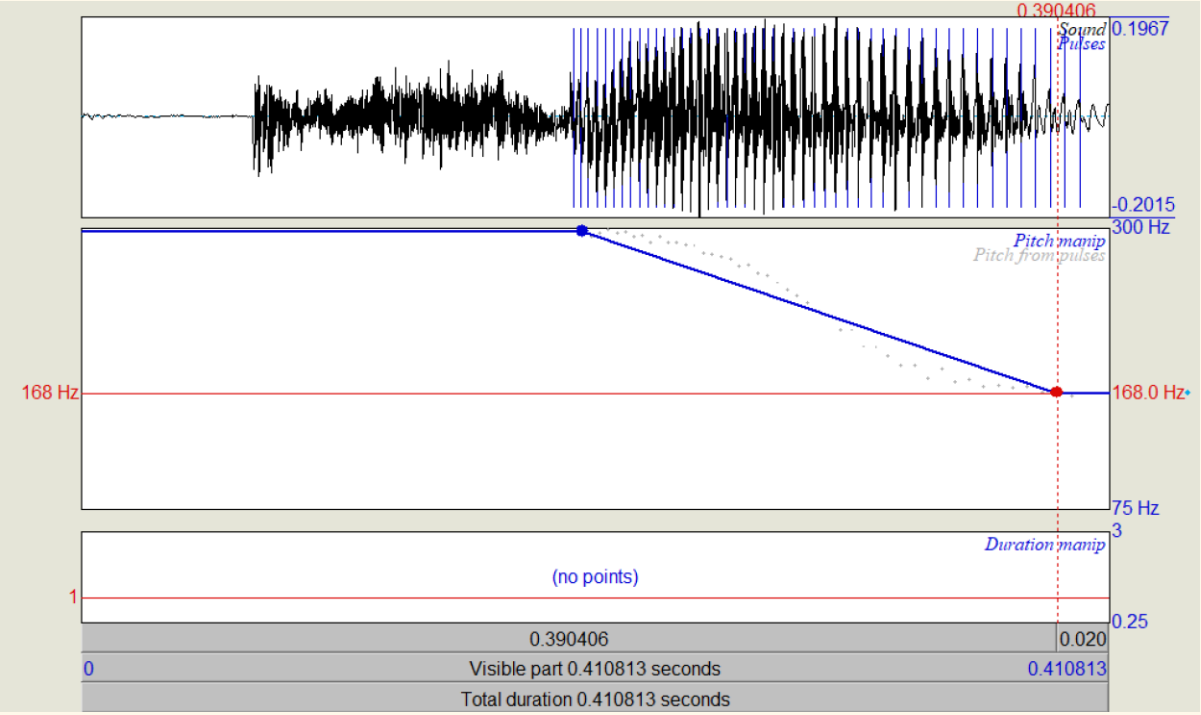
  2) manipulate recordings based on the conversion.
- Formula: T = [(lg X - lg L) / Lg H - Lg L] * 5

Pseudo-tone (Chao digits)	Speaker 1 (Hz)	Speaker 2 (Hz)
High-rising **45**	297-375	294-335
High-falling **41**	297-147	294-201
Level **44**	297-297	294-294
Falling-rising **425**	297-240-375	294-260-335

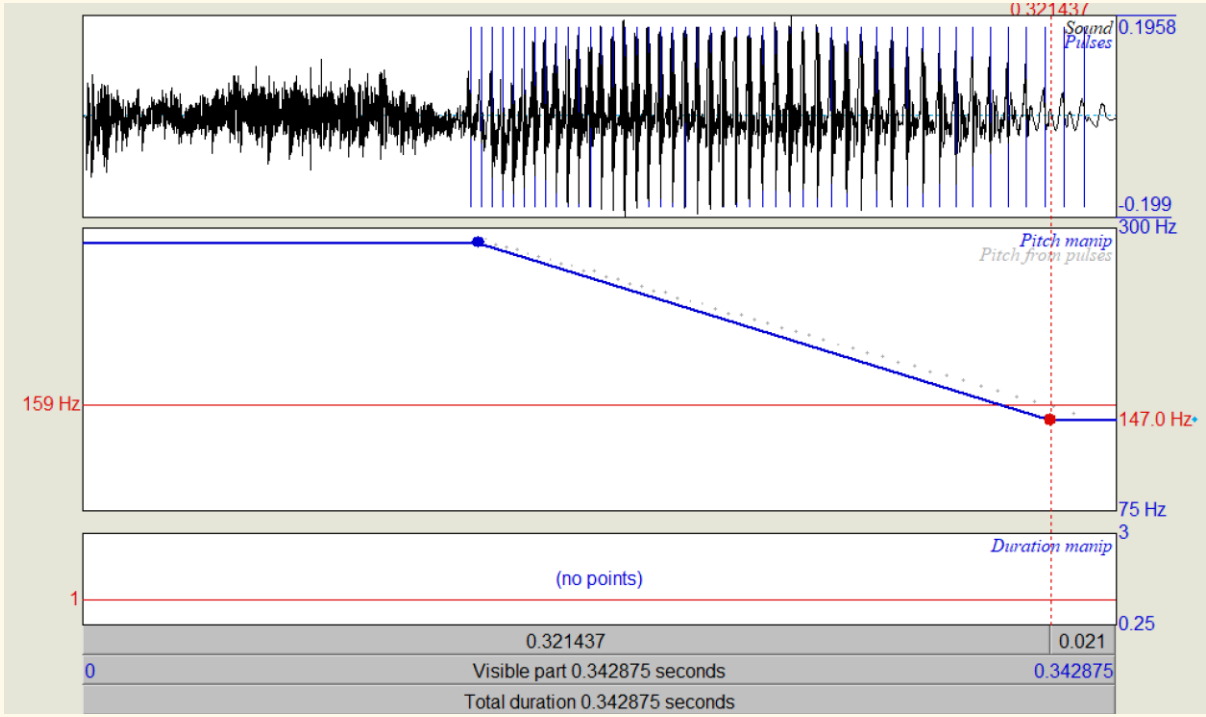
Table 1: F0 values (Hz) of the four pseudo-tones.

Wang et al. (2003)

