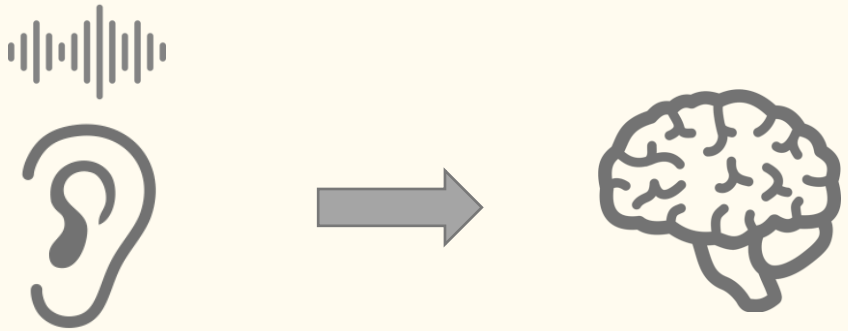
An incremental cue-based approach to extended lexical tone training

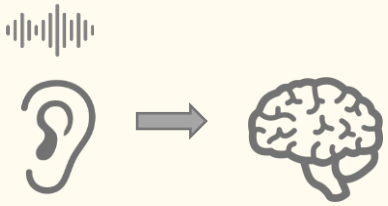
Yanyu Li, Laurence White & Ghada Khattab

Newcastle University

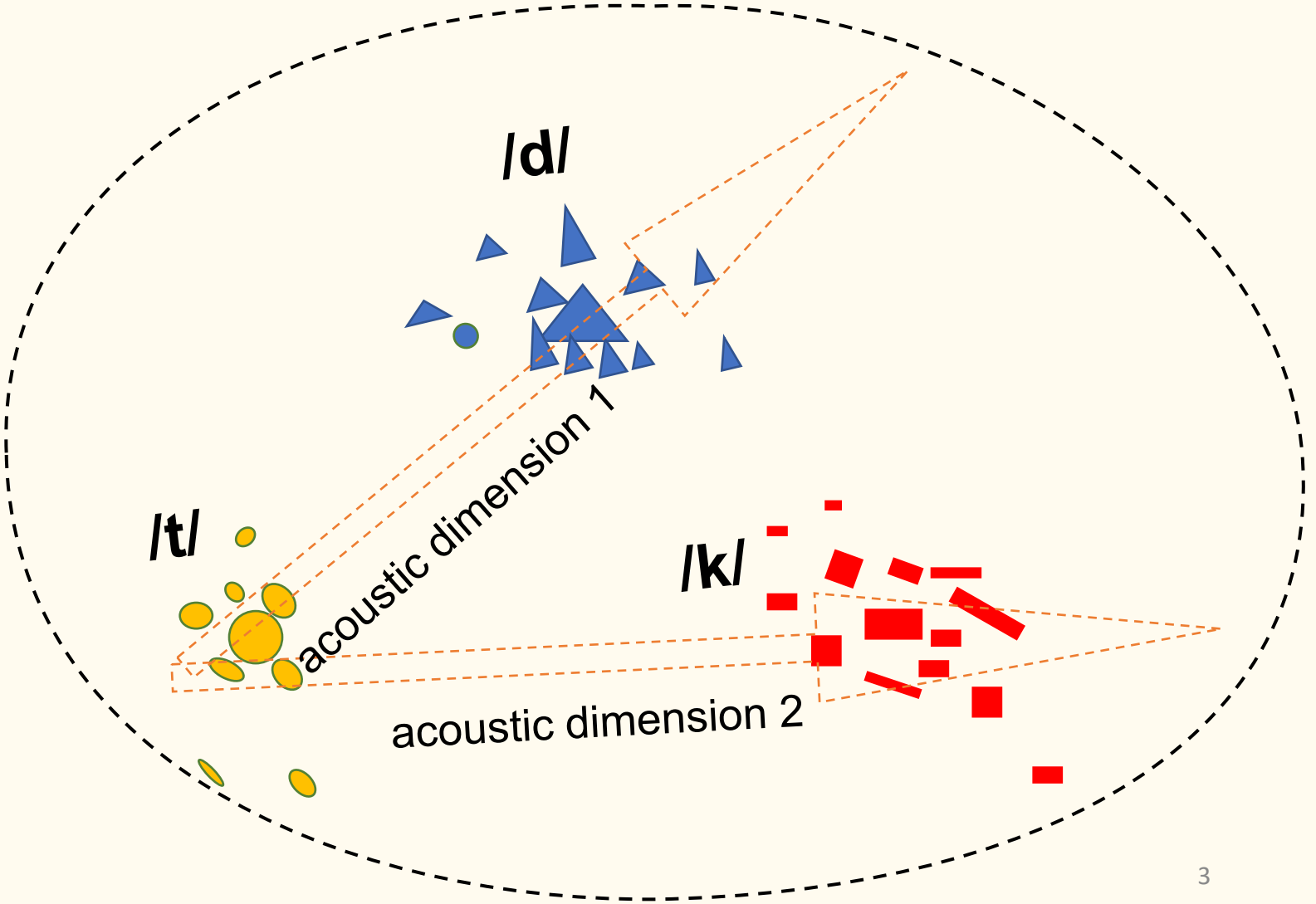


ta vs da

Perceiving speech categories



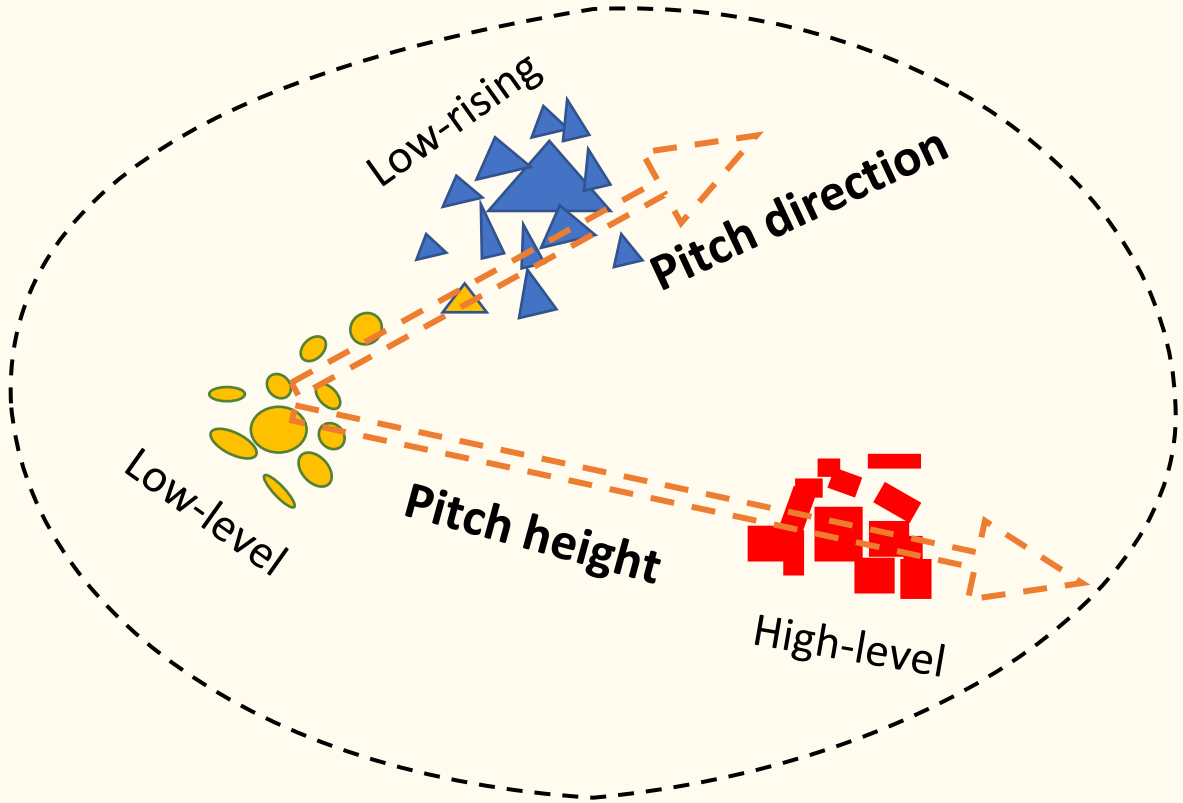
Theories such as
- Automatic Selective Perception (ASP) models
- Attention-to-dimension (A2D) models



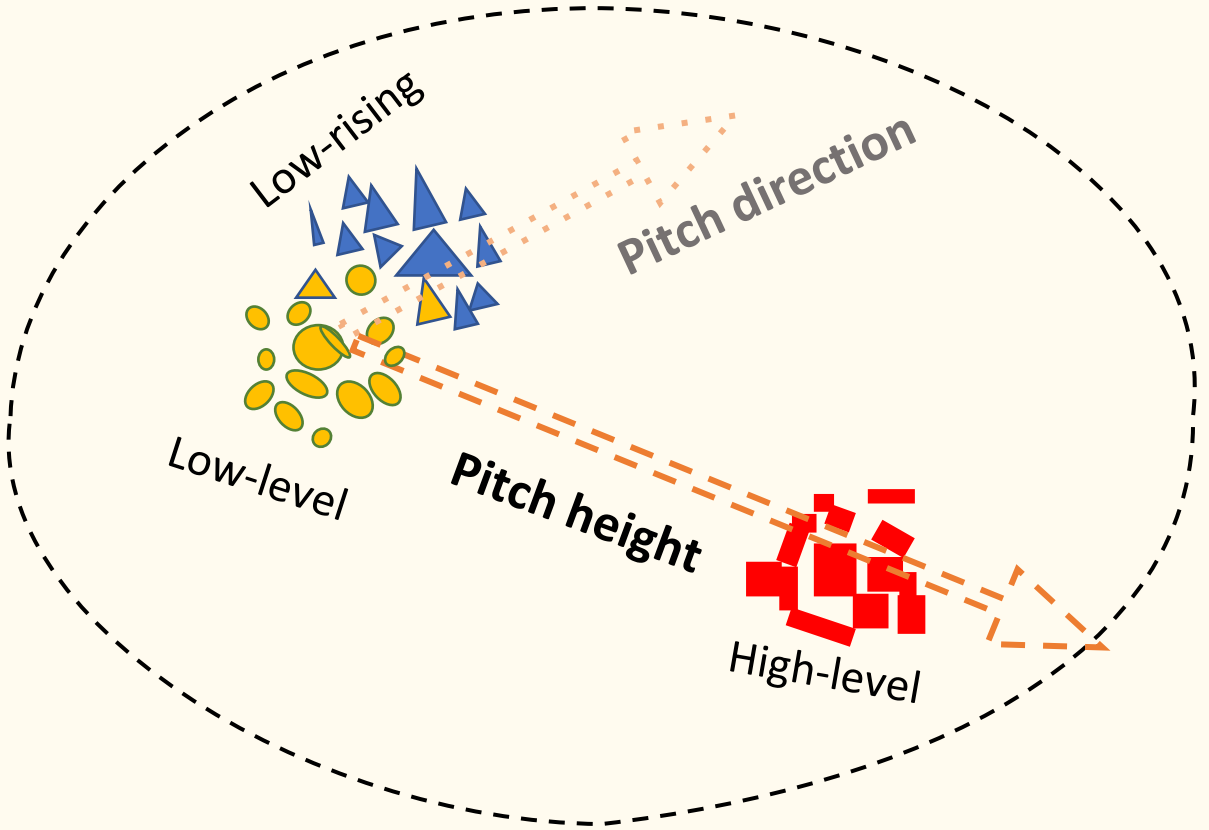
Francis & Nusbaum (2002), Strange (2011)

e.g., tú vs tù

Lexical tone



Tonal listeners



Non-tonal listeners

Chandrasekaran et. al (2010), Gandour (1983)

Cue exaggeration to redirect attention? – Incremental cue training

Cue exaggeration on “pitch direction”, i.e., tonal slope

Participants: L1 (first-language) English listeners with no tonal language experience.

