To voice or not to voice? Cross-linguistic effects on phonological representations

Maud Pélissier, Helene S. Øya, Linda Wheeldon, Allison Wetterlin

Université Paris Cité, CLILLAC-ARP, Paris France

University of Agder, Kristiansand, Norway

Université Paris Cité

UiA University of Agder

Research questions

- L2 processing of contrasts that do or do not exist in the L1
- How are the contrasts perceived and produced?
- Is there a link between perception and production?
- What are the consequences for lexical access?

- Norwegian learners of English
- Two voicing contrasts: /t-d/ (similar L1-L2) and /s-z/ (L2-specific)
- 4 tasks: categorisation task, lexical decision task, MMN, picture naming

Background – Perceiving non-native contrasts

- Adult learners often have **difficulties** perceiving contrasts that do not exist in their L1 (Best & McRoberts, 2003)
  - E.g., [s] vs [z] for Danish and Swedish speakers (Bohn & Ellegaard, 2019; Flege & Hillenbrand, 1986)
- Models of L2 speech perception predict that the **starting point for L2 phonology is L1 phonology** (PAM-L2, Best & Tayler, 2007; r-SLM, Flege & Bohn 2021; L2LP, Escudero, 2005; Van Leussen & Escudero, 2015)
- The development of new categories depends on a range of factors including how the L2 contrast maps onto L1 categories and L2 proficiency

Background – Asymmetries in phonological representations

- There are **asymmetries** in the perception of some phonological features pertaining to their specification in their representations (e.g., Lahiri & Reetz, 2010)

- The case for **voicing** is complex especially in Germanic languages – what cues are used? What is represented?
- **Cues differ** in languages for voicing in stops
  - Norwegian has aspiration for voicelessness and pre-voicing for voicing
  - English lacks pre-voicing and relies on VOT for voicing
  - Fricative voicing contrasts do not occur in Norwegian

Background – Asymmetries in phonological representations

- Models differ in terms of:
  - What they say is represented
  - What features can be specified
  - What features are used as cues
- Hypothesis 1 (Laryngeal phonology): [SPREAD GLOTTIS] is specified in English (Avery & Idsardi, 2001)
  - Supported by MMN studies showing underspecification of voicing in English (Hestvik & Durvasula, 2016; Monahan et al., 2022)
- Hypothesis 2 (FUL model): Voicing is specified in English (Lahiri & Reetz, 2010)
  - [SPREAD GLOTTIS] is a predictable cue for voicelessness in English and doesn't need to be specified
  - Supported by role of voicing in phonological & morphophonological processes
  - Behavioural studies on asymmetries (Hwang et al, 2010)
  - MMN evidence for Japanese (Hestvik et al., 2020), and Danish (Højlund et al., 2019)

Methods

1 – Picture naming task
2 – Lexical decision task
3 – MMN task
4 – ABX categorisation task

34 L1 Norwegian, L2 English (>B2)

ABX categorisation task

- 24 minimal pairs of pseudowords (12 for /s-z/, 12 for /t-d/, CVCVC structure)
- Choose if 3rd pseudoword matches 1st or 2nd one

Lexical Decision task

- 48 word-nonword pairs for each phoneme (e.g., /ˈpɔɪzn/ - */ˈpɔɪsn/ ‘poison’)
- 96 pairs with manipulated vowel contrast as distractors
- 2 lists – only heard one member of the pair

Picture naming task

- 48 pictures
- 12 words per target phoneme (/s/, /z/, /t/, /d/)
- In initial, medial, final position (4 each)
- Named 3 times

MMN oddball task

- VCV pseudowords (/ɑ:tɑ:/, /ɑ:dɑ:/, /ɑ:sɑ:/, /ɑ:zɑ:/)
- Oddball paradigm with 4 standards / 1 deviant
- 4 blocks (s-deviant among z-standard, z-deviant among s-standard, etc.)

Is the L2-specific contrast (/s-z/) more difficult to process?