Resynthesise



Mandarin *falling* 51



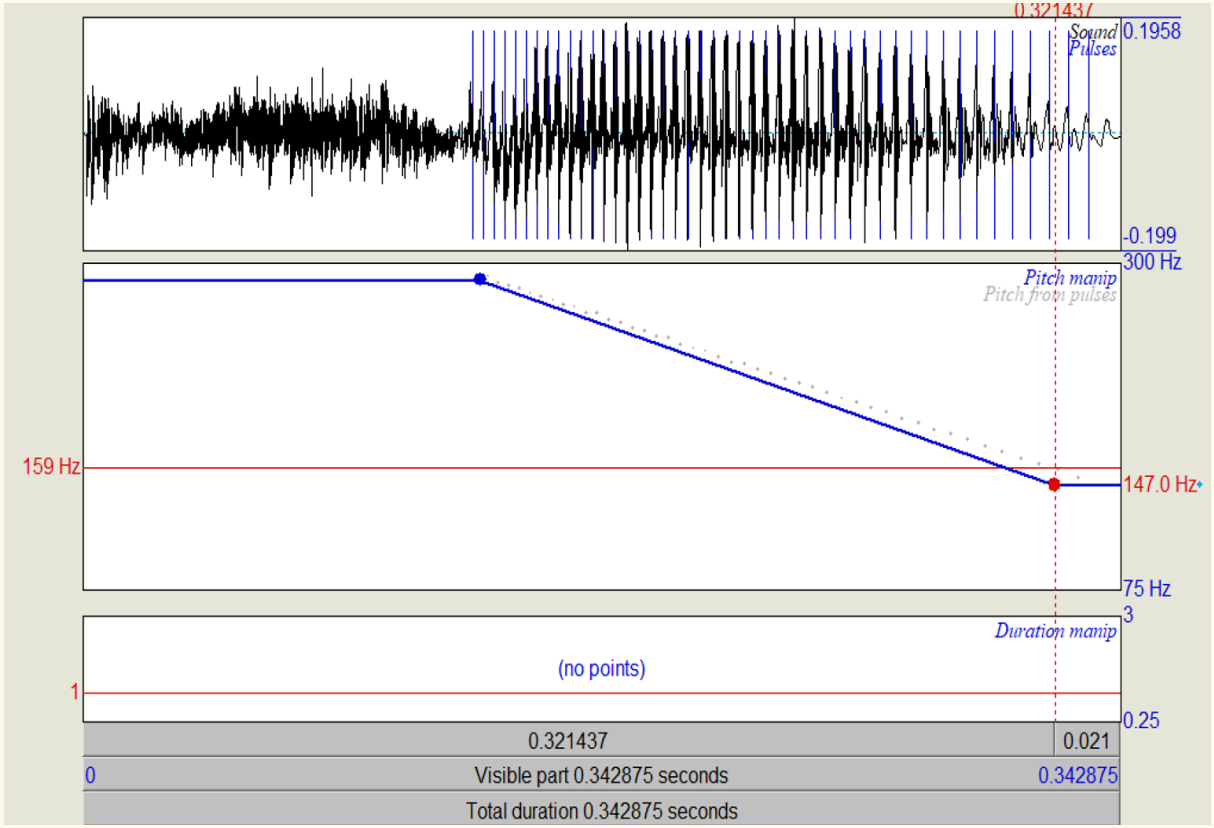
Pseudo-tone *falling* 41

Exaggerated stimuli

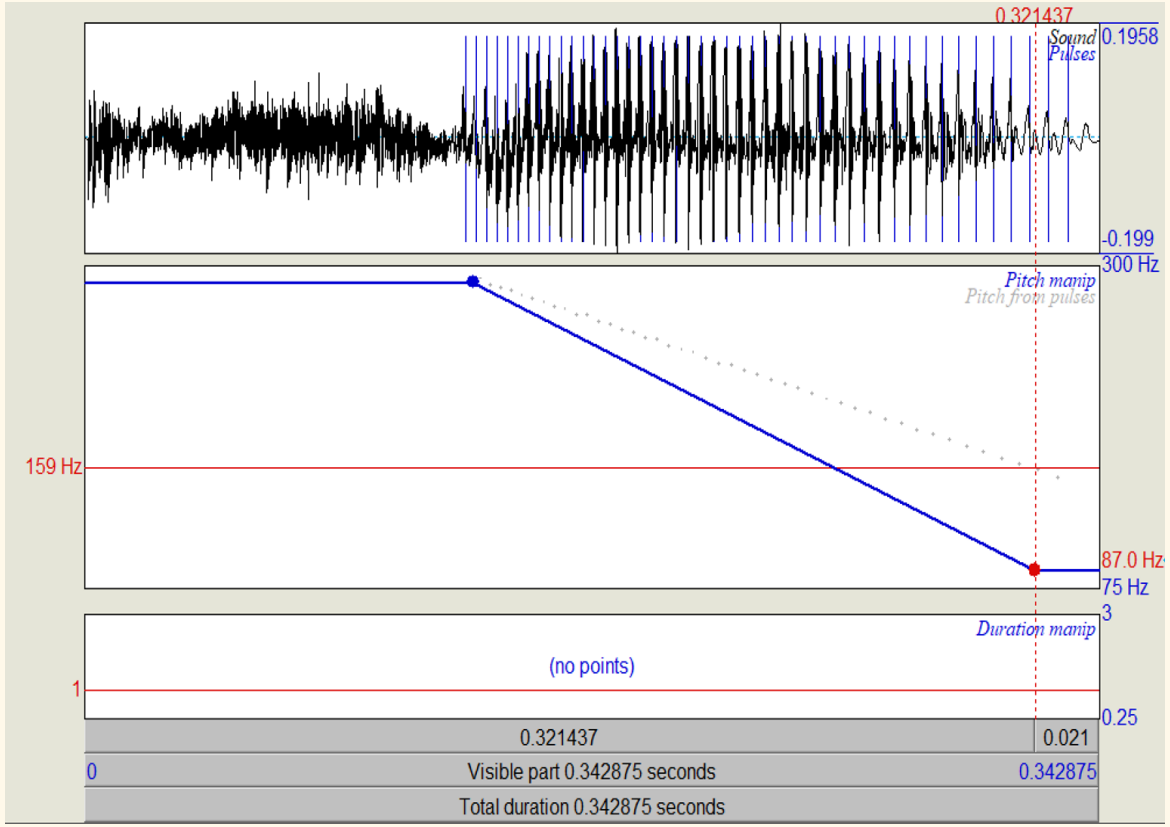
- Tone ***high-rising* 45 & *high-falling* 41**

Pseudo-tone (Chao digits) by block	Speaker 1 (Hz)	Speaker 2 (Hz)
High-rising **45**		
Block 2	297-435	294-395
Block 3	297-405	294-365
Block 4 (baseline)	297-375	294-335
High-falling **41**		
Block 2	297-87	294-141
Block 3	297-117	294-171
Block 4 (baseline)	297-147	294-201