Training 1 (**Most** exaggerated) → Training 2 (**Less** exaggerated) → Training 3 (**Least** exaggerated)

- /i/-/ɪ/ (Kondaurova and Francis, 2010)
- /l/-/r/ (Iverson et al., 2005)

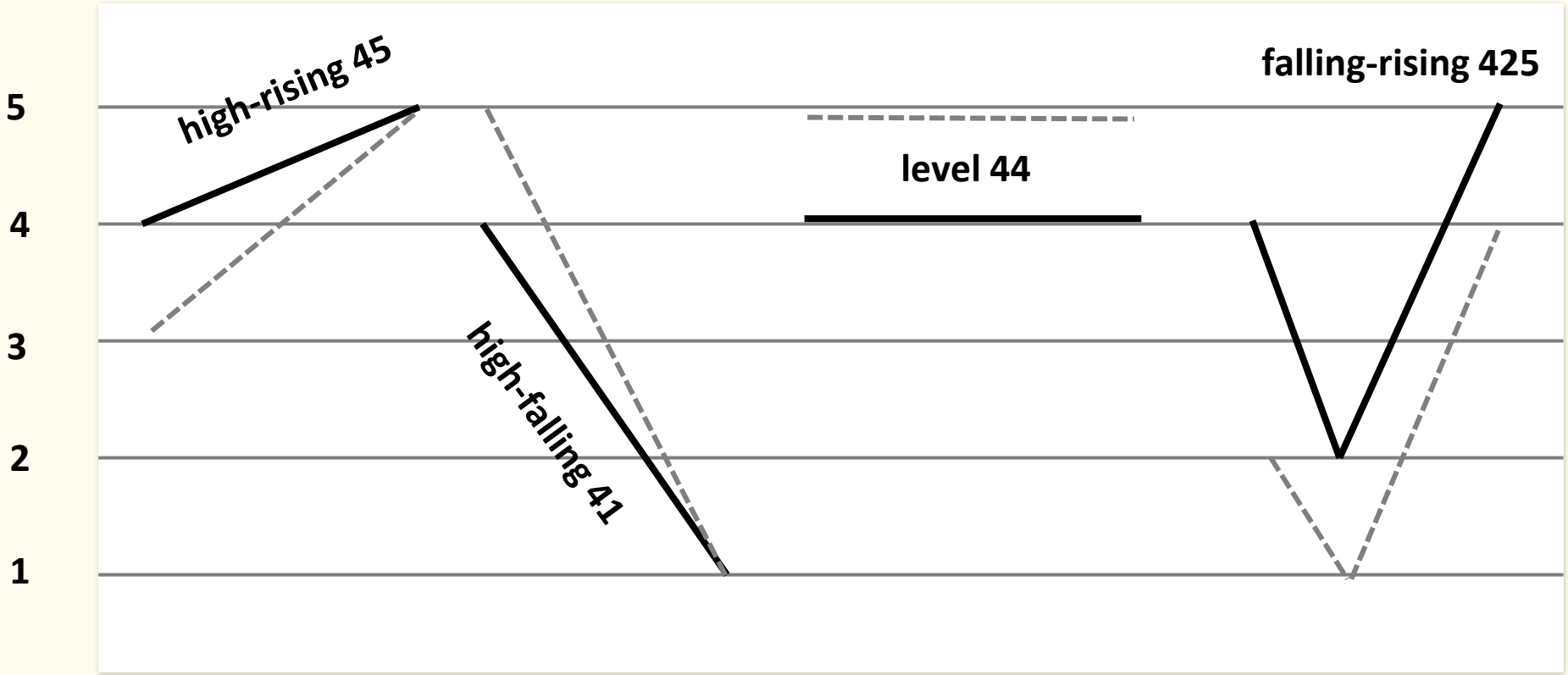
Extended incremental cue training

- Is **incremental cue training** more beneficial than **fixed unmanipulated stimuli exposure** in lexical tone perceptual learning?
- How does the effect change for a training course lasting **multiple days**?

Extended incremental cue training

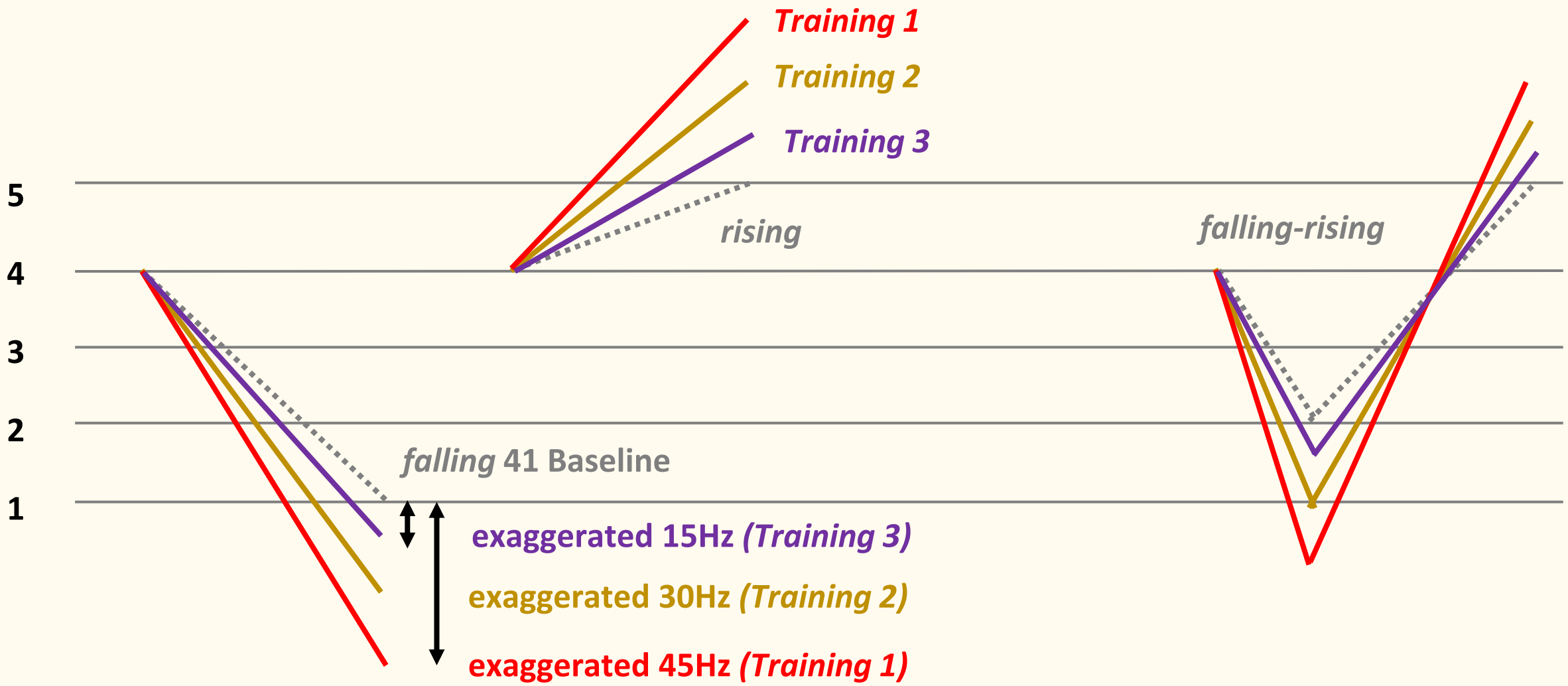
- Is **incremental cue training** more beneficial than **fixed unmanipulated stimuli exposure** in lexical tone perceptual learning?
- How does the effect change for a training course lasting **multiple days**?