YES

- ABX task: effect of Contrast (z = -3.94, p < .001)
- LDT: effect of Contrast (z = 10.38, p < .001)

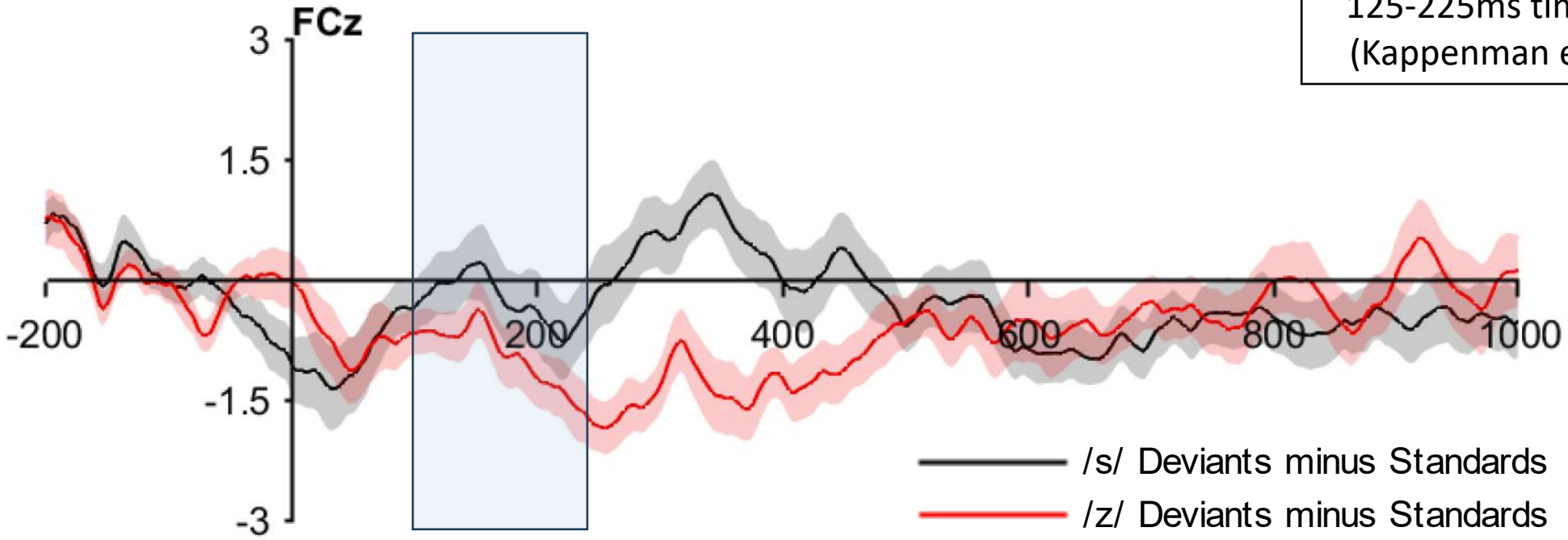
	/s-z/ contrast	/t-d/ contrast
ABX Accuracy rate	75.08%	87.91%
LDT Accuracy rate on nonwords	23.49%	64.64%

Are there asymmetries in the phonological representations in L2? – MMN

YES/NO

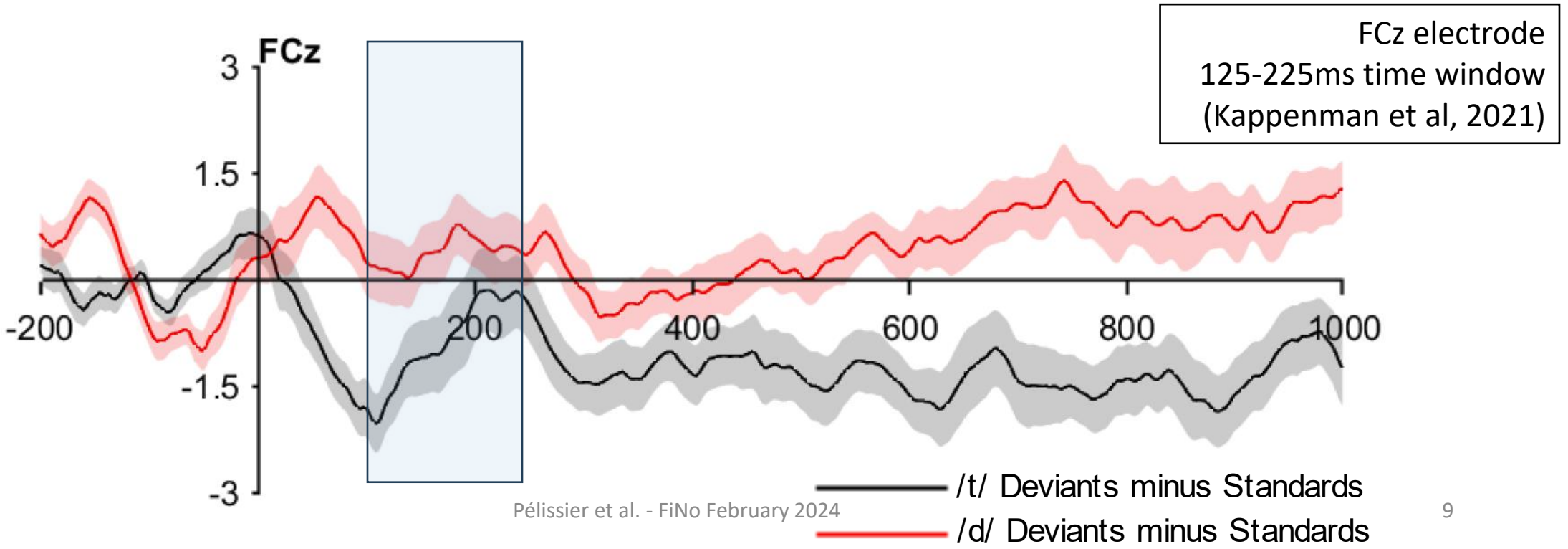
- For the /s-z/ contrast
  - MMN to both
  - No asymmetry between the s-MMN and the z-MMN

