A Quantitative Model Of The Language Familiarity Effect In Infancy

Craig Thorburn, Naomi Feldman, Thomas Schatz | NECPhon 2019

UNIVERSITY OF
MARYLAND

The Language Familiarity Effect

English Learner *(native)* ✓

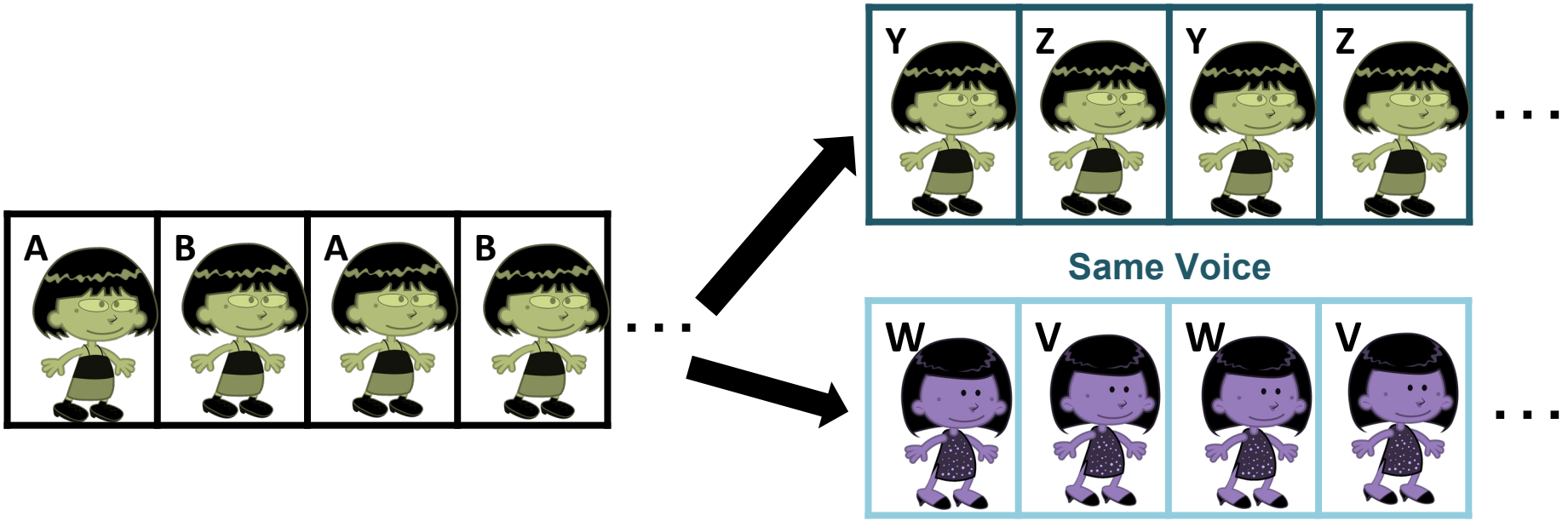
Japanese Learner *(non-native)* ✗

English Learner *(non-native)* ✗

Japanese Learner *(native)* ✓

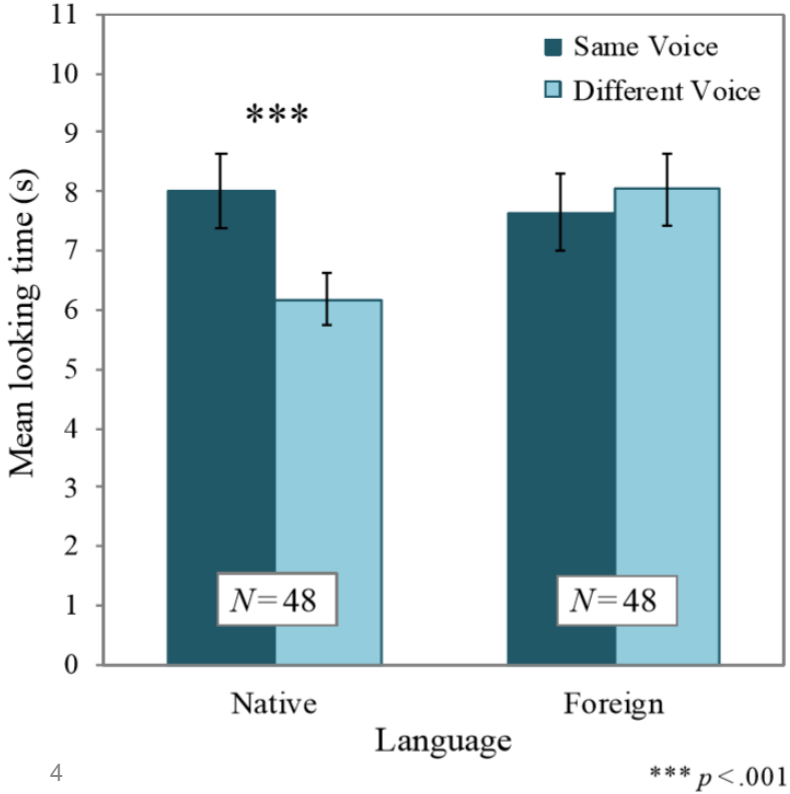
Johnson, Westreck & Cutler 2011; Fecher & Johnson 2019