Tonal stimuli

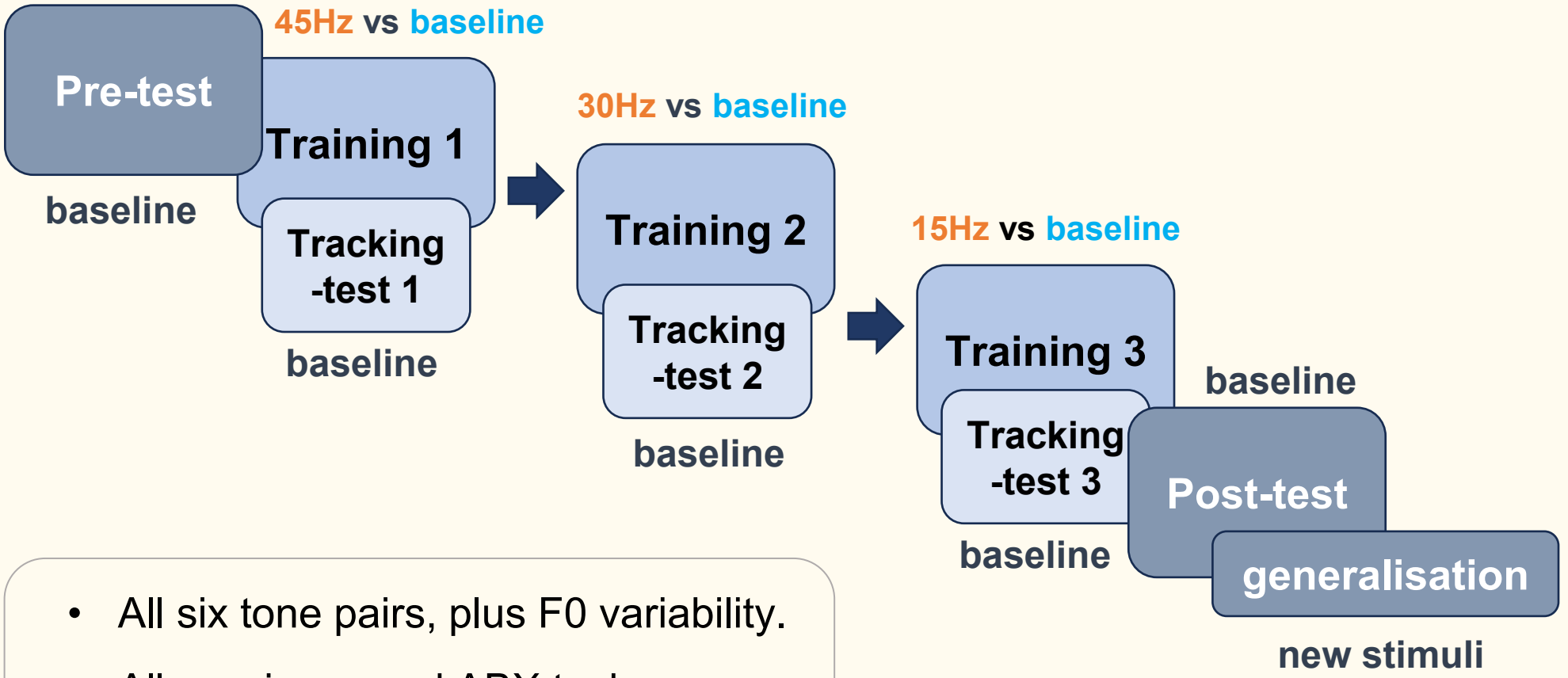


9 Syllables: /pa:/, /ta:/, /ka:/, /pi:/, /ti:/, /ki:/, /pu:/, /tu:/, and /ku:/

Cue exaggeration for incremental group: tonal slope



Extended incremental cue training procedure



- All six tone pairs, plus F0 variability.
- All sessions used ABX tasks
- Feeback only in Training 1, 2, & 3

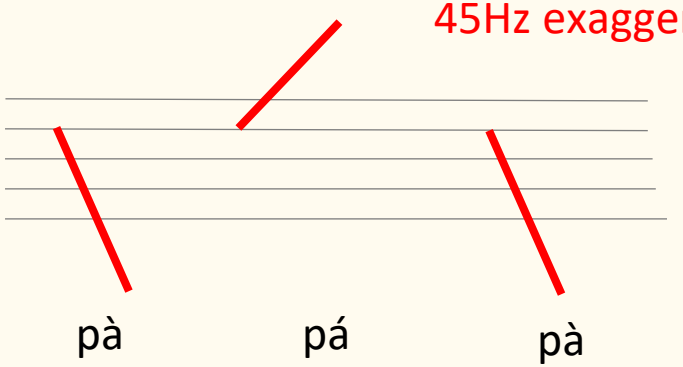
ABX trial examples: pà pá pà

In Training 1 vs Training 2

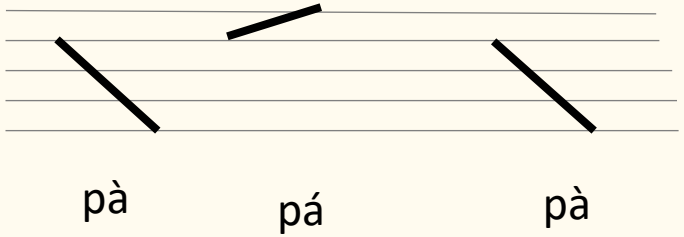
Training 1

Incremental group

45Hz exaggeration

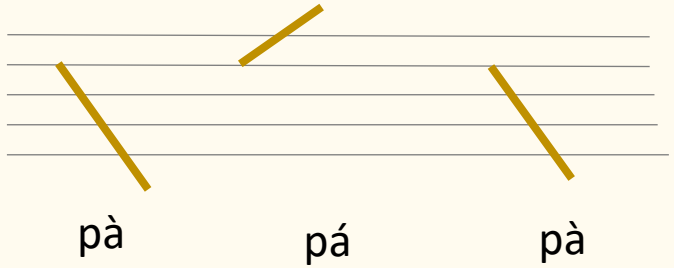


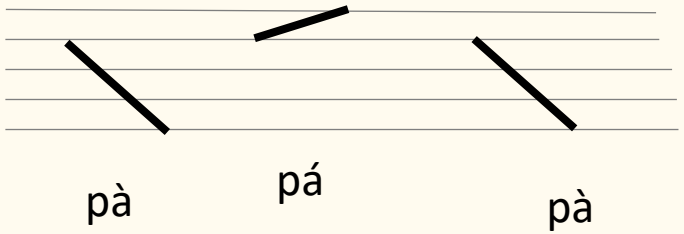
Fixed group



Training 2

30Hz exaggeration





Data analysis

- Mixed-effect logistic/linear regression models.
- Training or tests data are analysed separately.

training 1, 2, 3

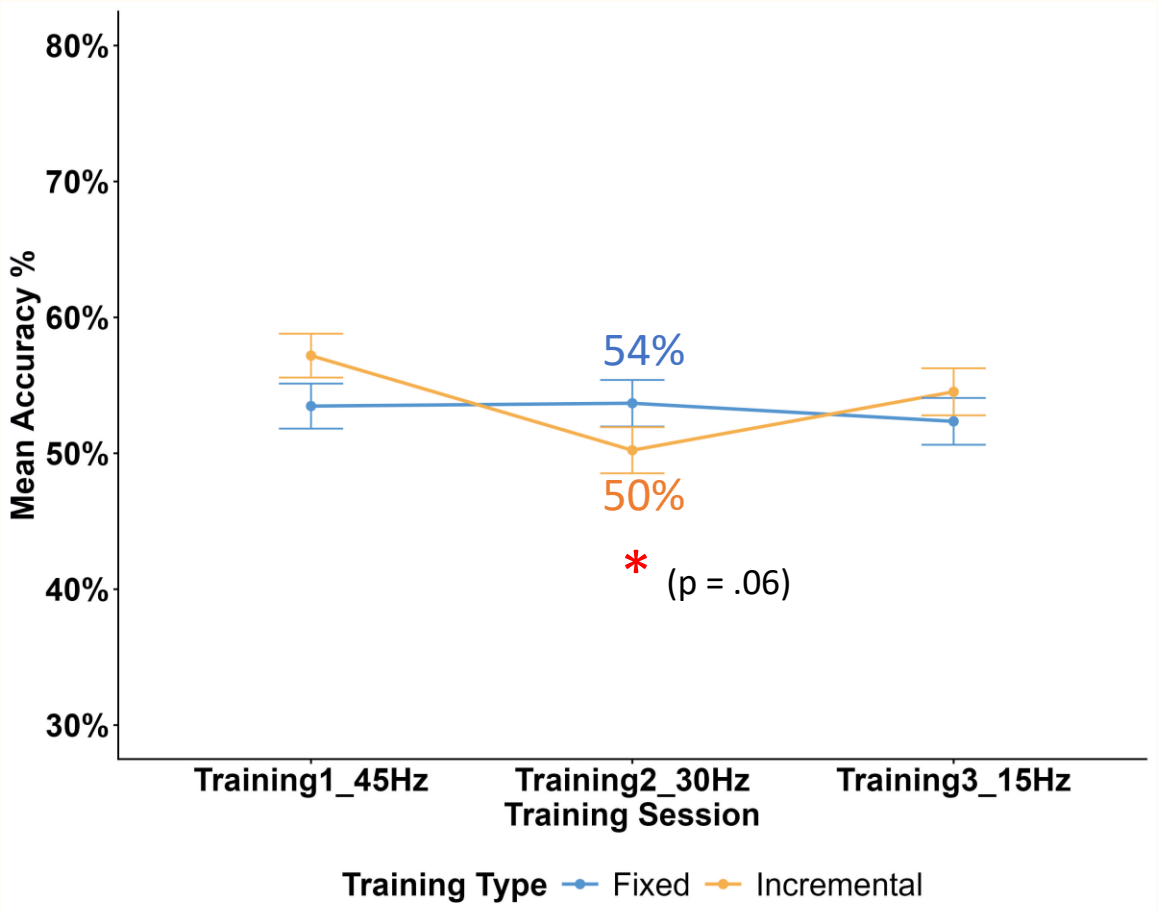
tracking-tests 1, 2, 3

pre- vs post-tests

- Model structure:
  - Measurement ~ block * training_type + (1+block|participant) + (1|syllable)
  - Measurement: RT, accuracy, combined accuracy and RT, sensitivity (d')

BIS (Balanced Integration Score)

Results in training: Accuracy

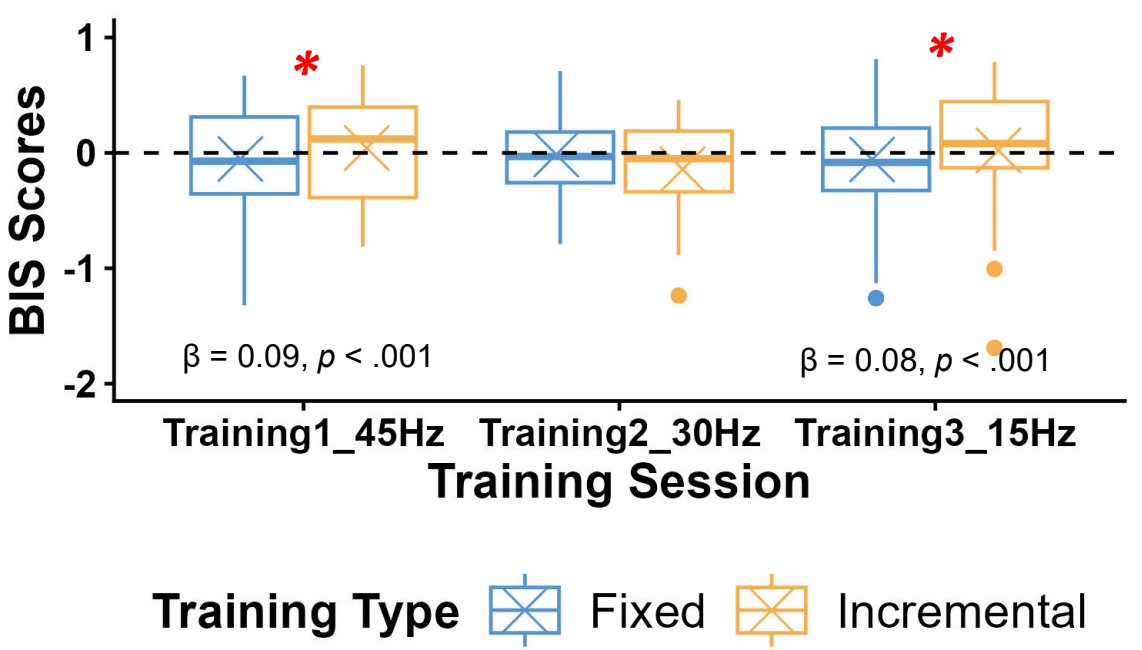


Training: Training Session (p = .07)