FCz electrode
125-225ms time window
(Kappenman et al, 2021)



Are there asymmetries in the phonological representations in L2? – MMN

- For the /t-d/ contrast
  - There is a significant effect of phoneme ($p < .001$)
  - The t-MMN is larger than the d-MMN (MMN to deviant /t/ only)



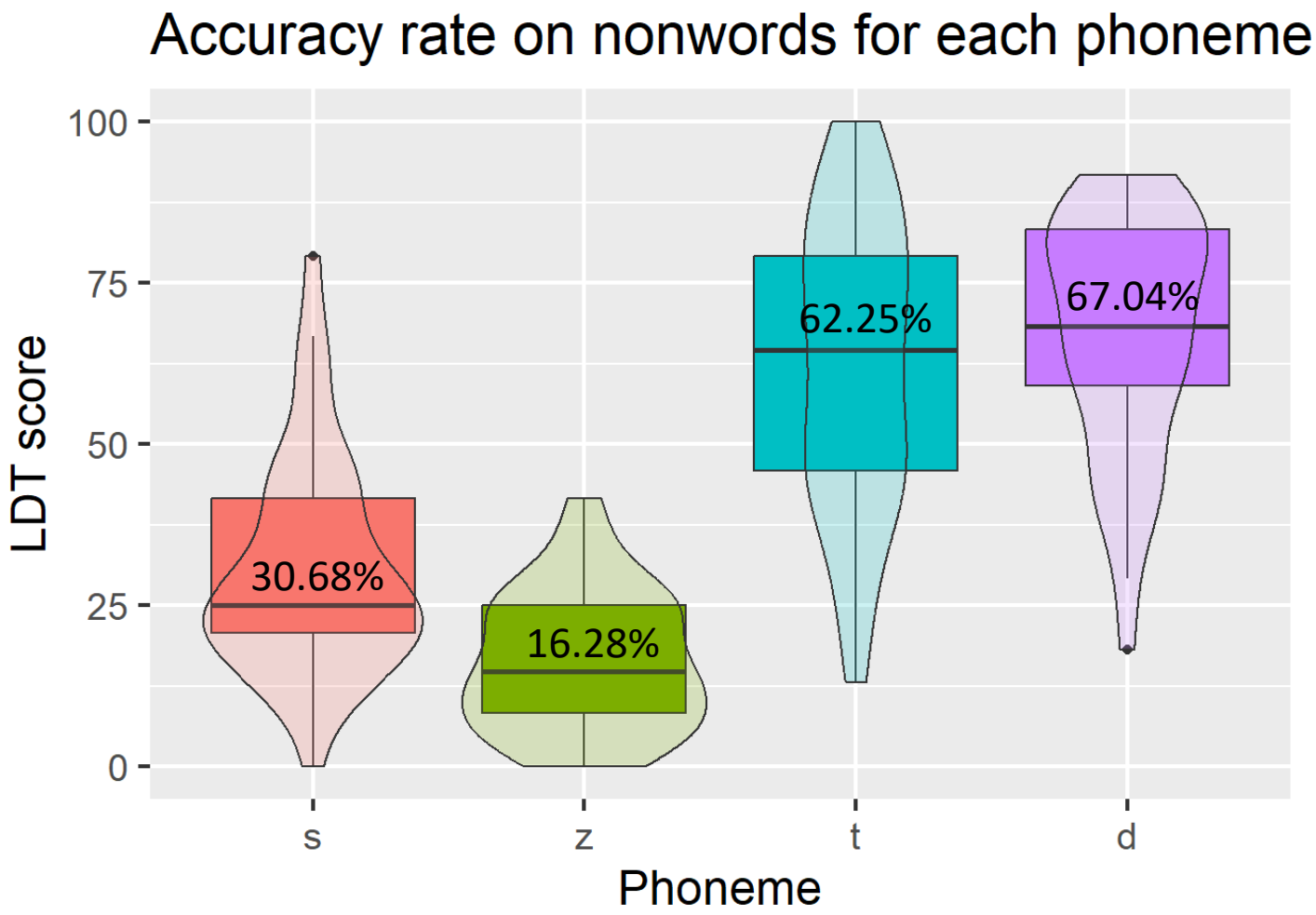
Are there asymmetries in the phonological representations in L2? – MMN