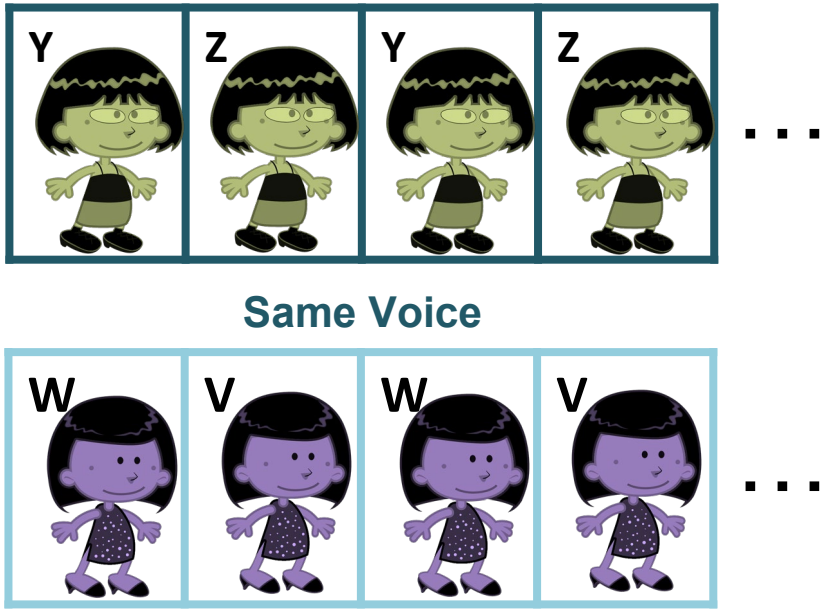
Visual Fixation Procedure



Fecher & Johnson 2019

Visual Fixation Procedure





Fecher & Johnson 2019

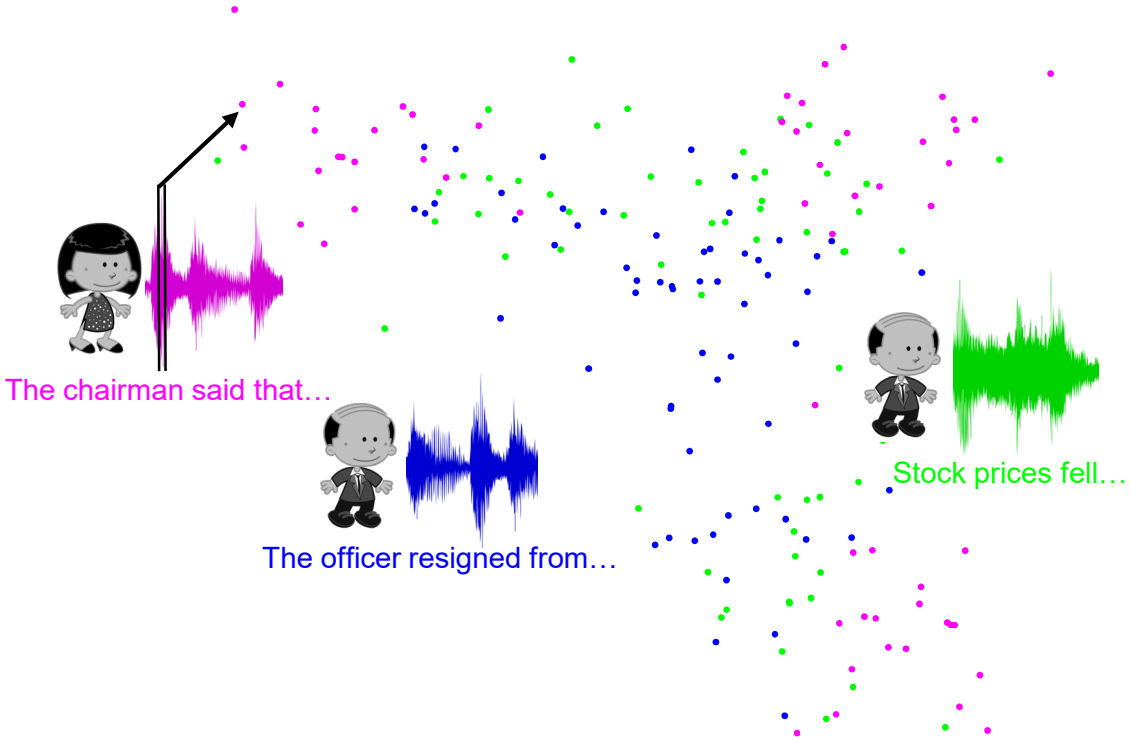
How could infants learn this?

- We can look to adults
  - Adults are also known to show a language familiarity effect
  - Thought to require some abstract phonological knowledge or comprehension
- We can model this using acoustic properties of input
  - We will look at higher and lower level features at two timescales