Balanced Integration Score

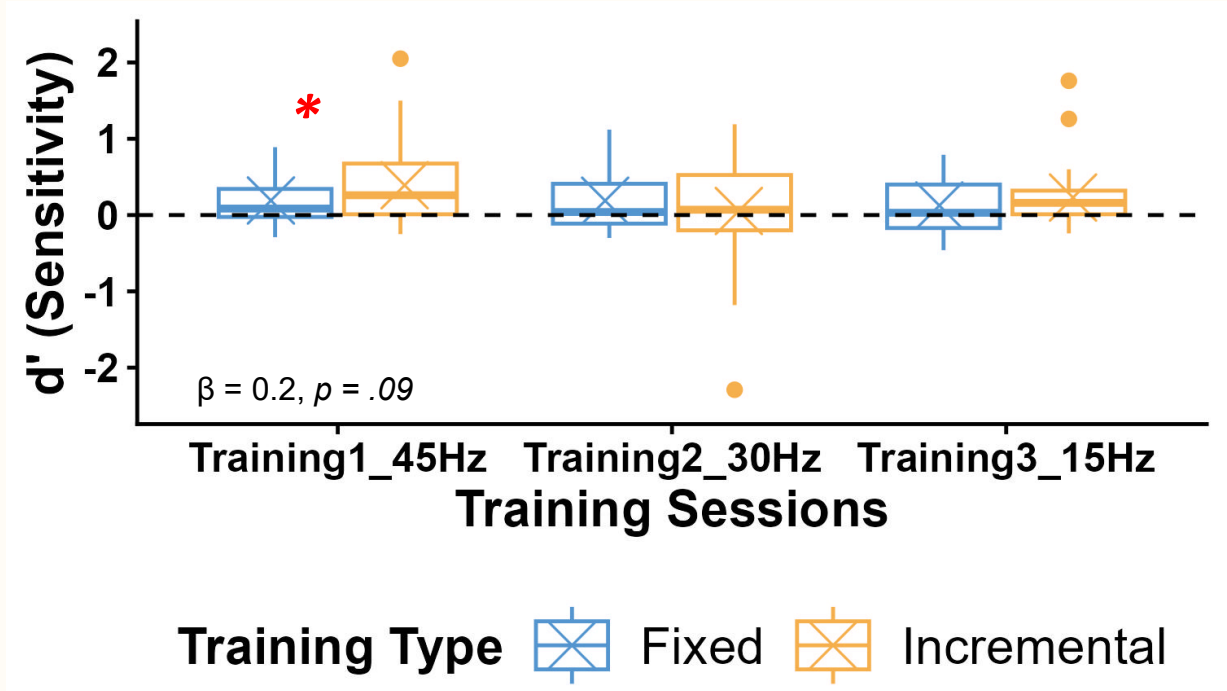
BIS & Sensitivity scores (d')

in training



Training: BIS scores

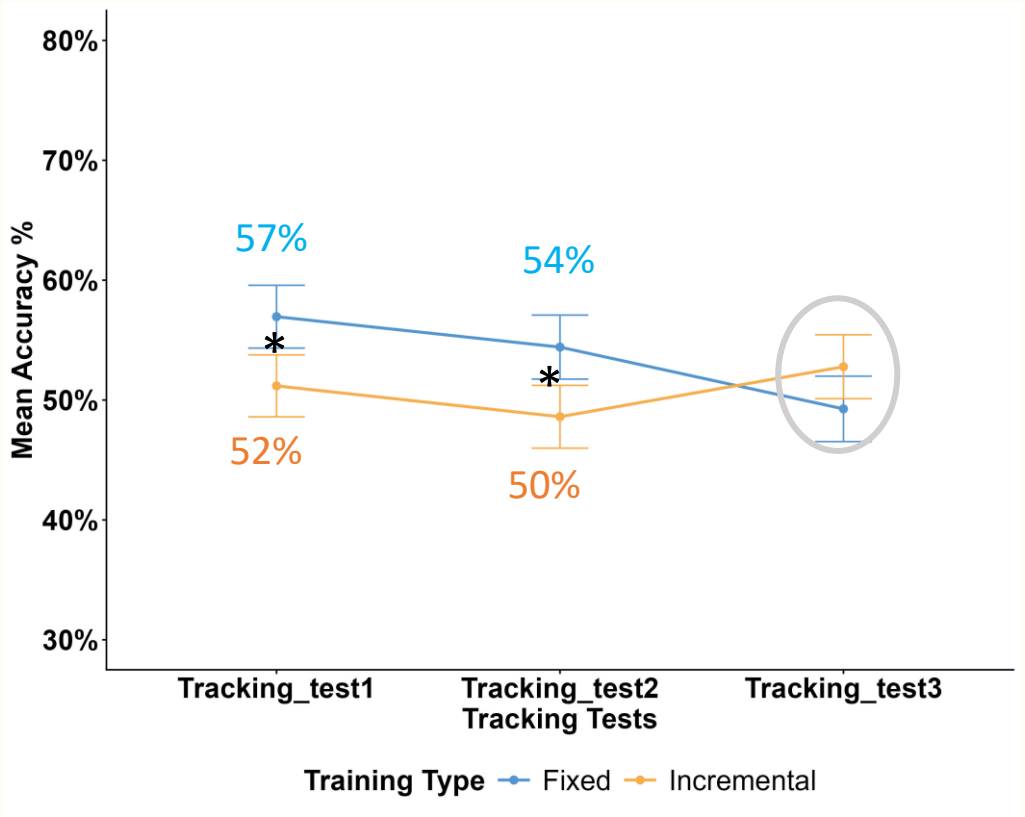
Training Session * Training Type



Training: Sensitivity scores (d')

No main/interaction effect

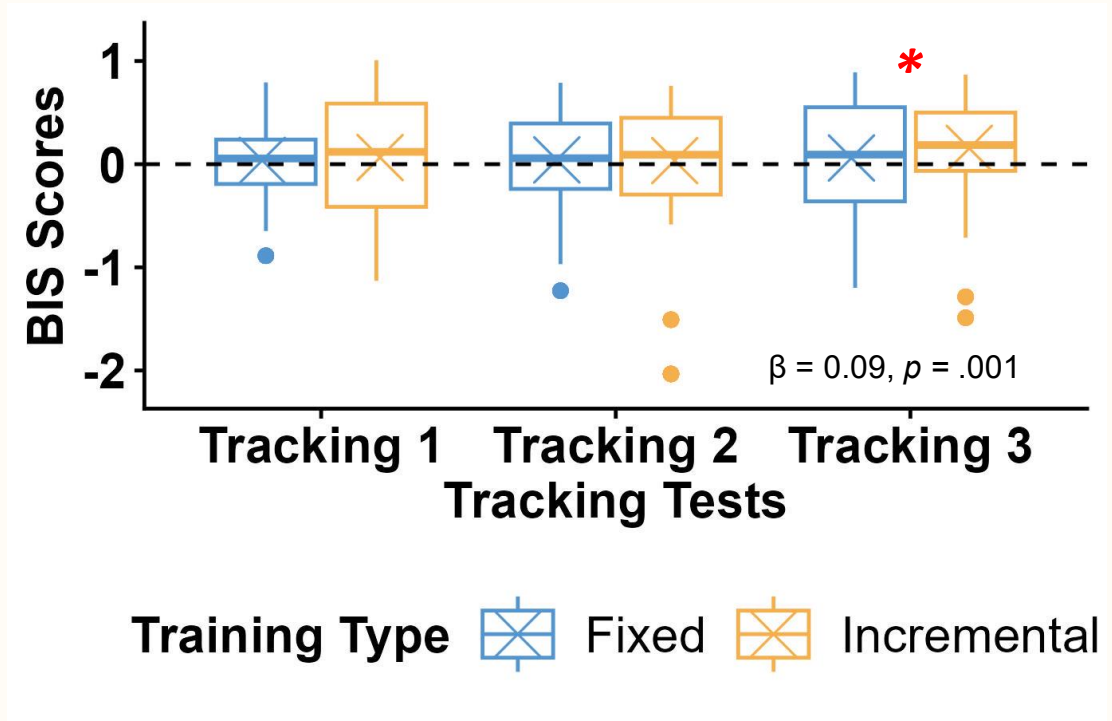
Results in tracking-tests: Accuracy



Tracking Tests: Training_type * Tracking_test

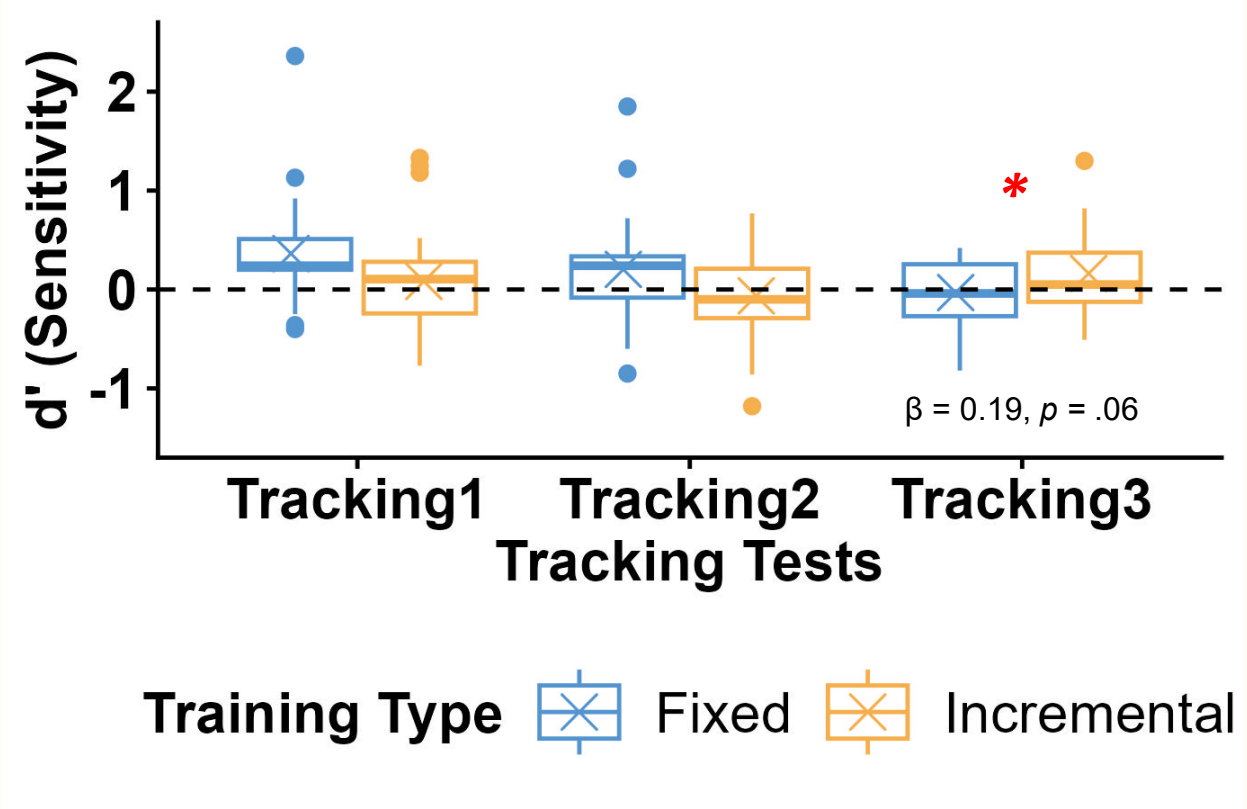
BIS & Sensitivity scores (d')

in tracking-tests



Tracking-test: BIS scores

Training Type : Tracking-test



Tracking-test : Sensitivity scores

Training Type * Tracking-test

Key findings

BIS & Sensitivity (d')

- perceptual boost in training
- sustainable benefits in tracking-tests?