Exaggeration

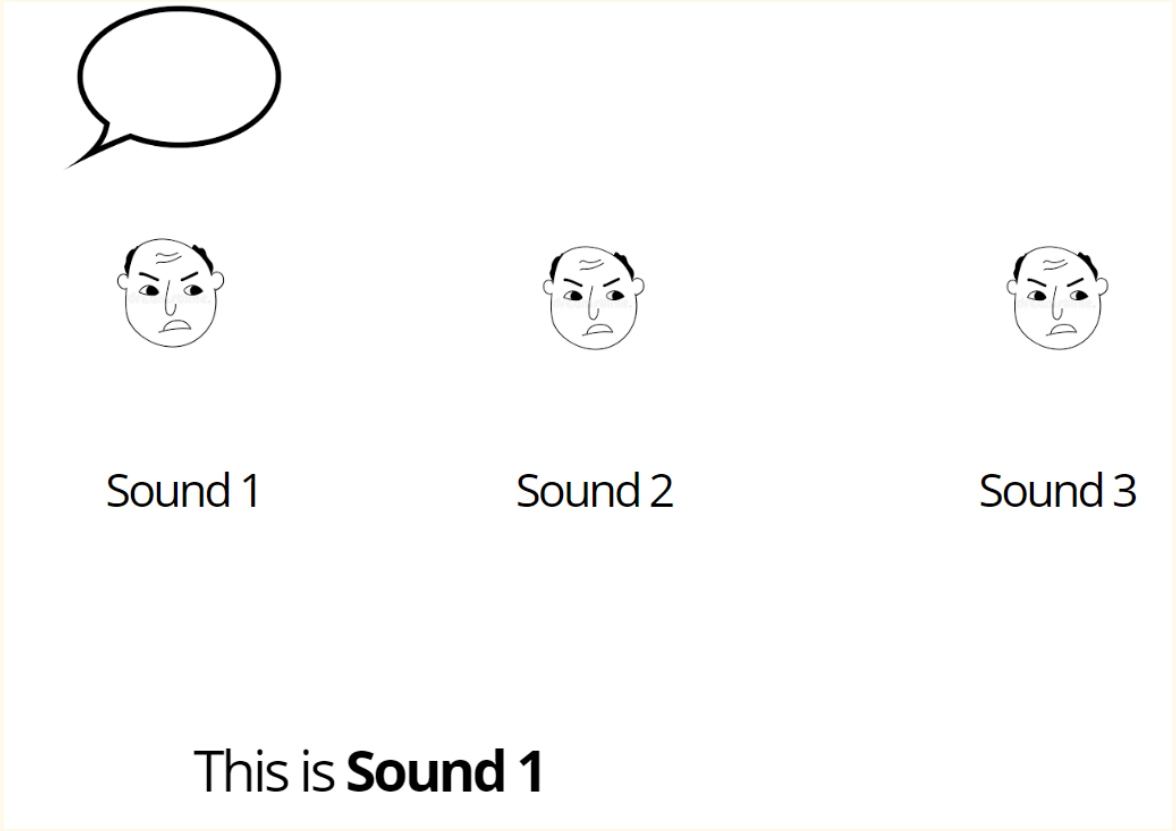


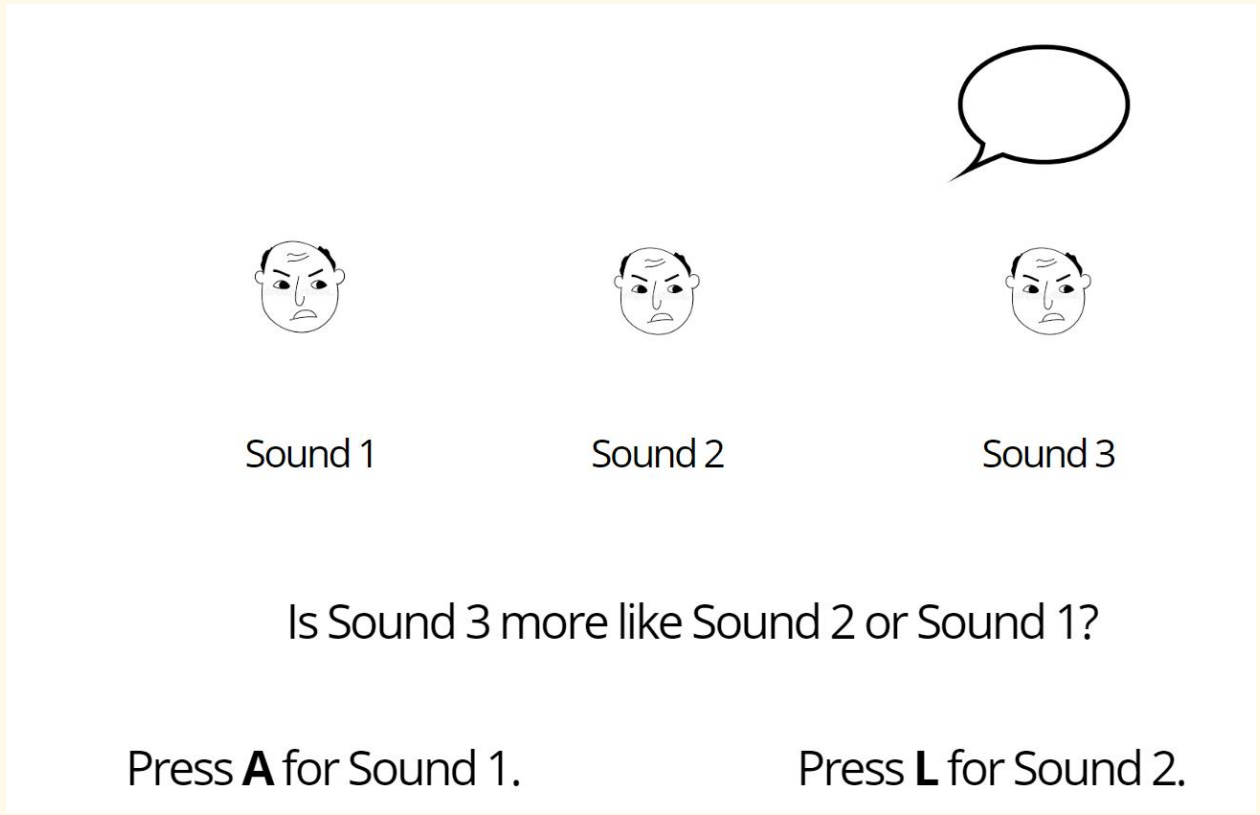
Baseline



Exaggeration 60Hz

ABX tasks





ká **kà** **ká**

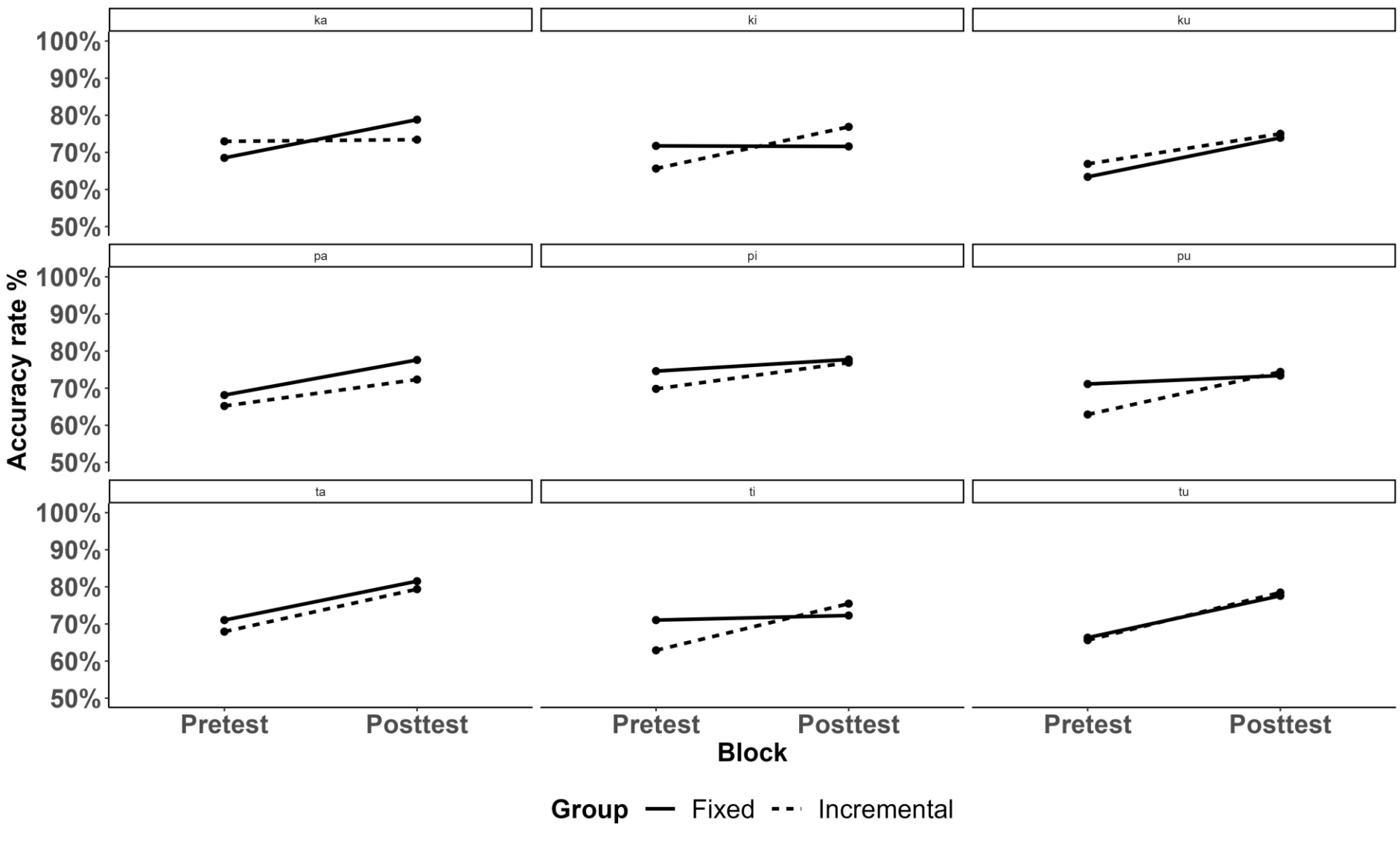
Speaker1 ***Speaker2*** ***Speaker1***