Short timescale
e.g. Phonetic information

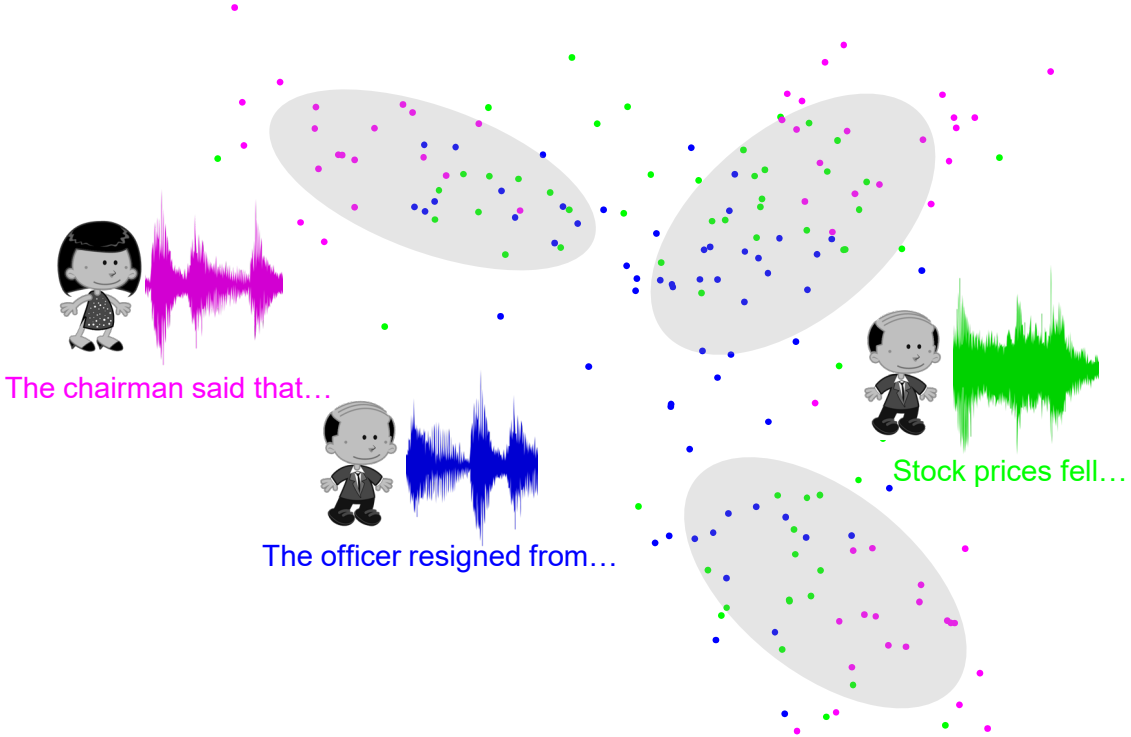
Long timescale
e.g. Speaker information

Short Timescale – Phonetic Information



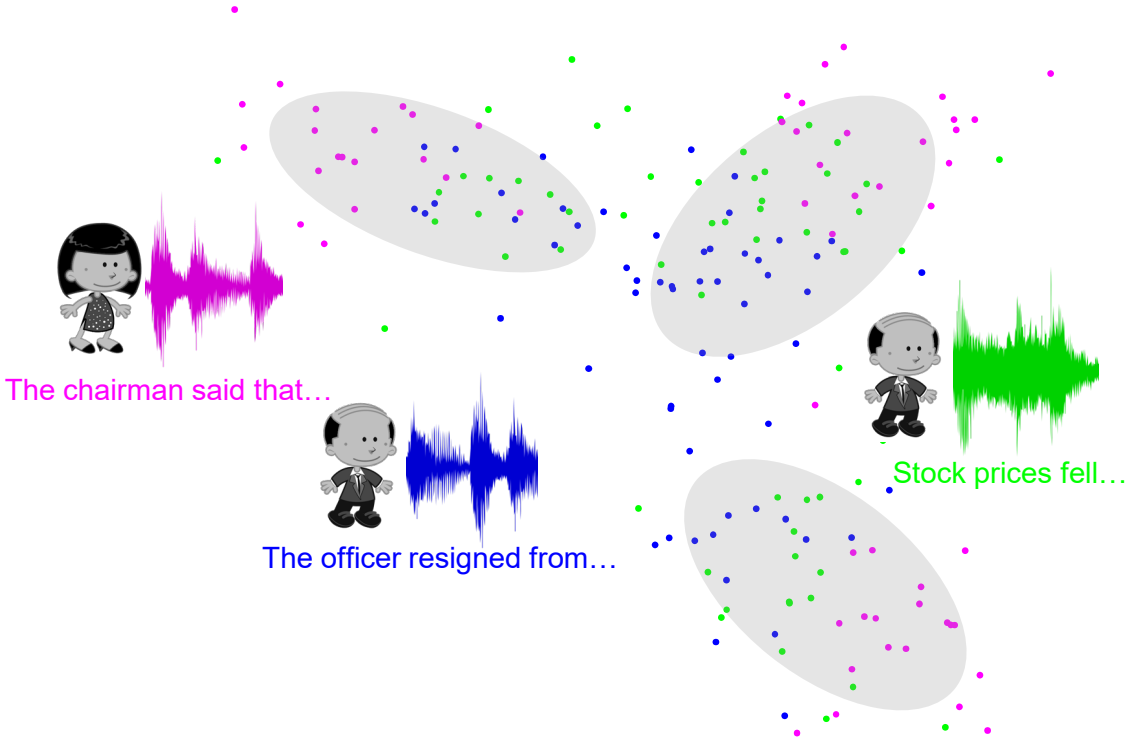
Schatz et al 2019

Short Timescale – Phonetic Information



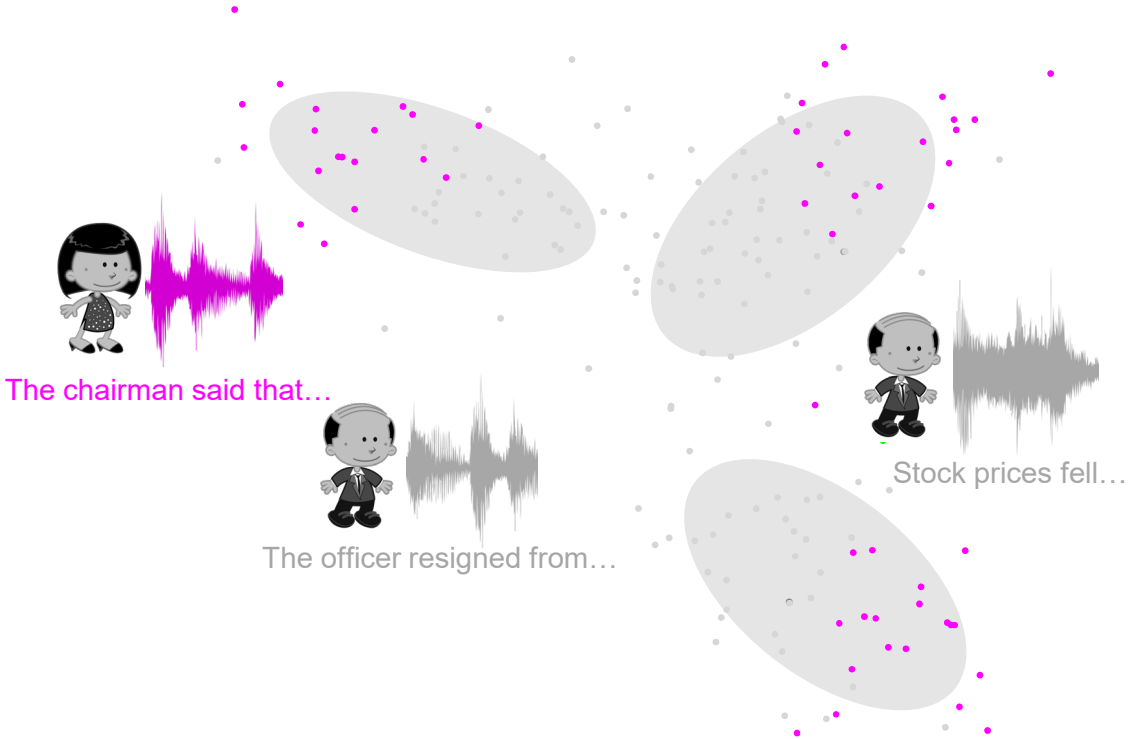
Schatz et al 2019

Long Timescale – Speaker Information