Auditory salience

- Independent of language experience

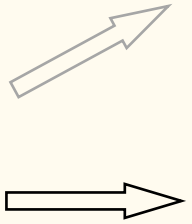
Perceptual salience

Statistical learning

Key findings

BIS & Sensitivity (d')

- perceptual boost in training
- sustainable benefits in tracking-tests?



Auditory salience

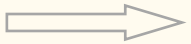
Perceptual salience

- Dependent of language experience

Key findings

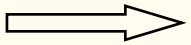
BIS & Sensitivity (d')

- perceptual boost in training
- sustainable benefits in tracking-tests?



Auditory / perceptual salience

Comparable accuracy



Statistical regularities

- Categorisation
- Informativeness

Take away…

- perceptual boost
- sustainable benefits
- Comparable accuracy

→ Auditory salience

→ Perceptual salience

→ Statistical learning

- Evidence for perceptual boost and sustained incremental benefits from BIS and d'.
- The modest incremental benefits might be more robust when training designs reflect statistical regularities of speech input.

Newcastle University

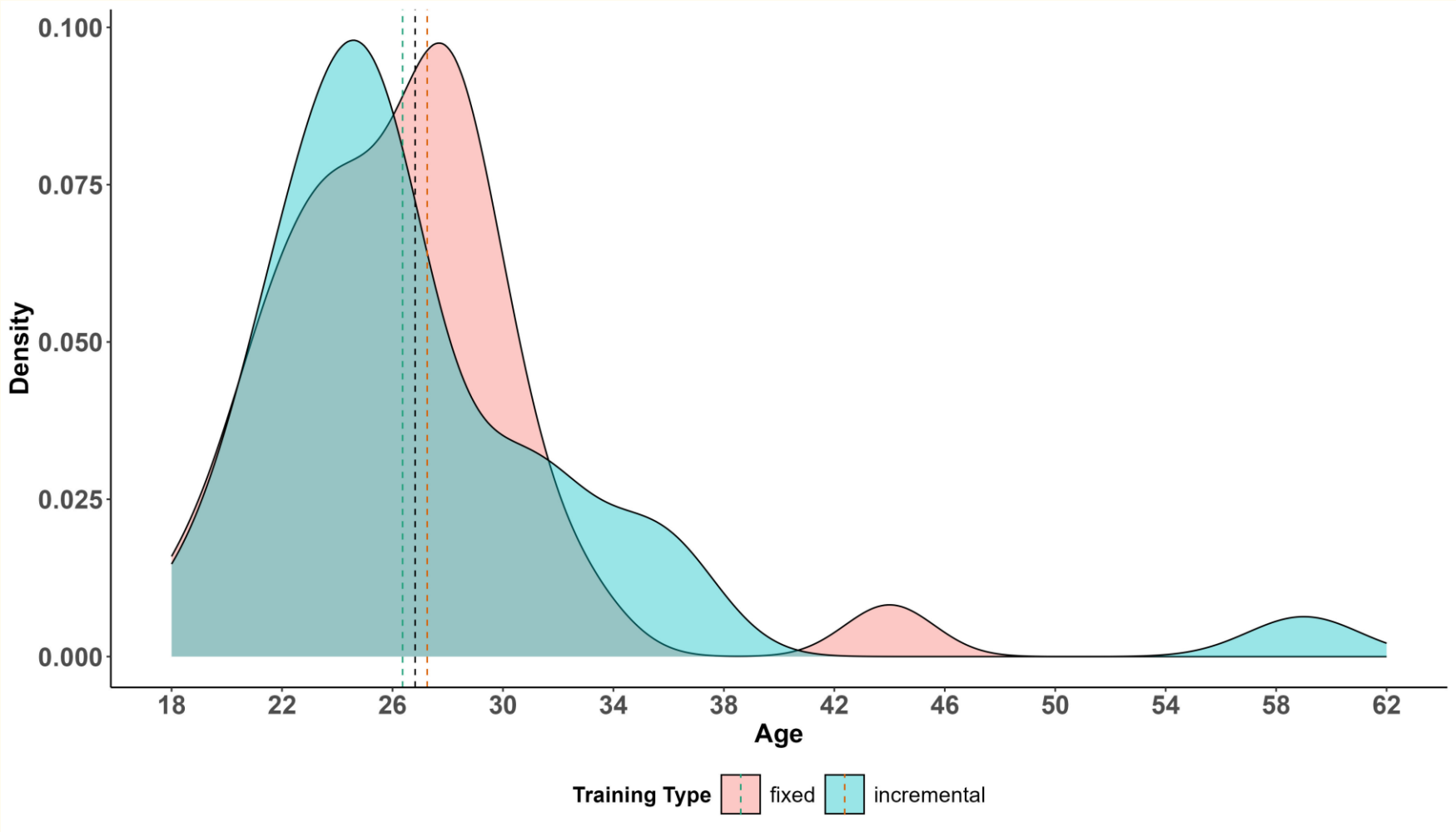
CHINA SCHOLARSHIP COUNCIL

@yanyuniverseli

References

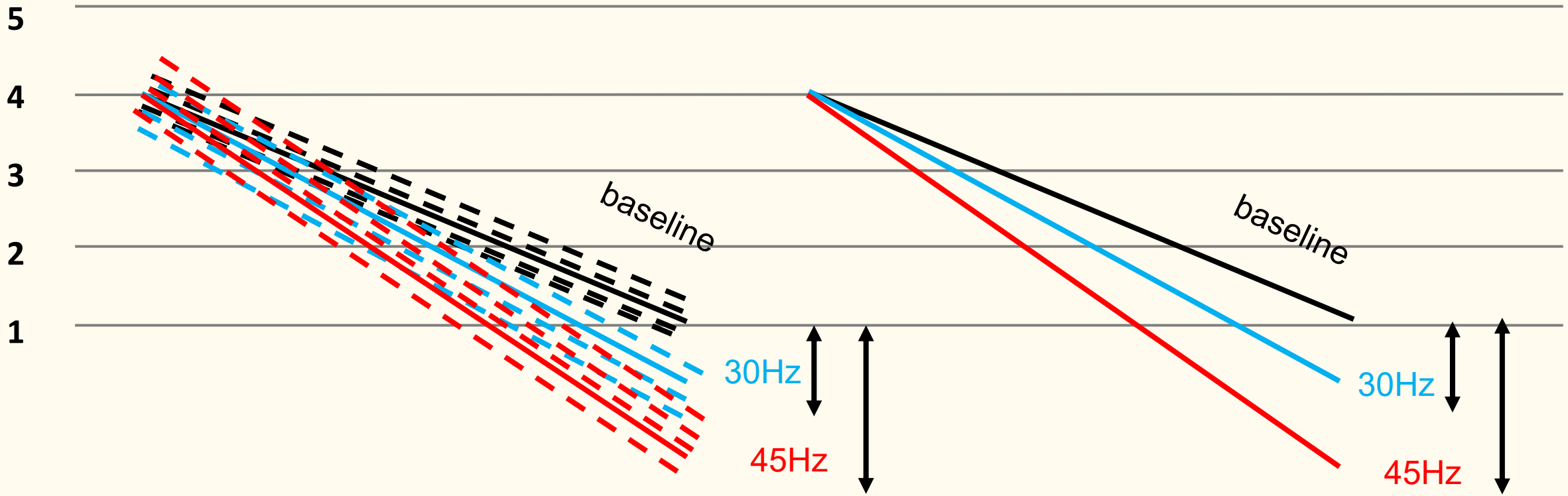
- Chandrasekaran, B., Sampath, P. D., & Wong, P. C. M. (2010). Individual variability in cue-weighting and lexical tone learning. *The Journal of the Acoustical Society of America*, *128*(1), 456–465. https://doi.org/10.1121/1.3445785
- Francis, A. L., & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 349–366. https://doi.org/10.1037/0096-1523.28.2.349
- Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics*, *11*(2), 149–175. https://doi.org/10.1016/S0095-4470(19)30813-7
- 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. https://doi.org/10.1121/1.412280
- 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. https://doi.org/10.1016/j.wocn.2010.08.003
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979–1000. https://doi.org/10.1098/rstb.2007.2154
- McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L. (2002). Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(2), 89–108. https://doi.org/10.3758/CABN.2.2.89
- Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science*, *271*(5245), 77–81.
- Pashler, H., & Mozer, M. C. (2013). When does fading enhance perceptual category learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(4), 1162–1173. https://doi.org/10.1037/a0031679
- Rvachew, S., & Jamieson, D. G. (1989). Remediating speech production errors with sound identification training. *Journal of Speech-Language Pathology and Audiology*, *16*, 201–208.