Syllable type – Accuracy test blocks

Newcastle University

Study1

Syllable: p = .16

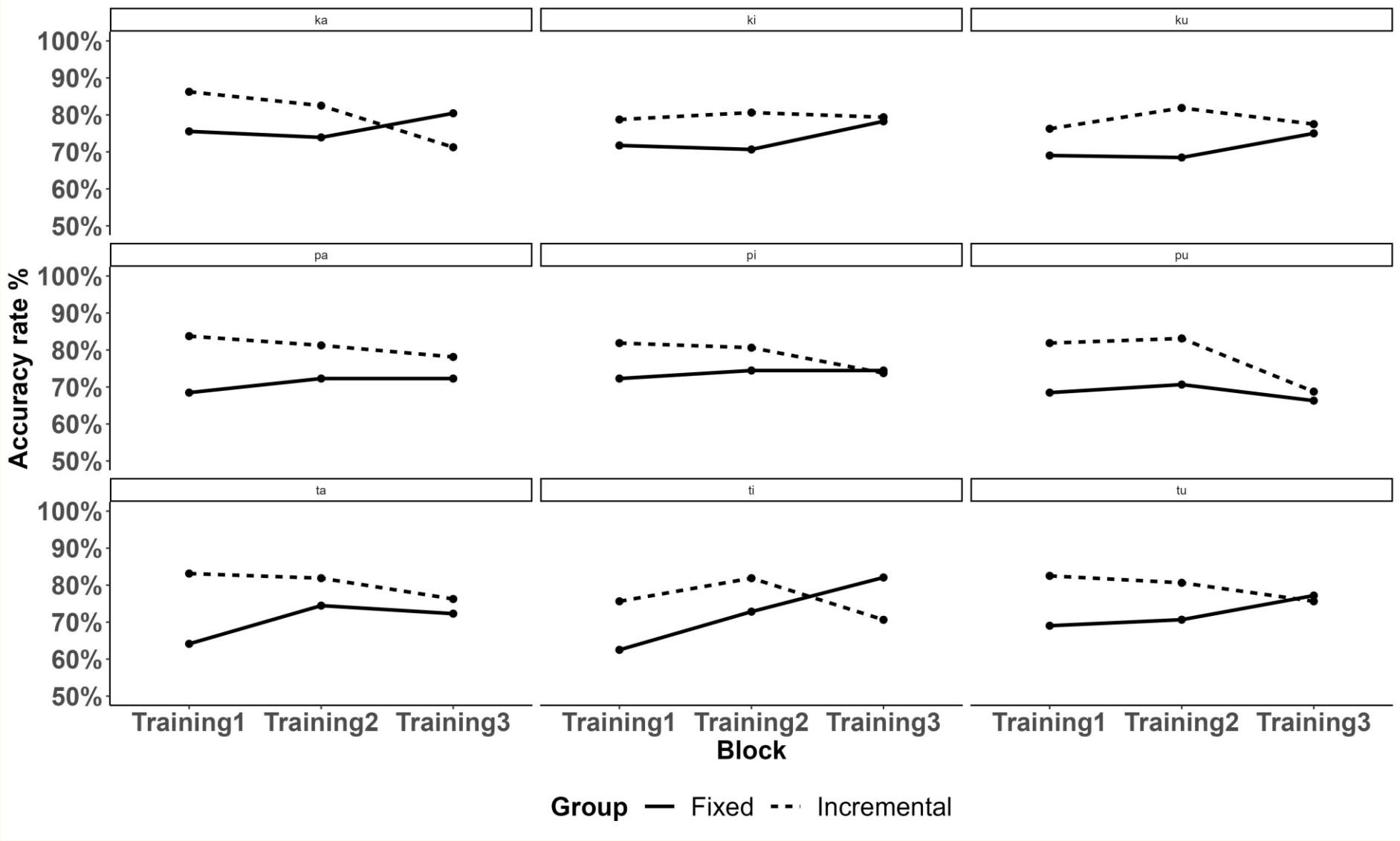


Syllable type – Accuracy training blocks

Newcastle University

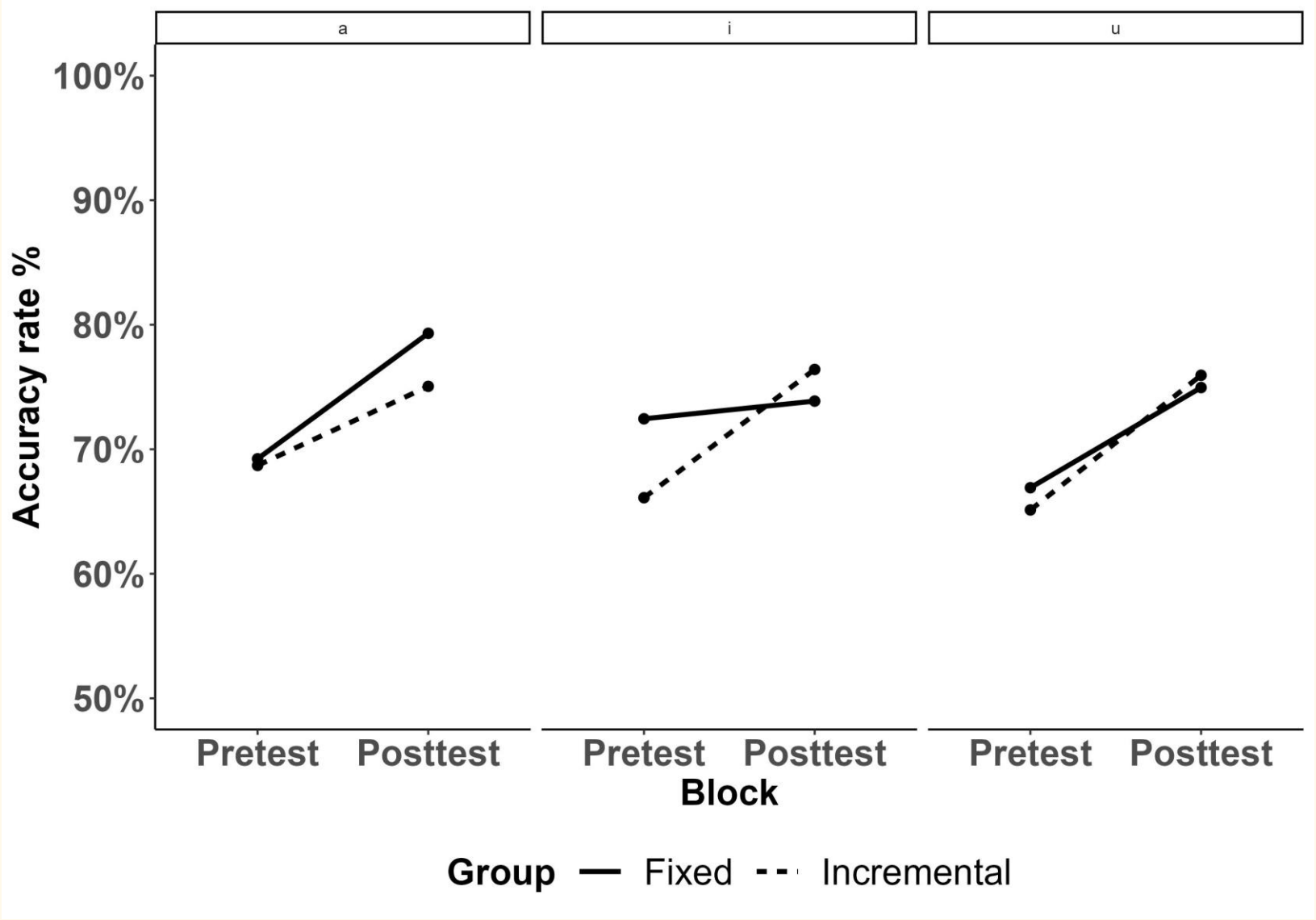
Study1

Syllable: p = .13