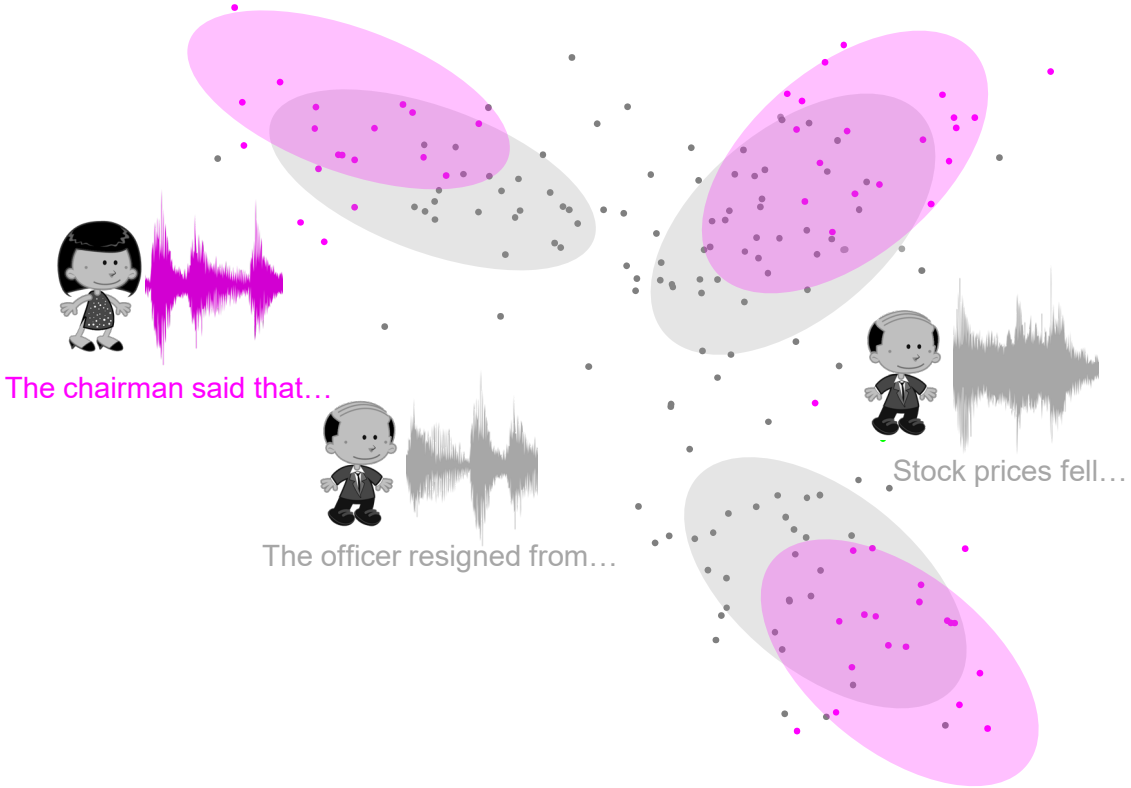
Dehak et al. 2010

Long Timescale – Speaker Information



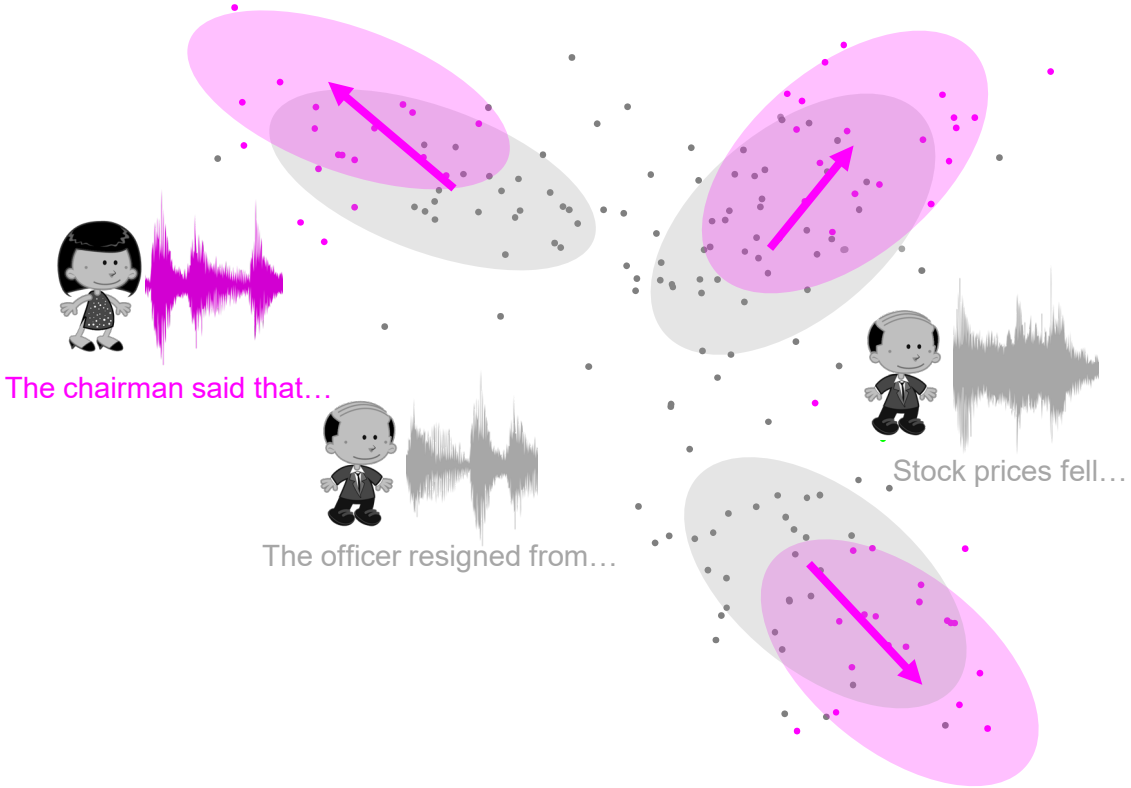
Dehak et al. 2010

Long Timescale – Speaker Information



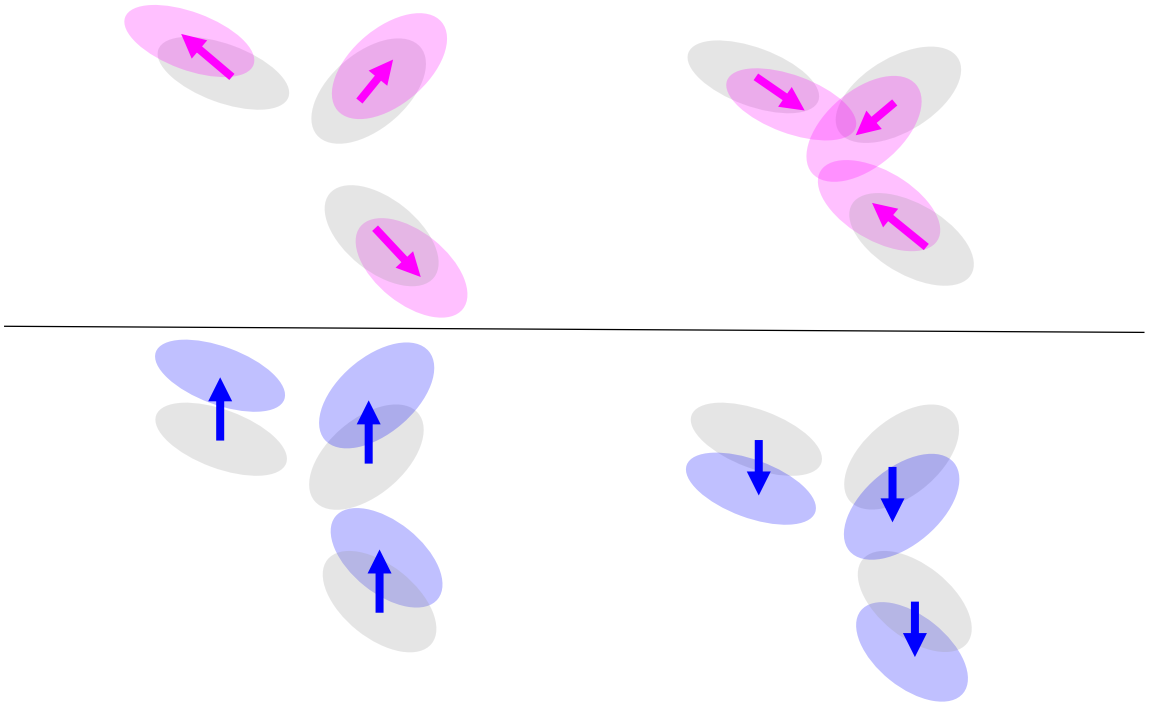
Dehak et al. 2010

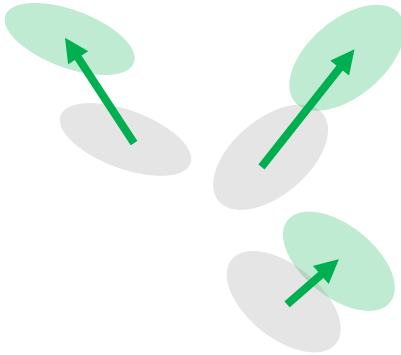
Long Timescale – Speaker Information



Dehak et al. 2010