Demographic information: age



For all tonal stimuli: F0 variation

Adding F0 height variability



BIS: measurement

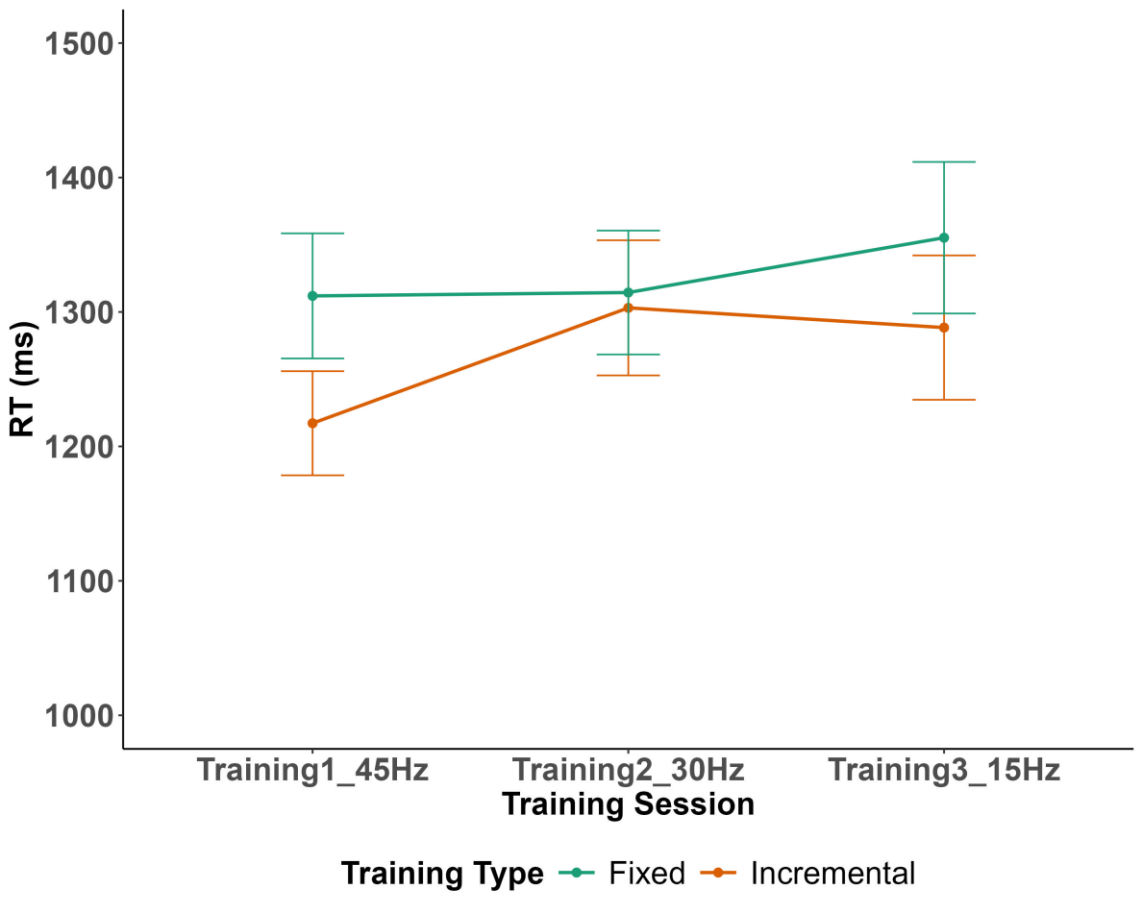
$$BIS_{i,j} = Z_{PC_{i,j}} - Z_{\overline{RT_{i,j}}}$$

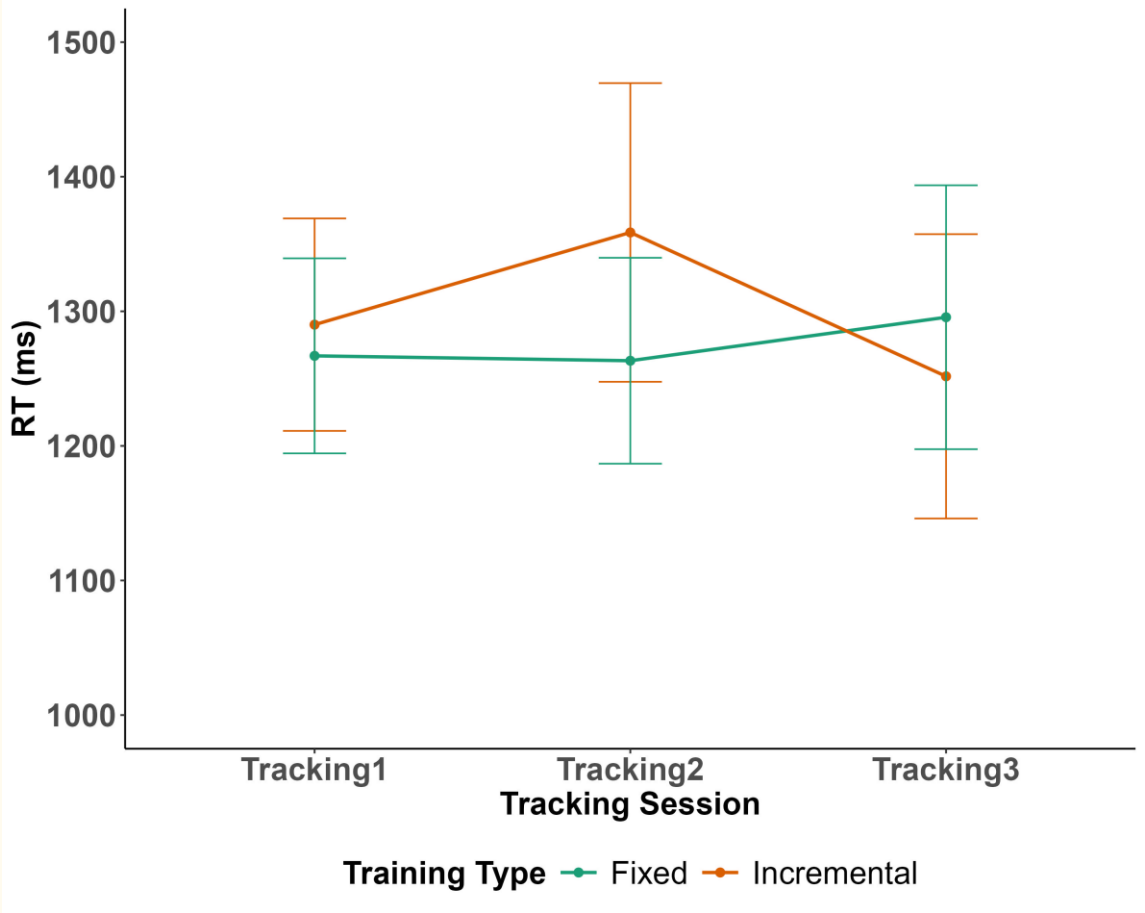
Where $Z_{x_{i,j}} = \frac{x_{i,j} - \bar{x}}{S_x}$ (S_x = sample standard deviation).

the means and sample standard deviation, i.e., $\overline{RT}$, $\overline{PC}$, SRT, and SPC, were calculated across all observed cells from the analysed experiment regardless of subjects and conditions. Therefore, positive values indicate good performance, i.e., higher accuracy with faster responses, while negative values indicate lower accuracy with longer responses. BIS values also directly express whether a group of participants performs above (BIS > 0) or below (BIS < 0) average.

Other effective measures included Rate-correct Score (RCS), Inverse Efficiency Score (IES), and Linear Integrated Speed-accuracy Score (LISAS) (Chignell et al., 2014; Liesefeld & Janczyk, 2019; Vandierendonck, 2017).

RT: training and tracking-tests

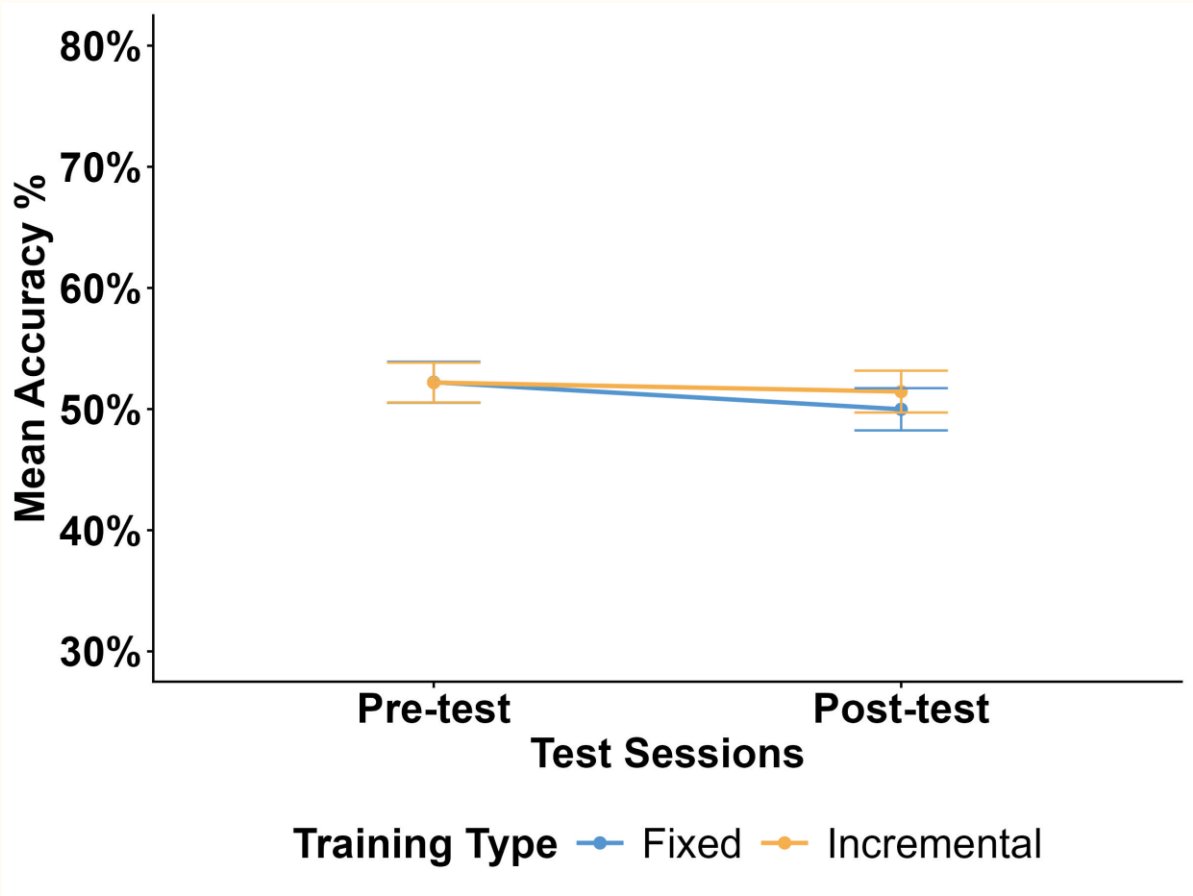




no effect of Training Type, $\chi2(1) = 0.000$, $p = .998$, Training Session, $\chi2(2) = 2.94$, $p = .23$, or Training Session x Training Type, $\chi2(2) = 2.85$, $p = .24$.

no effect of Tracking-test, $\chi2(2) = 1.58$, $p = .45$, Training Type, $\chi2(1) = 0.19$, $p = .66$, or Tracking-test x Training Type, $\chi2(2) = 0.82$, $p = .66$