Vowel type – Accuracy test blocks

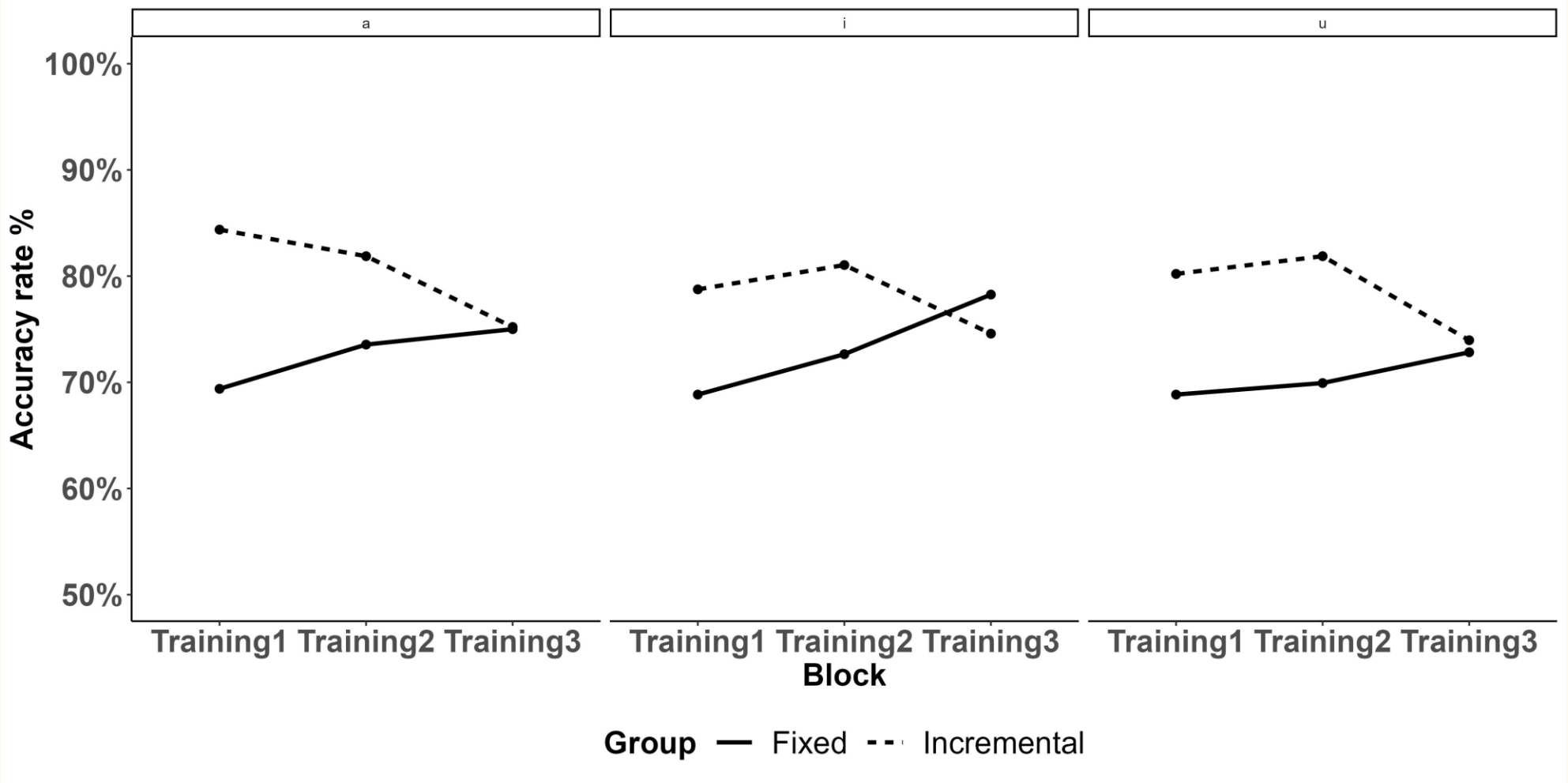
Vowel: p = .18



Vowel type – Accuracy training blocks

Newcastle University

Study1



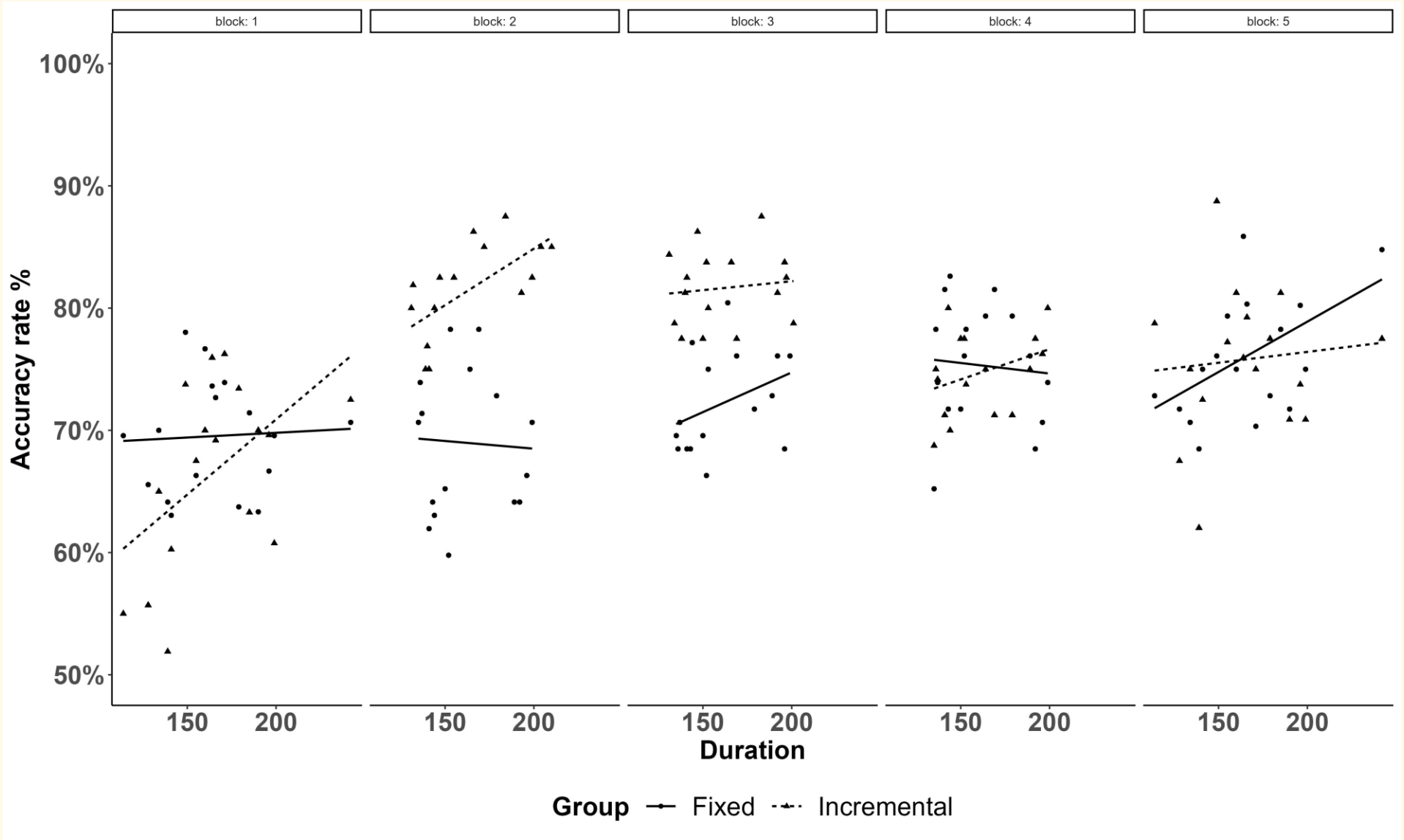
Vowel: p = .14

Tone duration– Accuracy

Duration:

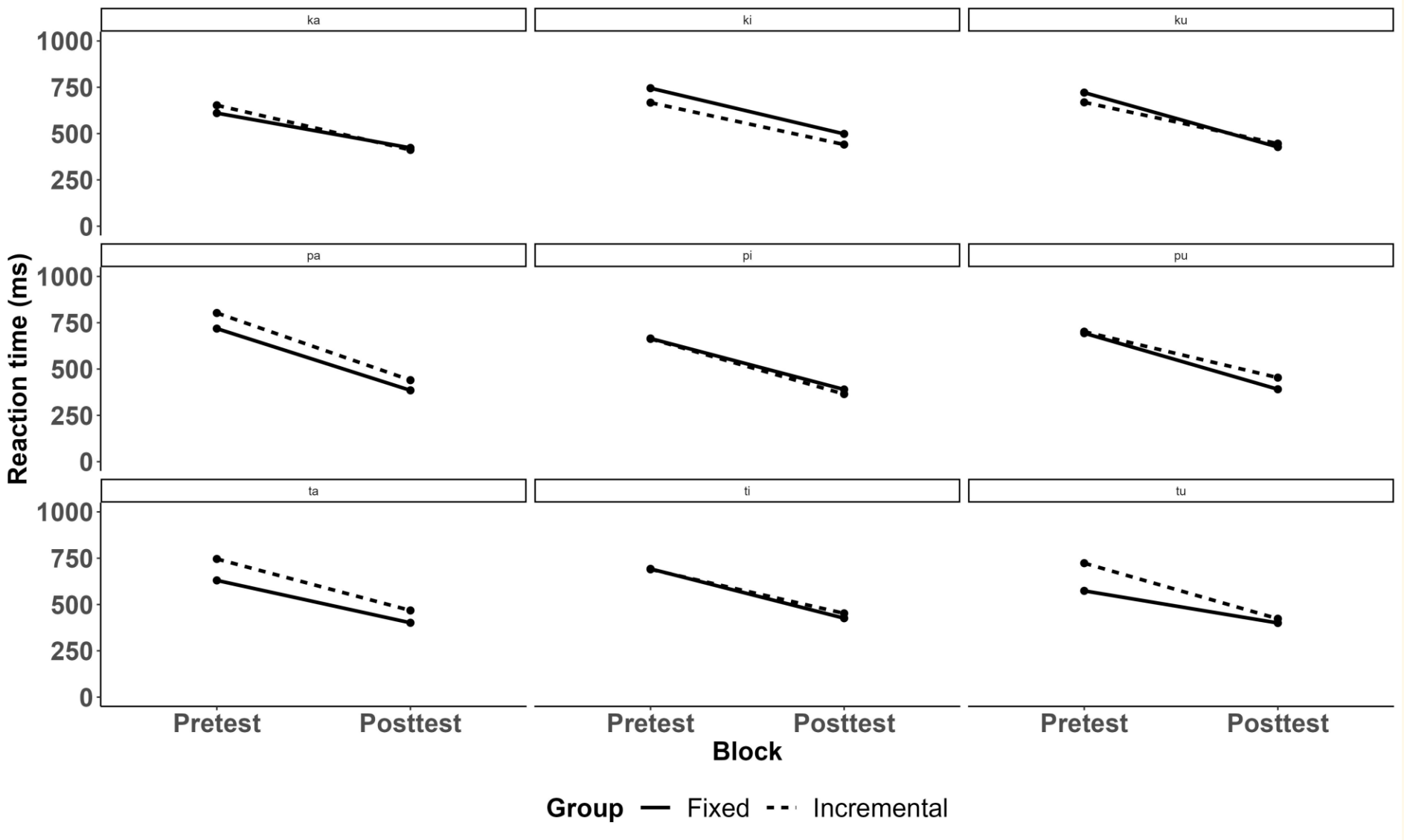
Test: p = .002 *

Training: p = .05 *



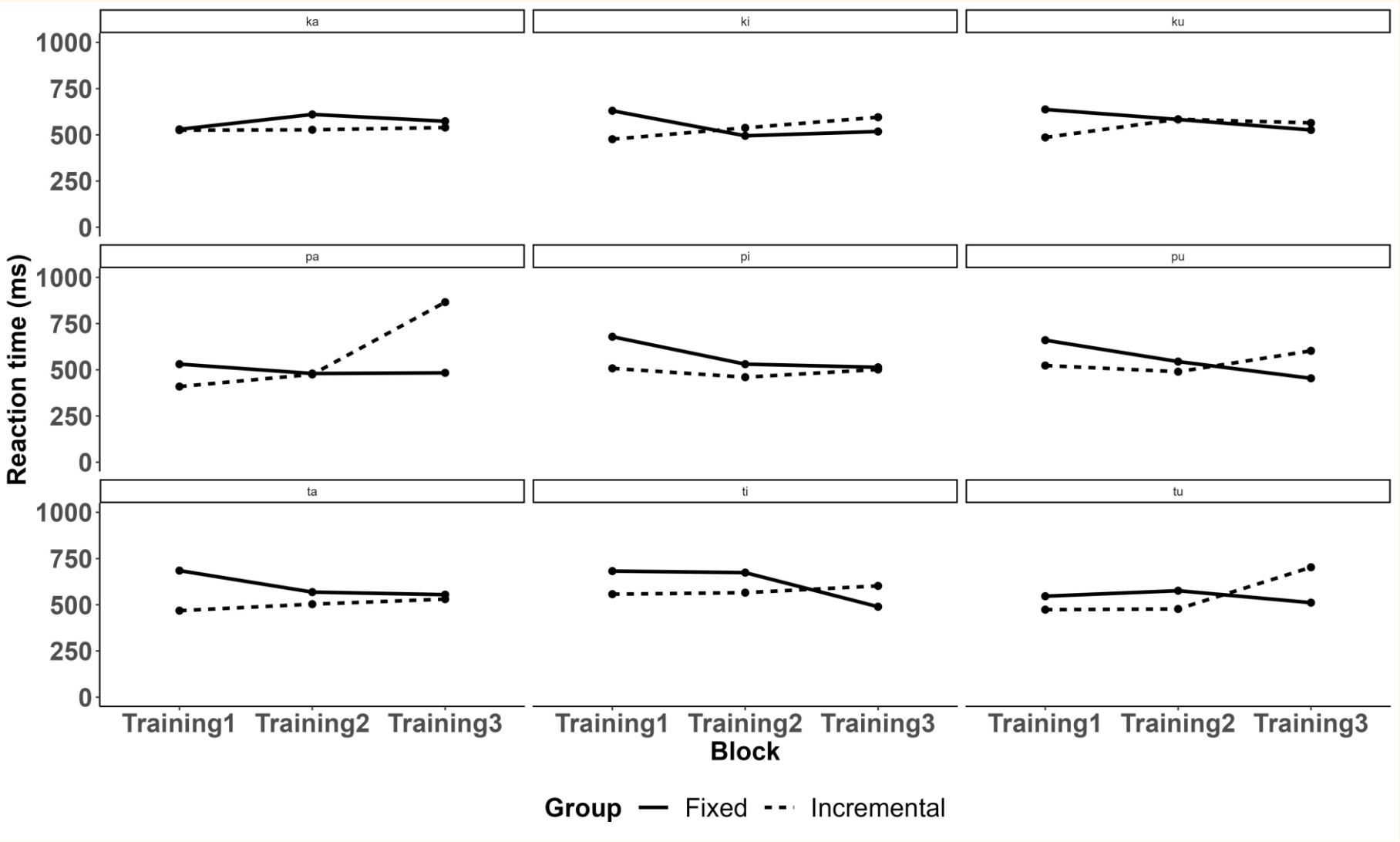
Syllable type – RT test blocks

Syllable: p = .33



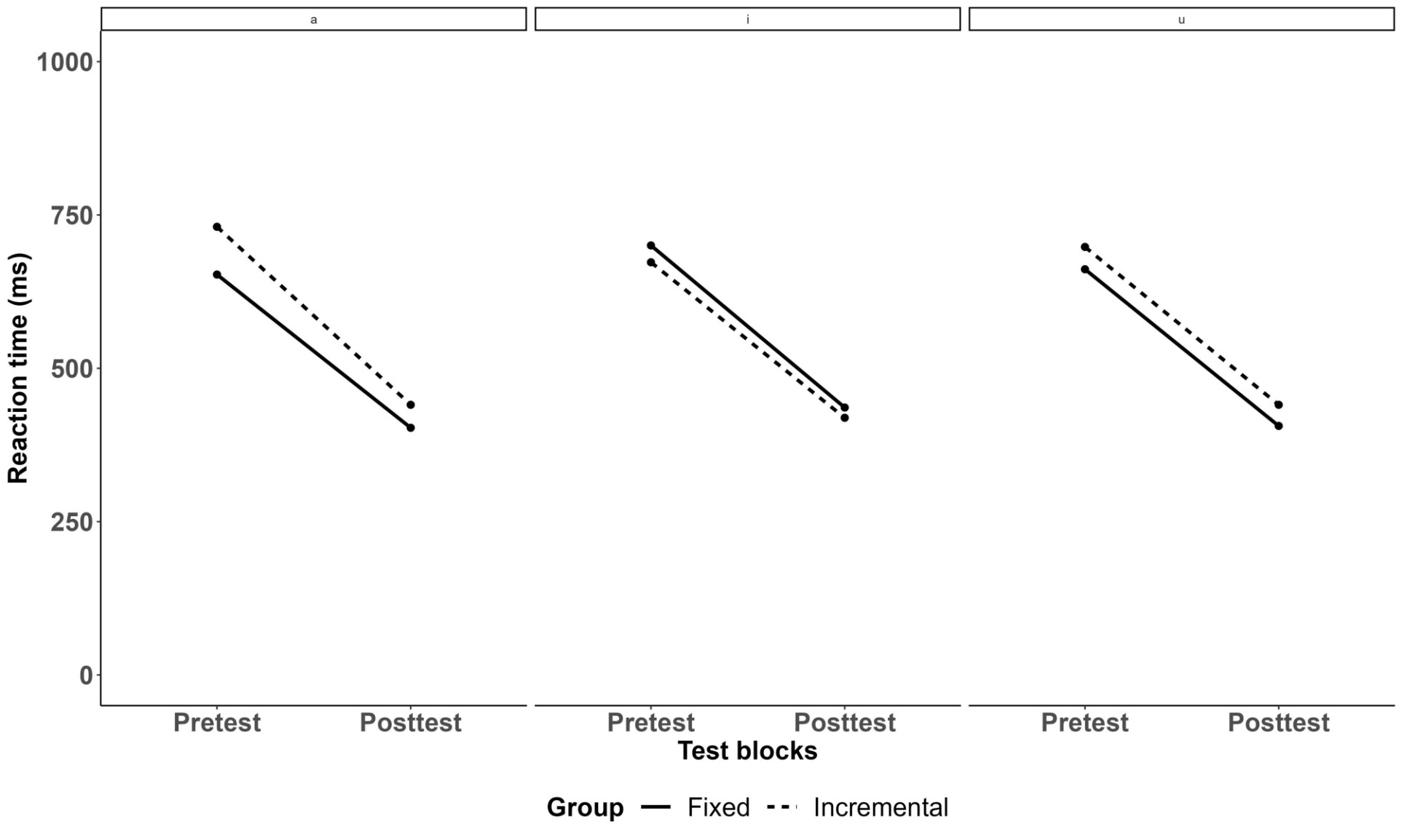
Syllable type – RT training blocks

Syllable: p = .05 *



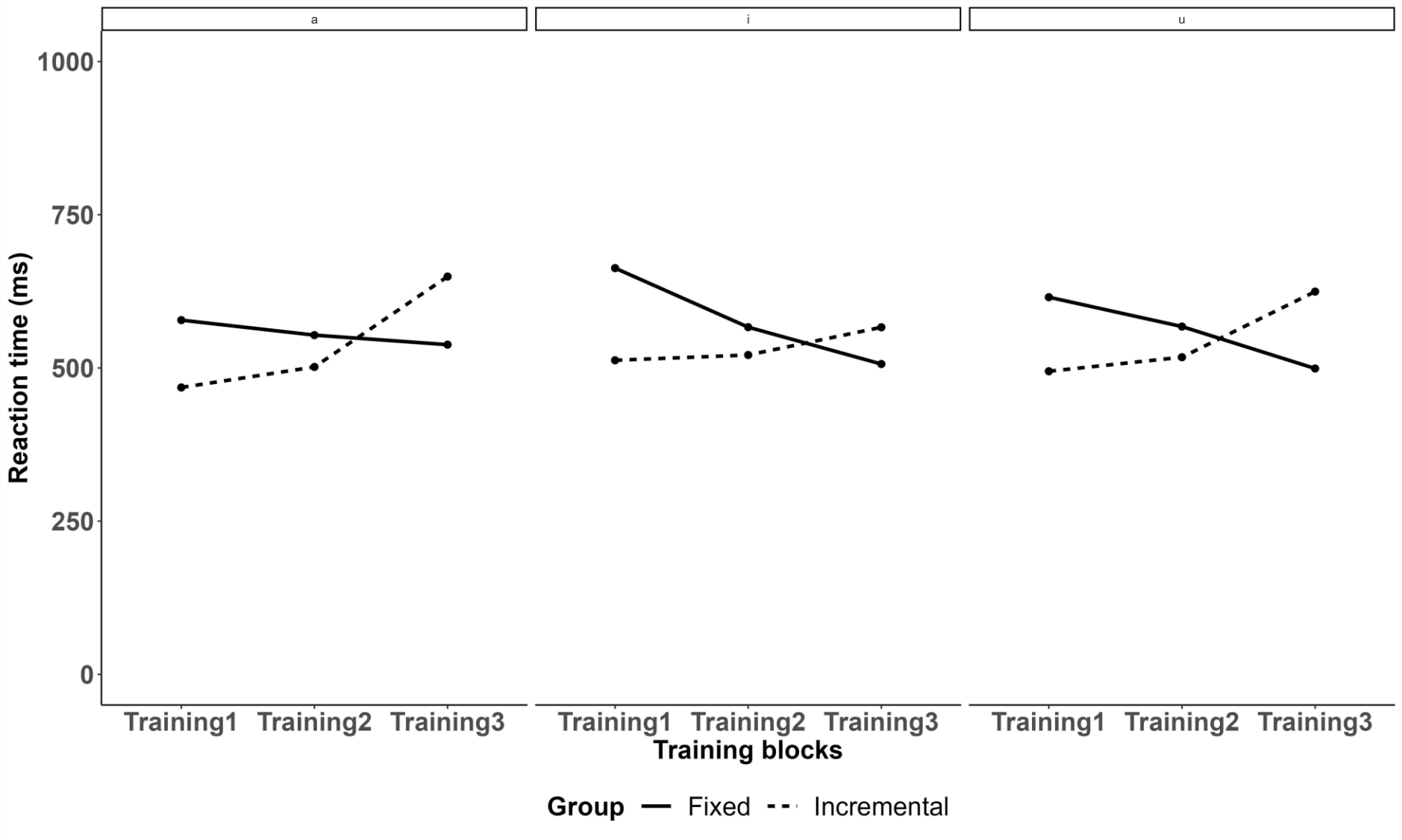
Vowel type – RT Tests

Vowel: p = .81

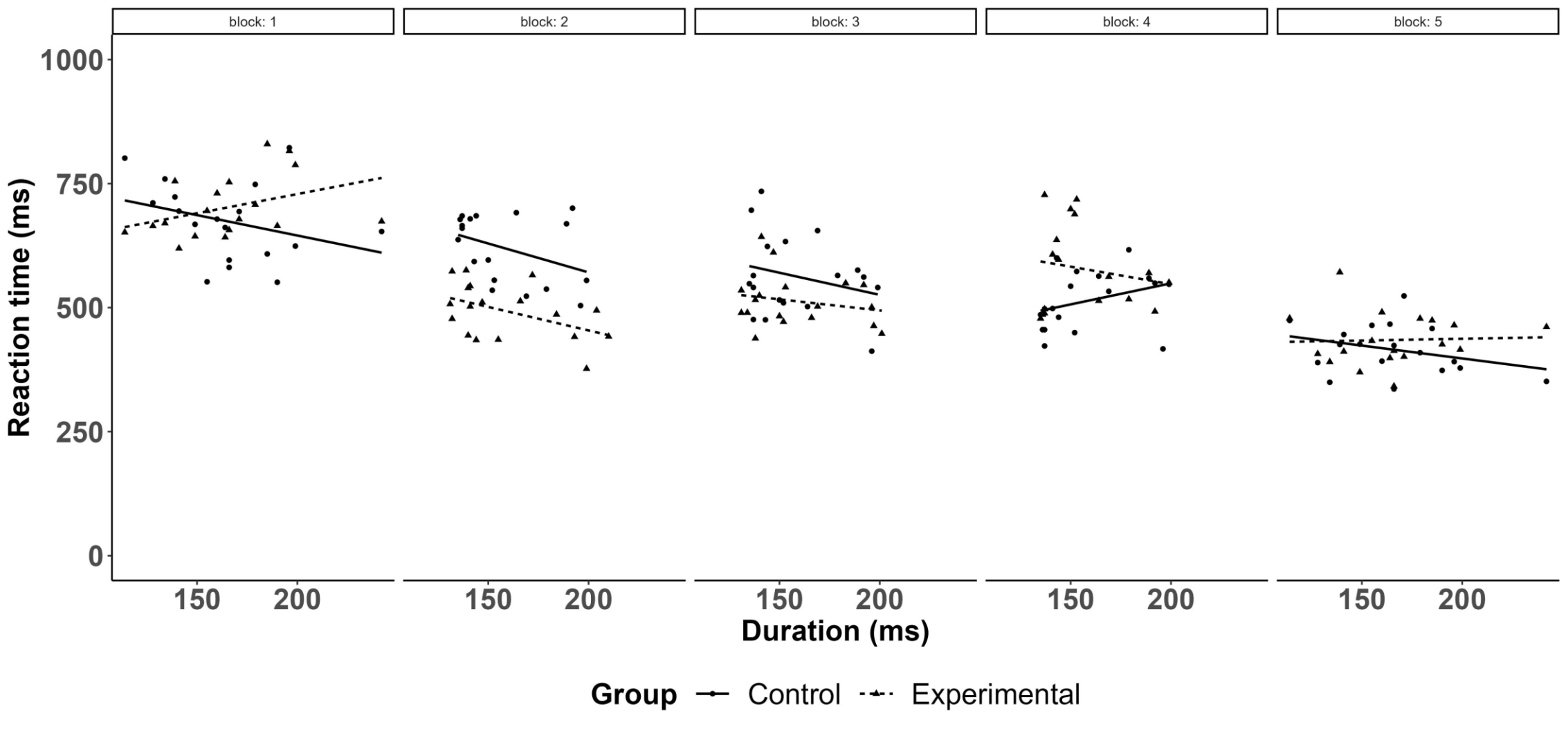


Vowel type – RT Training

Vowel: p = .01 *



Tone duration– RT



Duration: Test: p = .008 *; Training: p = .001 *

References

Chao, Y. R. (1930). A system of tone letters. Le maître phonétique.

Chandrasekaran, B., Sampath, P. D., & Wong, P. (2010). Individual variability in cue-weighting and lexical tone learning. The Journal of the Acoustical Society of America, 128(1), 456-465.

Gandour, J. (1983). Tone perception in Far Eastern languages. Journal of phonetics, 11(2), 149-175.

Kondaurova, M. V., & Francis, A. L. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: Comparison of three training methods. Journal of phonetics, 38(4), 569-587.

Iverson, P., & Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. The Journal of the Acoustical Society of America, 97(1), 553-562.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. Cognition, 87(1), B47-B57.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English/r/-/l/to Japanese adults. The Journal of the Acoustical Society of America, 118(5), 3267-3278.