Long Timescale – Speaker Information

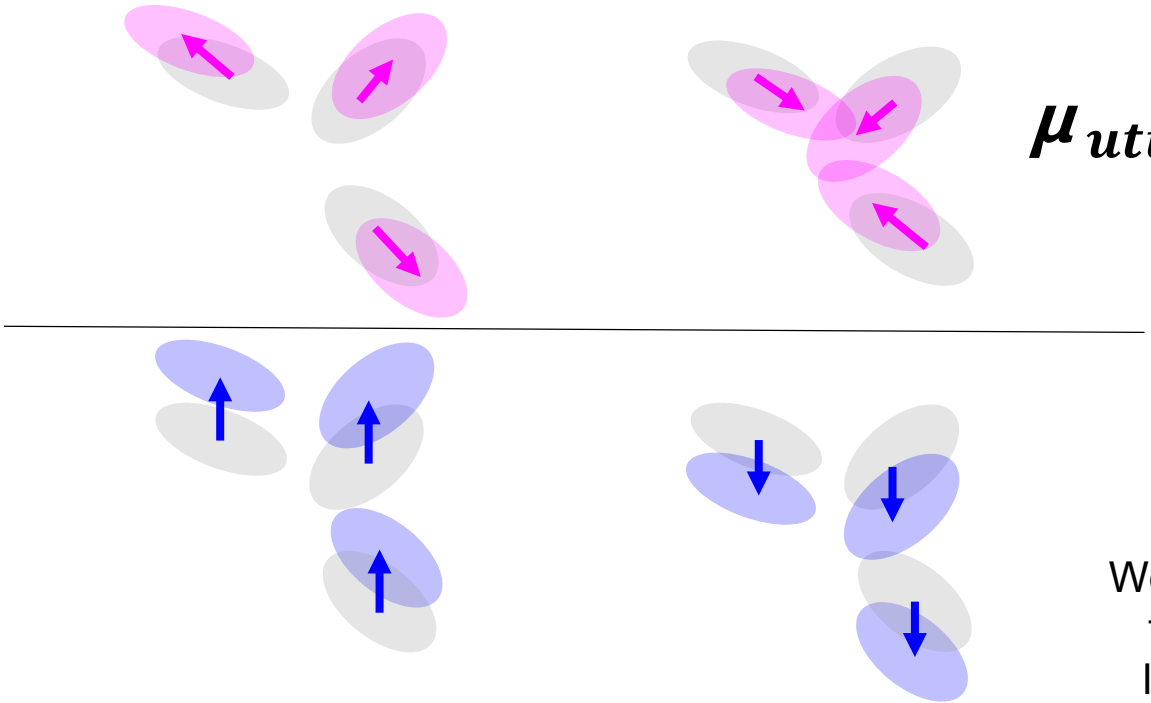




$$I_{utt} = \begin{bmatrix} 0.4 \\ \\ \\ 0.7 \end{bmatrix}$$

Dehak et al. 2010

Long Timescale – Speaker Information



$$\boldsymbol{\mu}_{utt} = \boldsymbol{\mu} + T * I_{utt}$$

$$I_{utt} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ \\ \dots \\ \\ I_{400} \end{bmatrix}$$