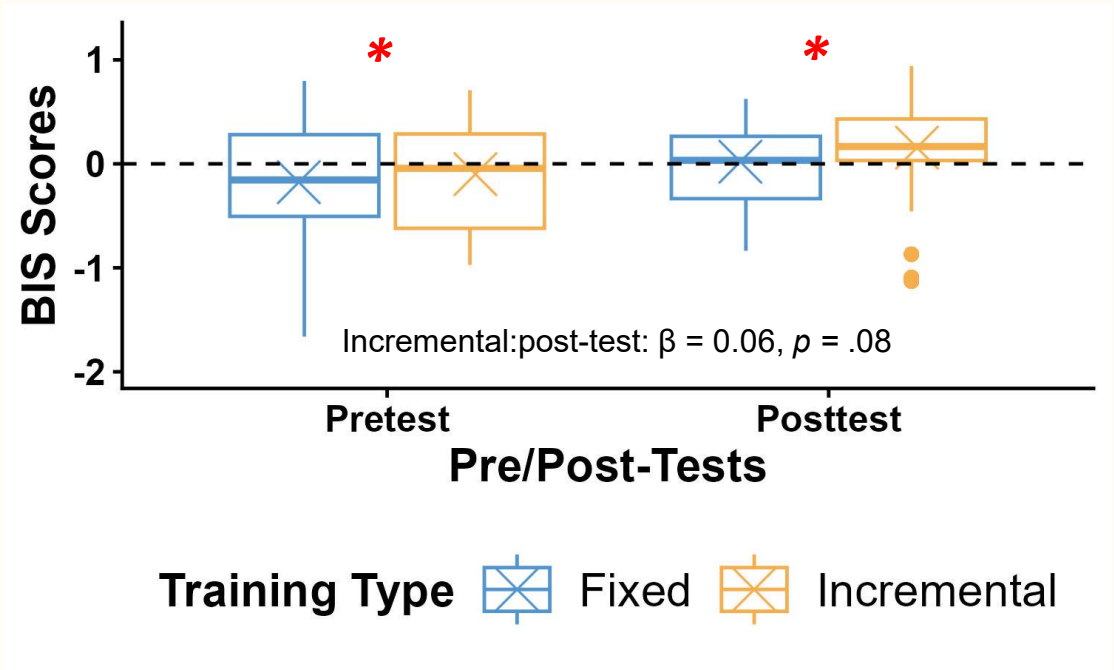
Results in pre-/post-tests: Accuracy



Pre-/post-tests: No main/interaction effects

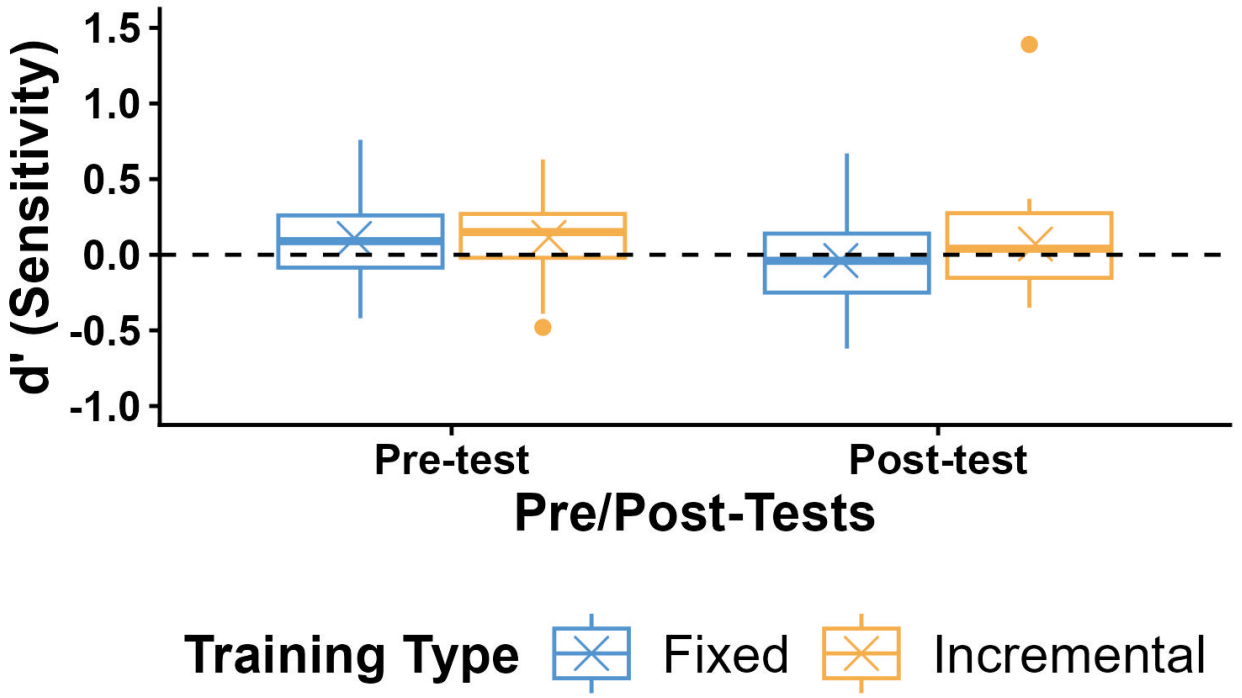
BIS & Sensitivity scores (d')

in pre-/post-tests



Pre-/post-test: BIS scores

Training Type x Test Session



Pre-/post-test: Sensitivity scores

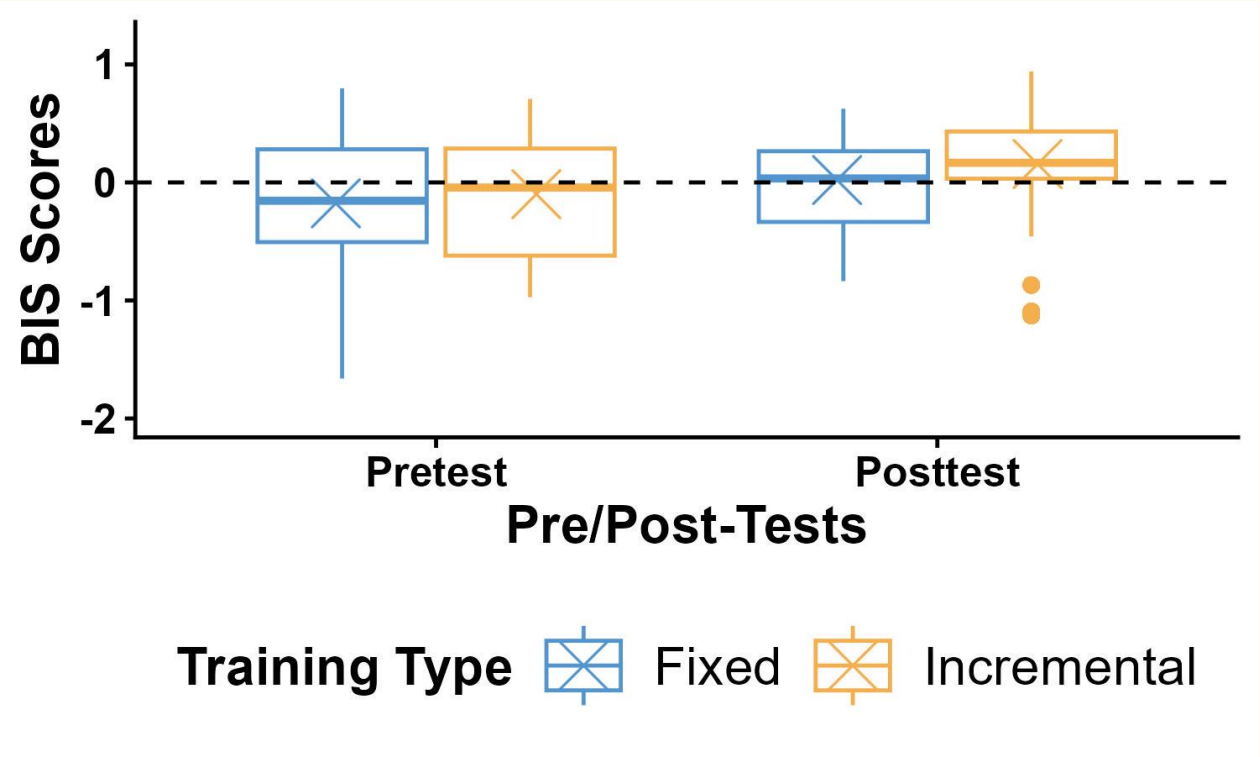
Test Session: *decrease in fixed group, p = .07*

BIS pre- to post-test

Best-fit model: lm(bis ~ training_type + session + training_type:session).

Both groups improved BIS from pre- to post-test; Incremental >> fixed in both tests.

"Incremental:Posttest" suggested that the Incremental group had slightly more improvement than the Fixed group after training, β = 0.06, SE = 0.03, p = .08