- /t-d/ contrast
  - This is in line with what was observed with Danish speakers (Højlund et al 2019)
  - And is consistent with [+VOICE] being specified in English
  - But the opposite pattern of English native speakers (Hestvik & Durvasula 2016; Monahan et al, 2022)
- /s-z/ contrast
  - No asymmetry: no representation of /z/ with specified voicing that could lead to a mismatch
  - But voicing is in the lexicon for other contrasts: the cues can be used to perceive the contrast

Are there asymmetries in the phonological representations in L2 at the lexical level? - LDT

YES/NO

- Contrast x Voicing interaction ($\chi^2 = 10.99$, $p < .001$)
- Asymmetry for the /s-z/ contrast ($p < .001$)
- No asymmetry for the /t-d/ contrast ($p = .29$)

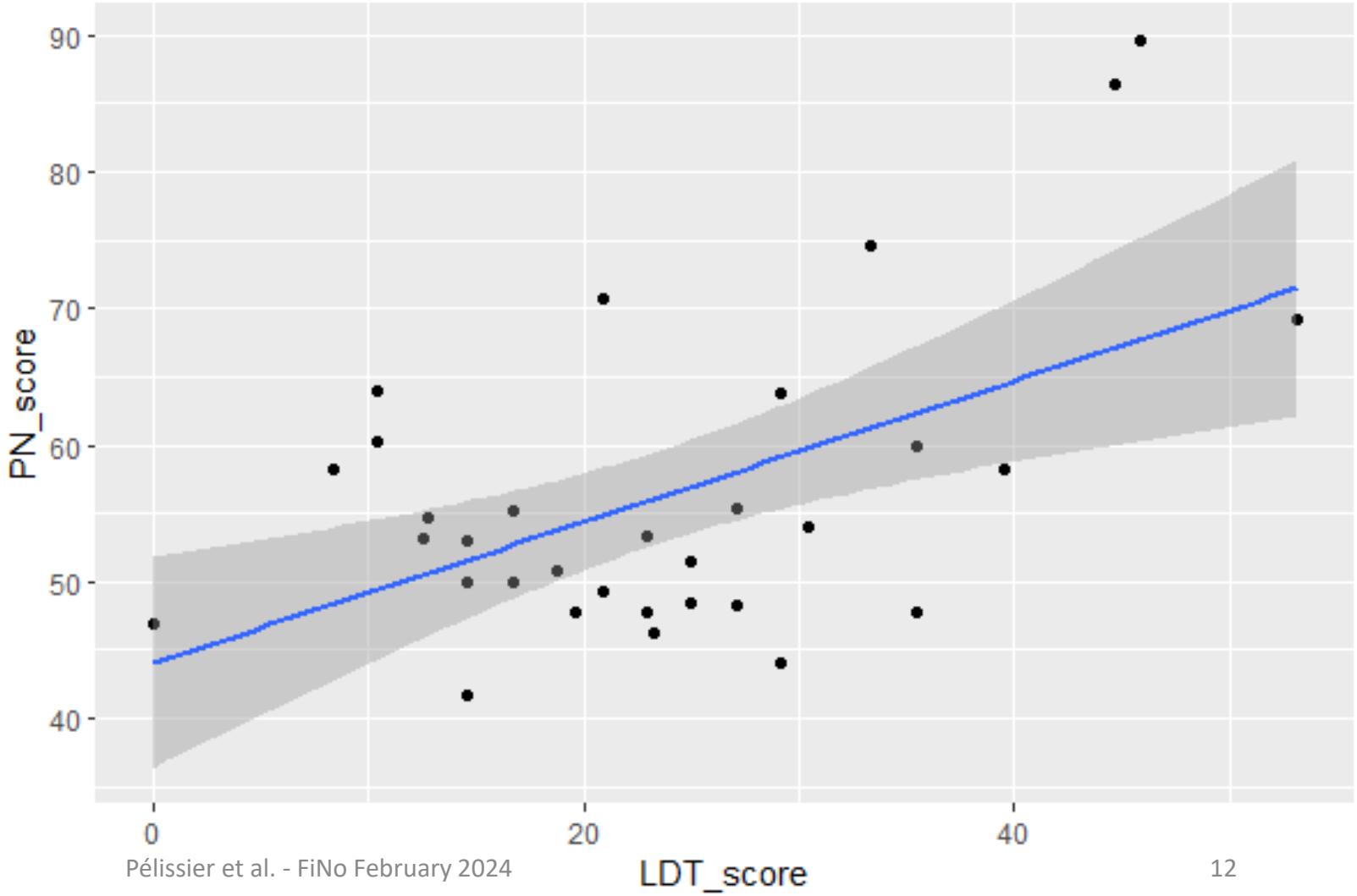


Is there a link between perception and production for the /s-z/ contrast?

YES

Production performance is predicted by :

- the LDT score (*β* = 5.67, *p* = .001)
- proficiency (LexTALE, *β* = 3.52, *p* = .03)
- only marginally by the ABX score (*β* = 3.06, *p* = .066)



Discussion

- Even highly proficient L2 learners rely heavily on L1 phonology
- Evidence of development towards sensitivity to L2 contrasts
- But persisting difficulties with the lexicalisation of the L2-specific contrast
- What matters most for production is lexicalisation (LDT score & vocabulary), not low-level perception (ABX)

Future steps

- Work in progress
- Investigating voicing asymmetries in the MMN and behavioural tasks with native English speakers

Thank you for your attention!

References

Avery, P., & Idsardi, W. J. (2001). Laryngeal dimensions. Completion and enhancement. In T. A. Hall (Ed.), *Distinctive Feature Theory* (pp. 41–70). Walter de Gruyter.

Best, C. C., & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech*, *46*(Pt 2-3), 183–216. https://doi.org/10.1177/00238309030460020701

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Second Language Speech Learning: The Role of Language Experience in Speech Perception and Production* (pp. 13–34). John Benjamins.