Weight matrix learned through maximum likelihood mapping

Dehak et al. 2010, Carbajal et al. 2016

Short timescale
Phonetic information

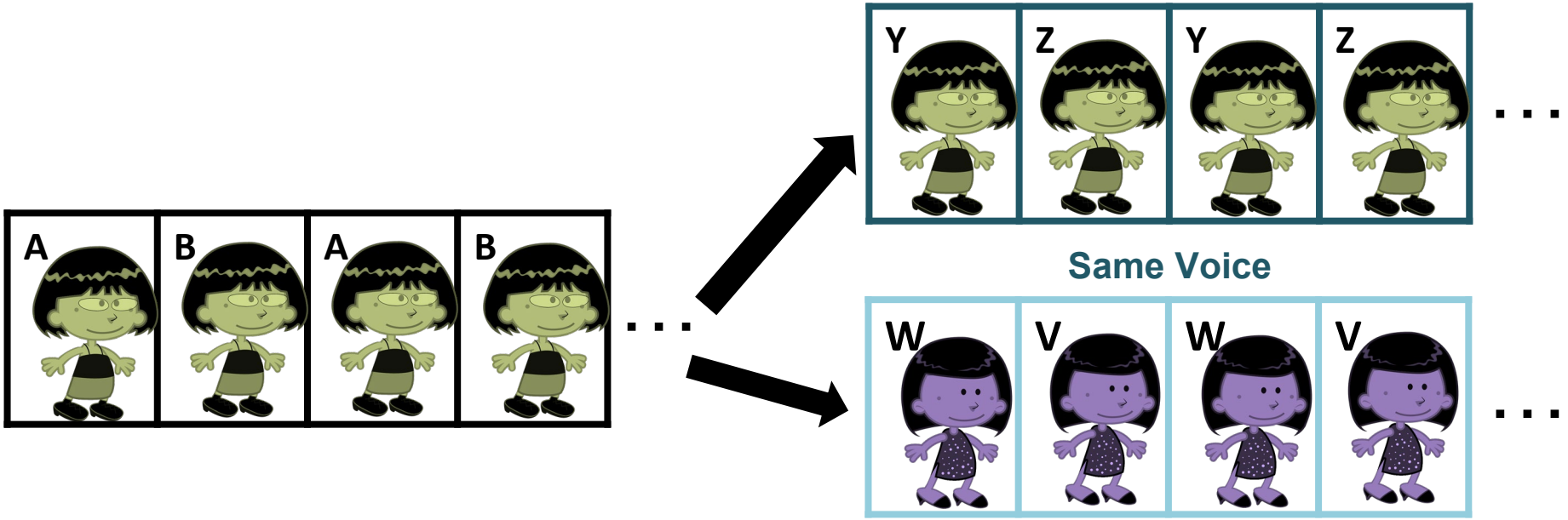
Frame-level Gaussian Activations

Long timescale
Speaker information

Utterance-level I-Vectors

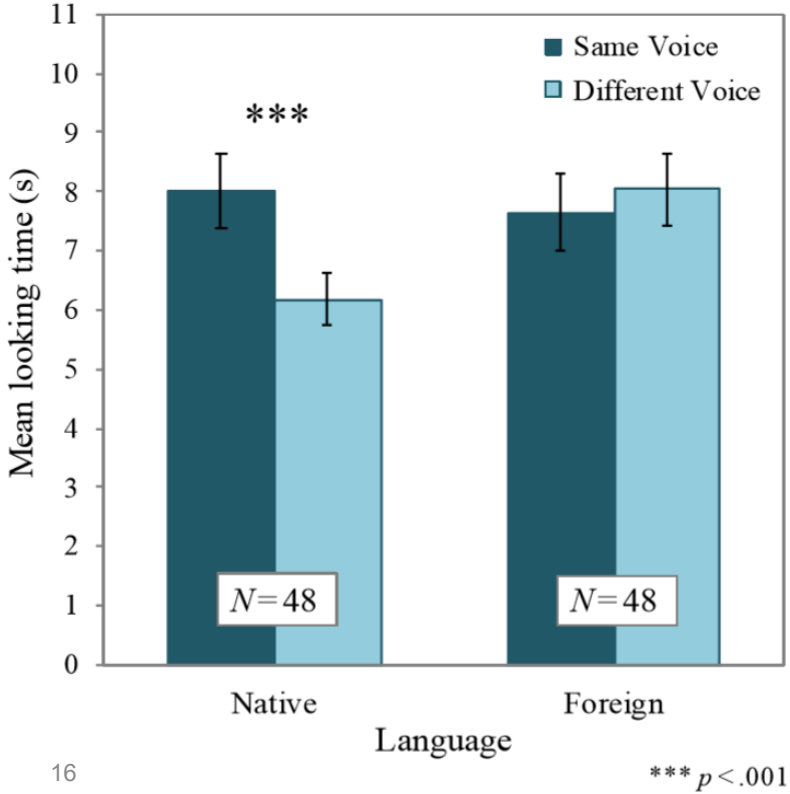
- Infants were able to discriminate between speakers of their native language but not of a non-native language
- We have built a model of how infants could represent speech and can now use this to test for a language familiarity effect