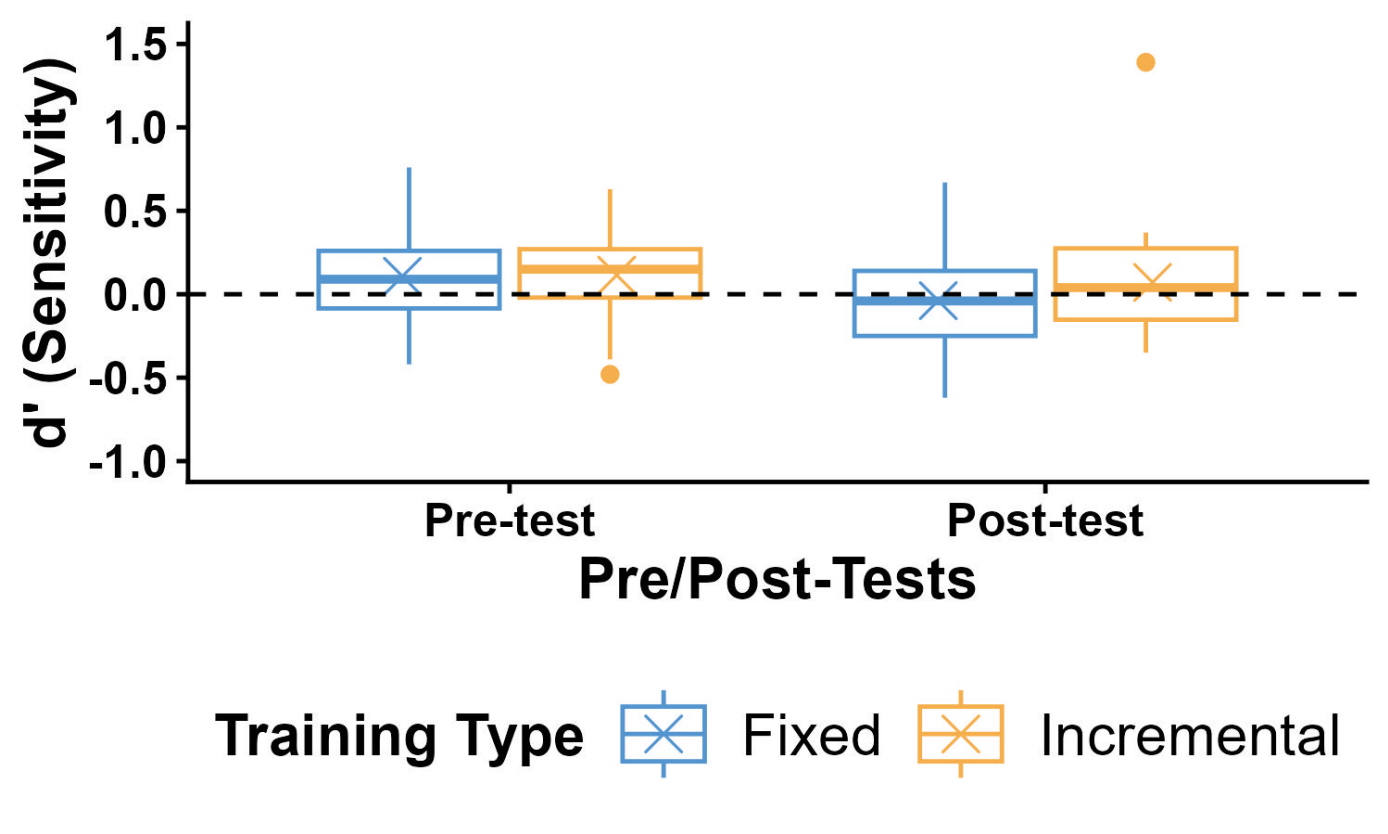


Sensitivity pre- to post-test

marginal effect of Test Session, F(1, 0.26) = 2.88, p = .09

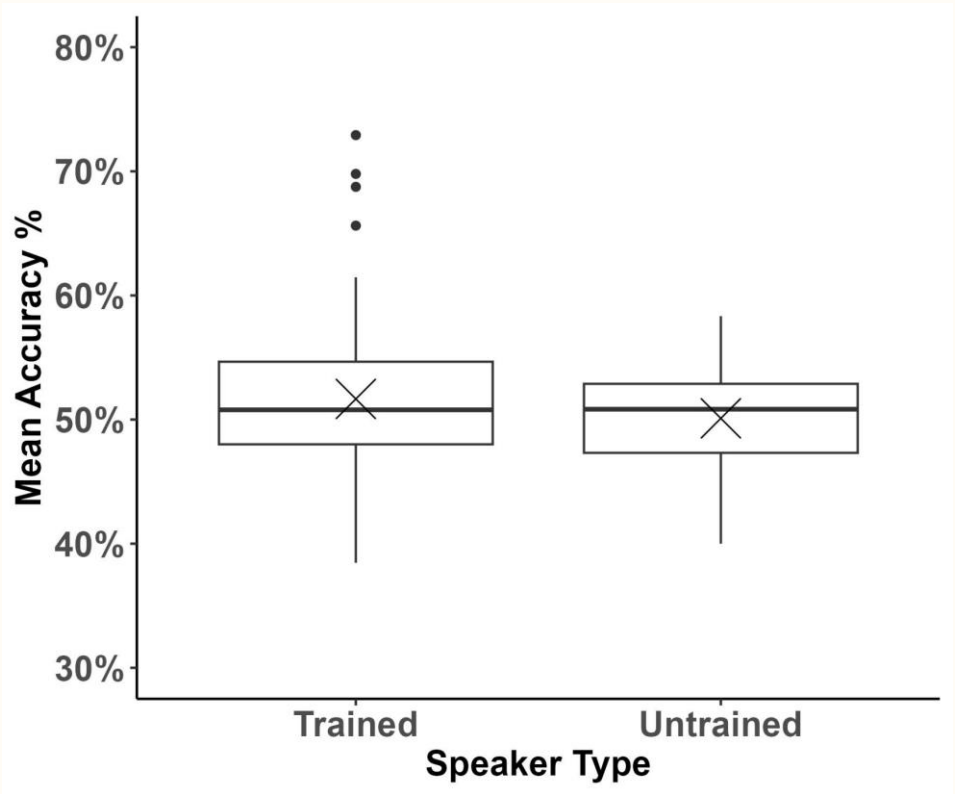
between-session pairwise comparisons: reduced sensitivity after training for the Fixed group, F(1, 0.31) = 3.44, p = .07

Others: non-significant



Results in generalisation: Accuracy

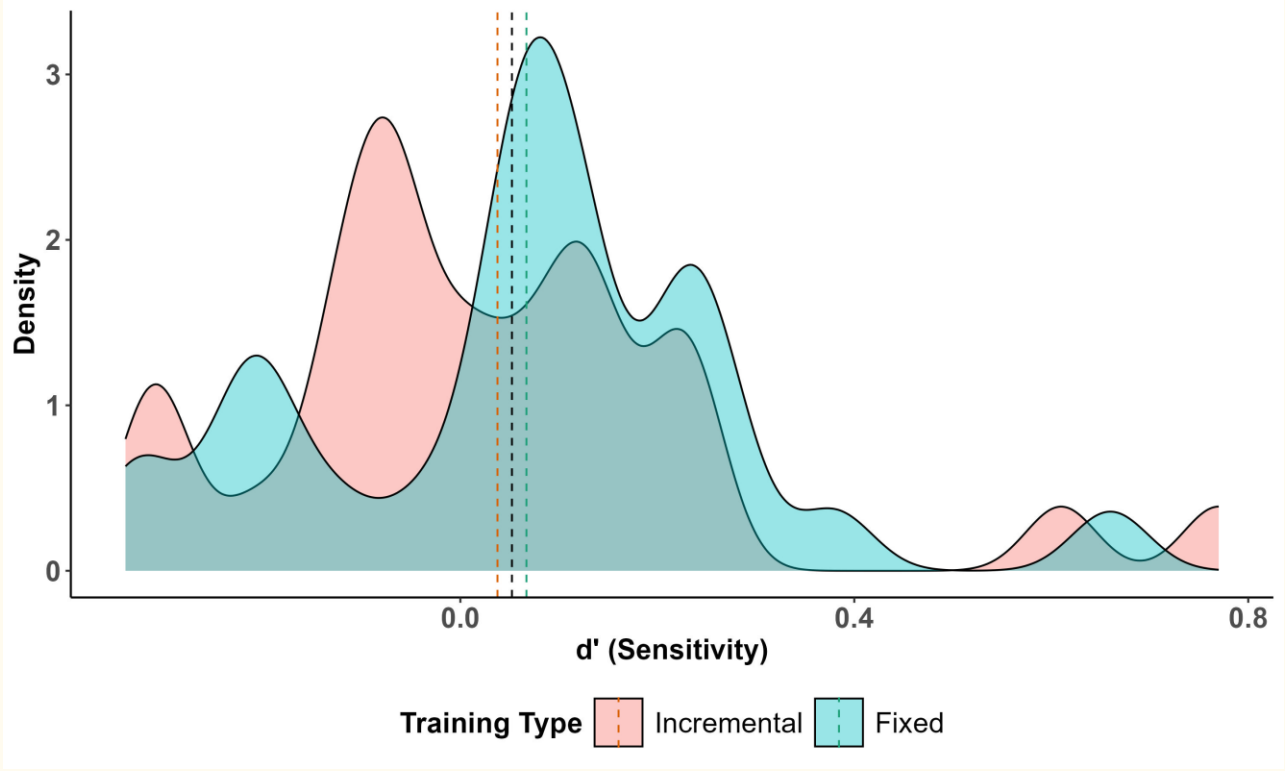
trained speaker *with untrained stimuli*
untrained speaker *with trained stimuli*



Generalisation-test: speaker (p = .06)

Generalisation test: BIS & sensitivity

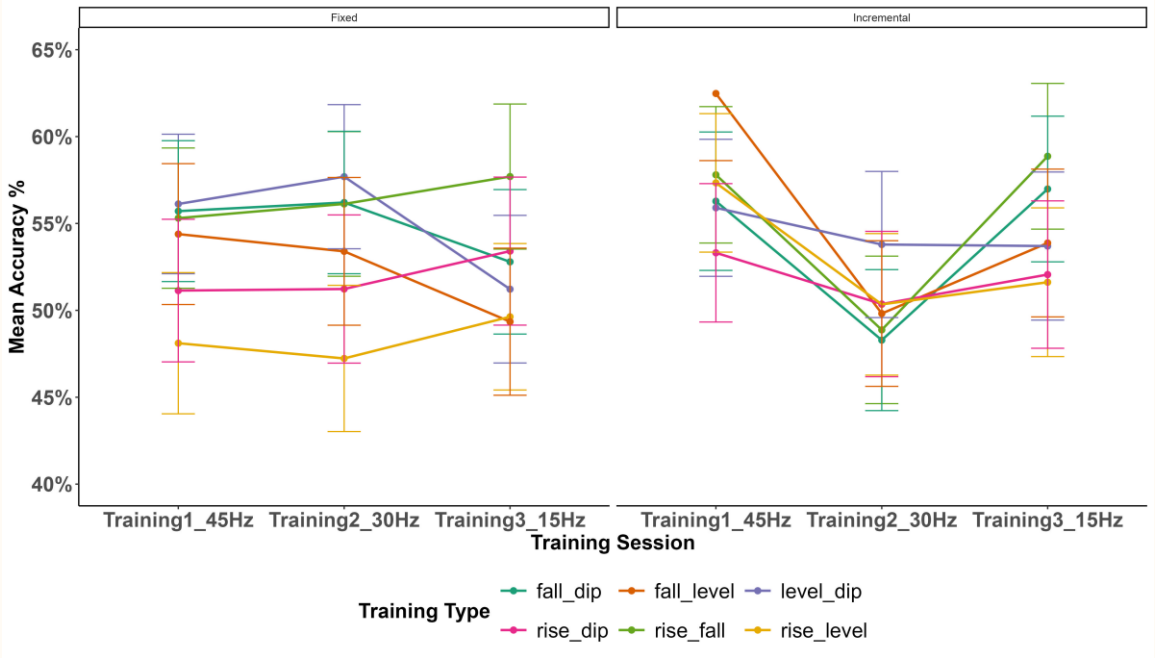
- Non-significant for BIS



- Training Type: F(1, 0.68) = 13.32, p < .001
  - Fixed > incremental

The effect of Tonal Contrast in training

- For the effect of Tonal Contrast on response accuracy in training, model comparisons found a main effect, $\chi2(5) = 14.72$, $p = .01$. There were also marginal interaction effects of Tonal Contrast x Training Session, $\chi2(10) = 18.12$, $p = .05$, and Tonal Contrast x Training Type, $\chi2(5) = 10.53$, $p = .06$.



Pairwise comparisons	Test statistics
Training 1	
falling/level vs rising/dipping	*β* = 0.26, *SE* = 0.08, *p* = .02
falling/level vs rising/level	*β* = 0.24, *SE* = 0.08, *p* = .05
Training 2	
level/dip vs rising/level	*β* = 0.29, *SE* = 0.09, *p* = .01
Training 3	
falling/level vs rising/fall	*β* = -0.27, *SE* = 0.09, *p* = .02
rising/falling vs rising/level	*β* = 0.31, *SE* = 0.09, *p* = .005

Table 6-17 Pairwise EMMs test statistics for the effect of Tonal Contrast and Training Session on accuracy data for the Fixed group in training sessions (Study 5).