Bohn, O.-S., & Ellegaard, A. A. (2019). Perceptual assimilation and graded discrimination as predictors of identification accuracy for learners differing in l2 experience: The case of Danish listeners'perception of English initial fricatives. *International Congress of Phonetic Sciences*. https://www.researchgate.net/publication/332715269

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. LOT.

Flege, J. E., & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3–83). Cambridge University Press.

Flege, J. E., & Hillenbrand, J. (1986). Differential use of temporal cues to the /s/-/z/ contrast by native and non-native speakers of English. *Journal of the Acoustical Society of America*, *79*(2), 508–517. https://doi.org/10.1121/1.393538

References (continued)

Hestvik, A., & Durvasula, K. (2016). Neurobiological evidence for voicing underspecification in English. *Brain and Language*, *152*, 28–43. https://doi.org/10/f77qfm

Hestvik, A., Shinohara, Y., Durvasula, K., Verdonschot, R. G., & Sakai, H. (2020). Abstractness of human speech sound representations. *Brain Research*, *1732*, 146664. https://doi.org/10/grn7r7

Højlund, A., Gebauer, L., McGregor, W. B., & Wallentin, M. (2019). Context and perceptual asymmetry effects on the mismatch negativity (MMNm) to speech sounds: An MEG study. *Language, Cognition and Neuroscience*, *34*(5), 545–560. https://doi.org/10/grn7r8

Hwang, S.-O. K., Monahan, P. J., & Idsardi, W. J. (2010). Underspecification and asymmetries in voicing perception. *Phonology*, *27*(2), 205–224. https://doi.org/10.1017/S0952675710000102

Kappenman, E. S., Farrens, J. L., Zhang, W., Stewart, A. X., & Luck, S. J. (2021). ERP CORE: An open resource for human event-related potential research. *NeuroImage*, 225, 117465.

Lahiri, A., & Reetz, H. (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics*, *38*(1), 44–59. https://doi.org/10.1016/j.wocn.2010.01.002

Monahan, P. J., Schertz, J., Fu, Z., & Pérez, A. (2022). Unified Coding of Spectral and Temporal Phonetic Cues: Electrophysiological Evidence for Abstract Phonological Features. *Journal of Cognitive Neuroscience*, *34*(4), 618–638. https://doi.org/10.1162/jocn_a_01817

Van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, *6*. https://doi.org/10/gsqb3p