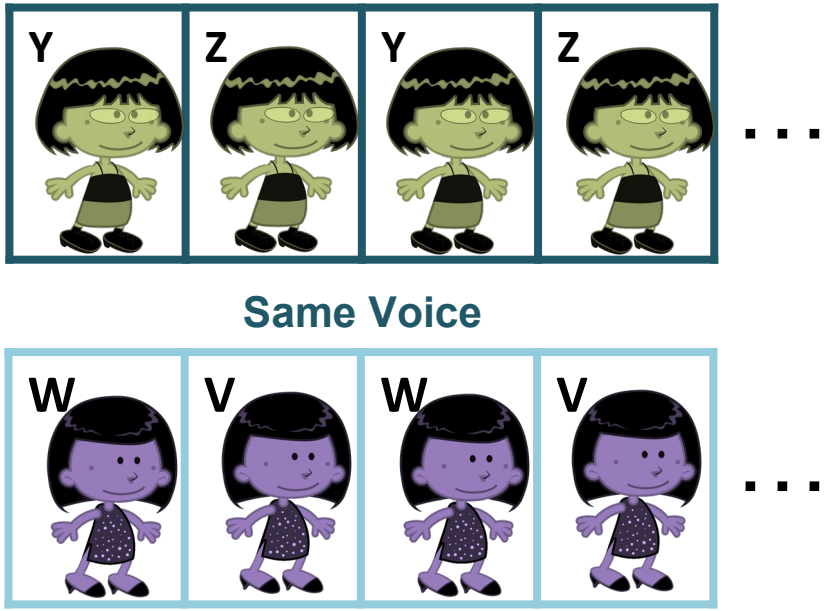
Visual Fixation Procedure



Fecher & Johnson 2019

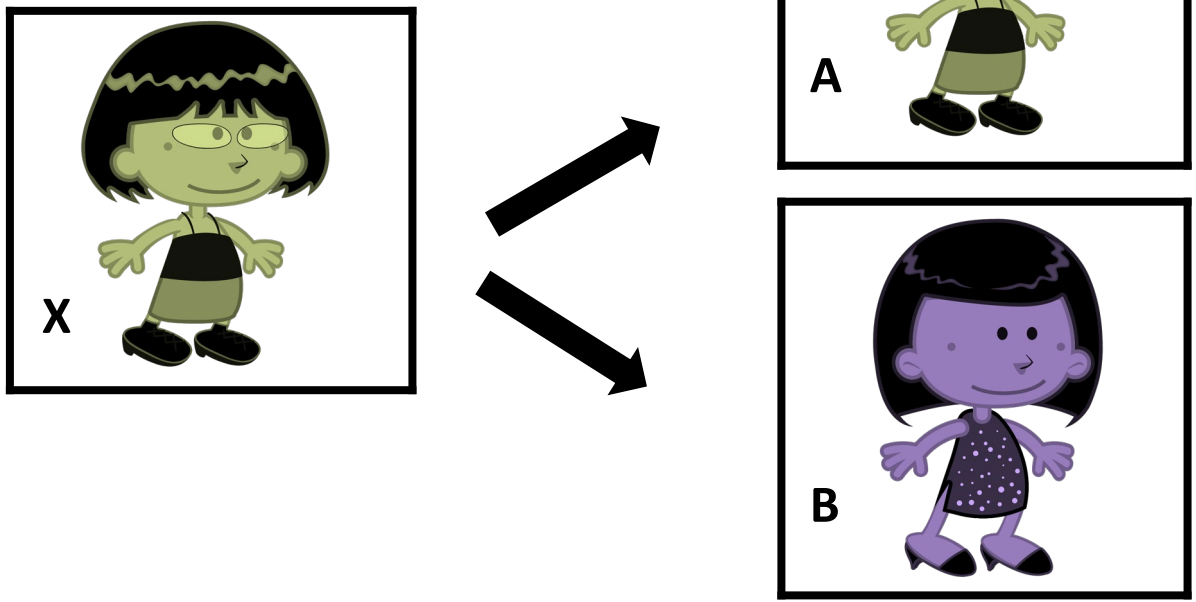
Visual Fixation Procedure





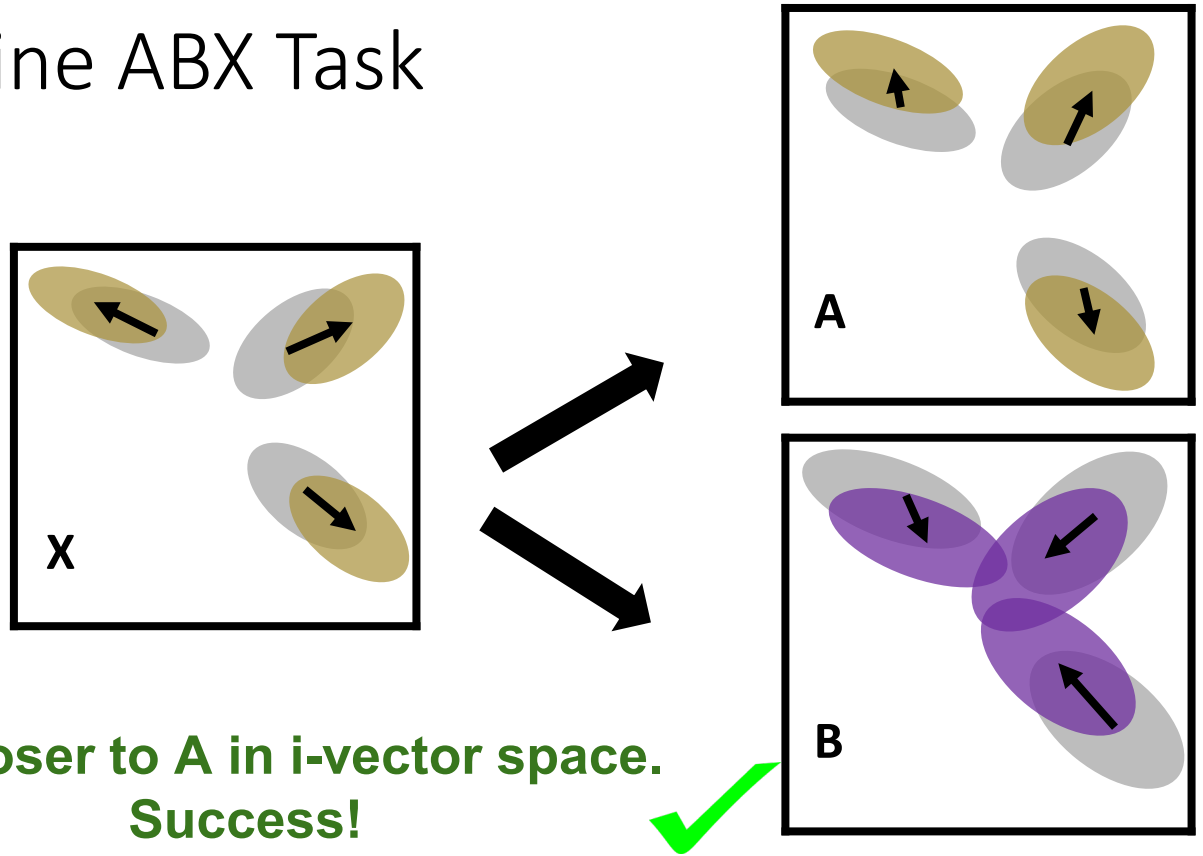
Fecher & Johnson 2019

Machine ABX Task



Schatz et al 2013, Schatz 2016

Machine ABX Task

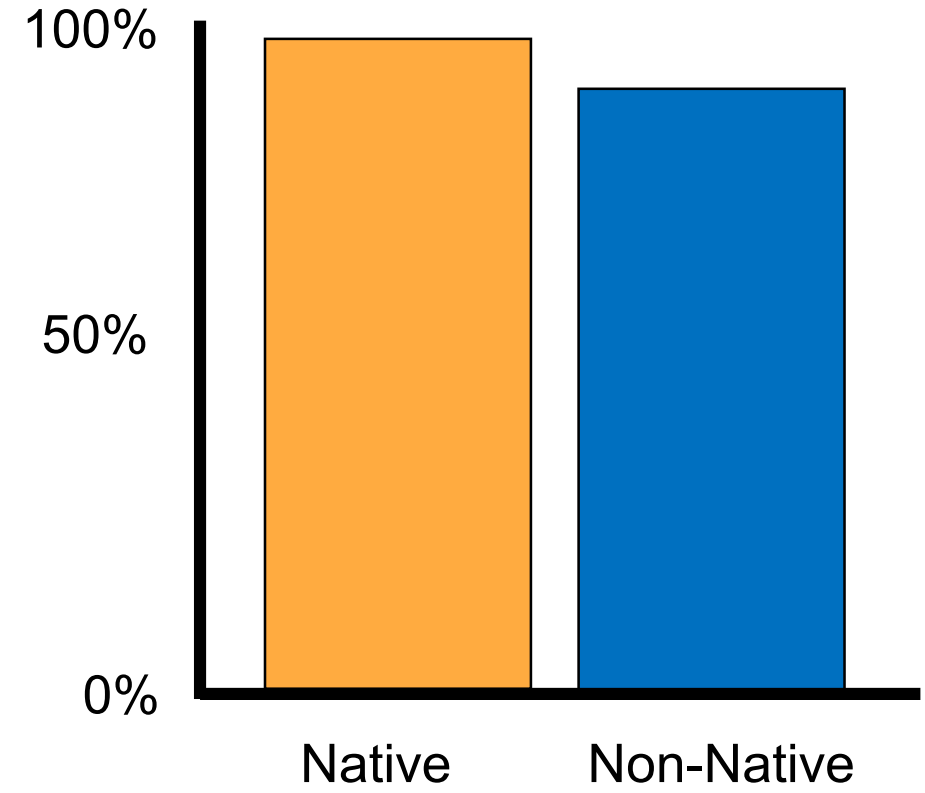


X is closer to A in i-vector space.
Success!

Schatz et al 2013, Schatz 2016

Predicted Results

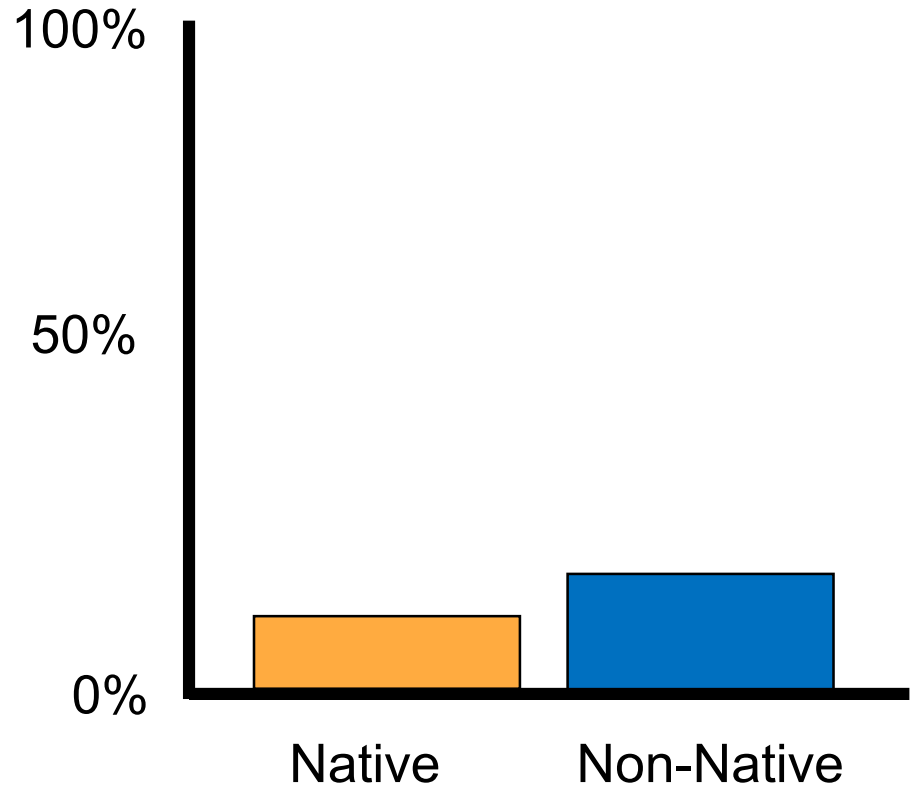
ABX Trials Performed Correctly (%)



Predicted Results

ABX Error Rate (%)

Probability of incorrect judgement



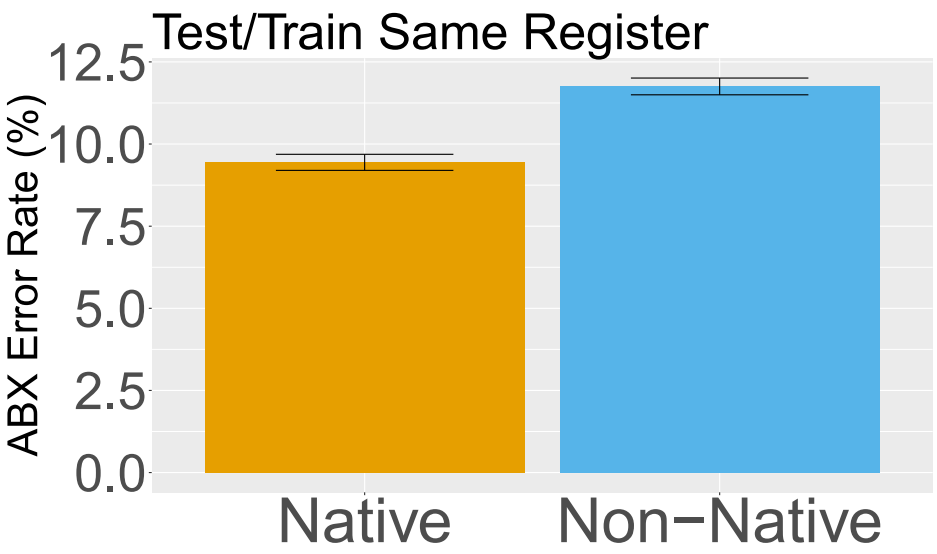
Experimental Paradigm

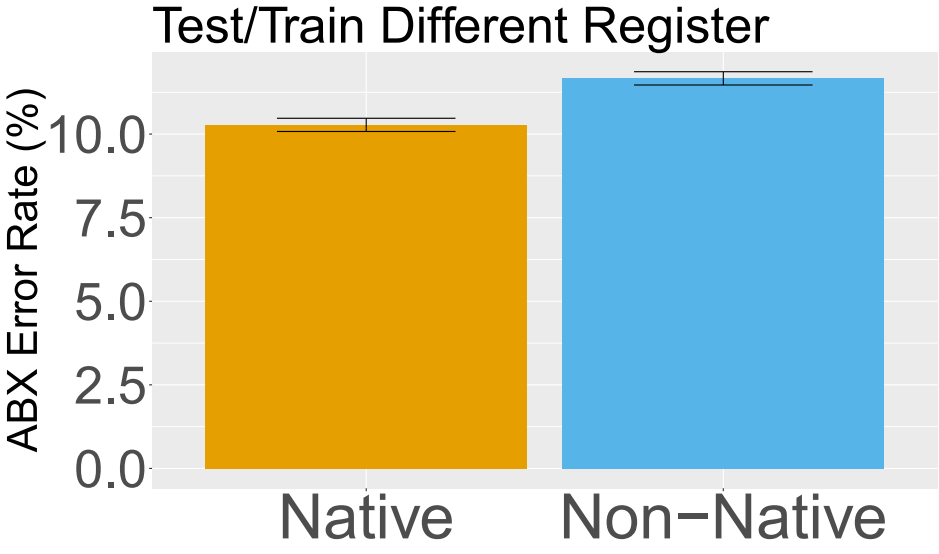
- Train models on four corpora
  - English and Japanese
  - Read and Spontaneous
- Test on utterances from all corpora

Training Language

Test Language	**English**	**Japanese**
English	Native	Non-Native
Japanese	Non-Native	Native

Results


Test/Train Same Register
ABX Error Rate (%)
12.5
10.0
7.5
5.0
2.5
0.0
Native
Non−Native


Test/Train Different Register
ABX Error Rate (%)
10.0
7.5
5.0
2.5
0.0
Native
Non−Native

Conclusion

- Language Familiarity Effect is found in infants at 4.5 months old
- This is the first model to account for this effect
- We propose that the infant uses acoustic variability hierarchically at multiple timescales
- No sophisticated linguistic knowledge required

Next Steps

- Validating against languages and conditions used in infant experiments (ie. Reversed speech)
- Investigating robustness to amount of training data
- Confirming this model empirically in other paradigms
- Connecting this model to the adult mechanism
- Showing other infant native language biases with this model

Acknowledgements

- Thomas Schatz
- Naomi Feldman

- Thanks to Elizabeth Johnson and the Maryland Phonology Circle for helpful discussions
- NSF grant BCS-1734245
- ESRC grant ES/R006660/1

UNIVERSITY OF MARYLAND

LINGUISTICS at MARYLAND

References

- Carbajal, M. J., Fer, R., & Dupoux, E. (2016). Modeling language discrimination in infants using i-vector representations. *In Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 889–894).
- Dehak, N., Kenny, P. J., Dehak, R., Dumouchel, P., & Ouellet, P. (2010). Front-end factor analysis for speaker verification. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(4), 788–798.
- Fecher, N., & Johnson, E. K. (2019). By 4.5 months, linguistic experience already affects infants talker processing abilities. *Child Development*.
- Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science*, 14(5), 1002–1011.
- Schatz, T. (2016). ABX-discriminability measures and applications Doctoral dissertation. Universite Paris 6 (UPMC).
- Schatz, T., Feldman, N., Goldwater, S., Cao, X. N., & Dupoux, E. (2019). Early phonetic learning without phonetic categories – Insights from machine learning (preprint). PsyArXiv.
- Schatz, T., Peddinti, V., Bach, F., Jansen, A., Hermansky, H., & Dupoux, E. (2013). Evaluating speech features with the minimaln pair ABX task: Analysis of the classical MFC/PLP pipeline. *In Proc. INTERSPEECH*