

# Leveraging the uniformity framework to examine crosslinguistic similarity for long-lag stops in spontaneous Cantonese-English bilingual speech

Khia A. Johnson

Department of Linguistics  
University of British Columbia

Interspeech 2021 | Paper ID 1780

**Do highly similar speech sounds from different languages arise from the same underlying category?**

# The SpiCE Corpus

## Speech in Cantonese & English

- Casual interview data
- 34 early bilingual talkers
- Sample of **9,458** long-lag stops

Documentation:  
<https://spice-corpus.readthedocs.io>



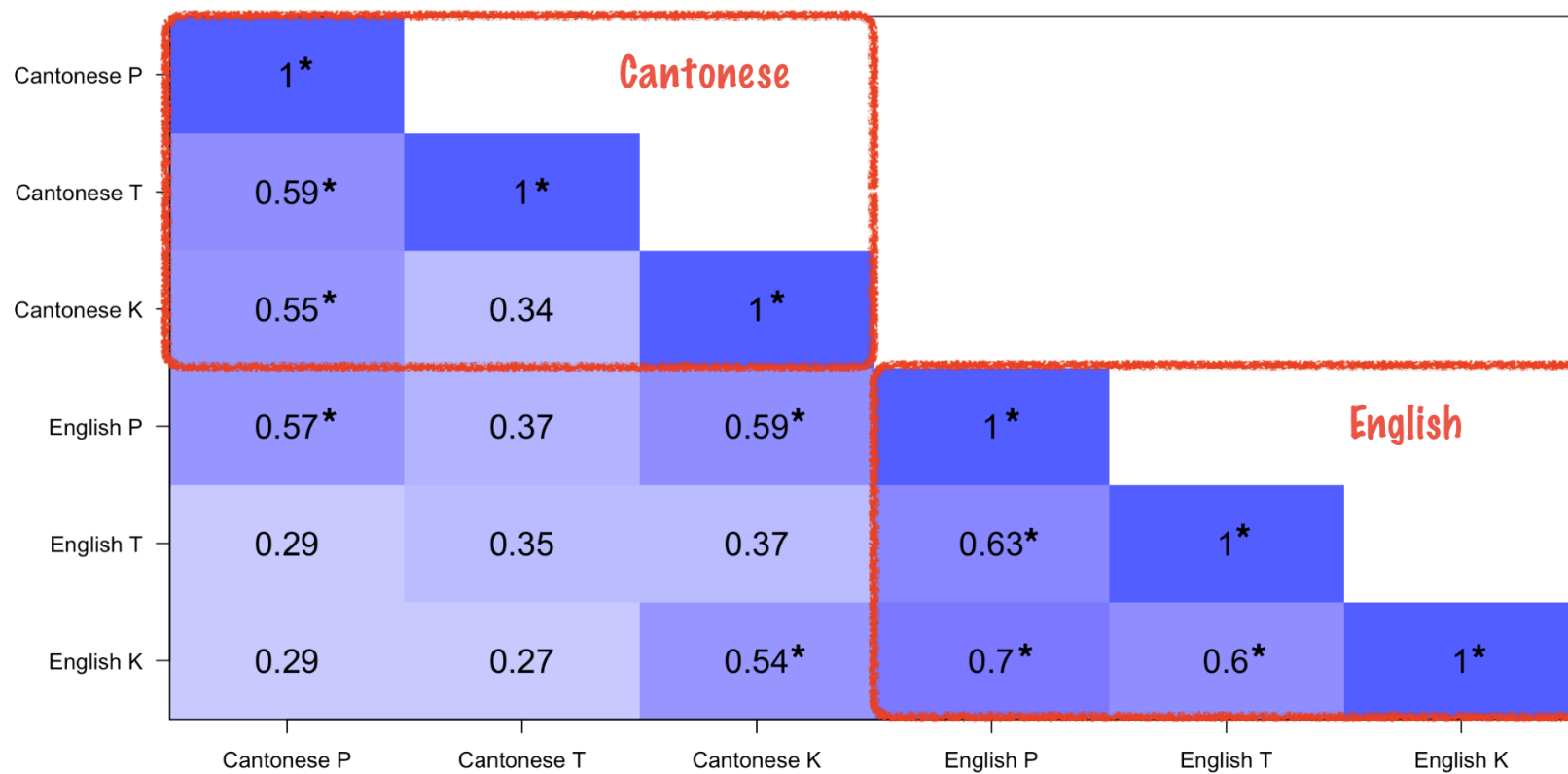
# Identifying *uniform* structure in phonetic variation:

- Ordinal relationships
- Correlations
- Mixed effects model

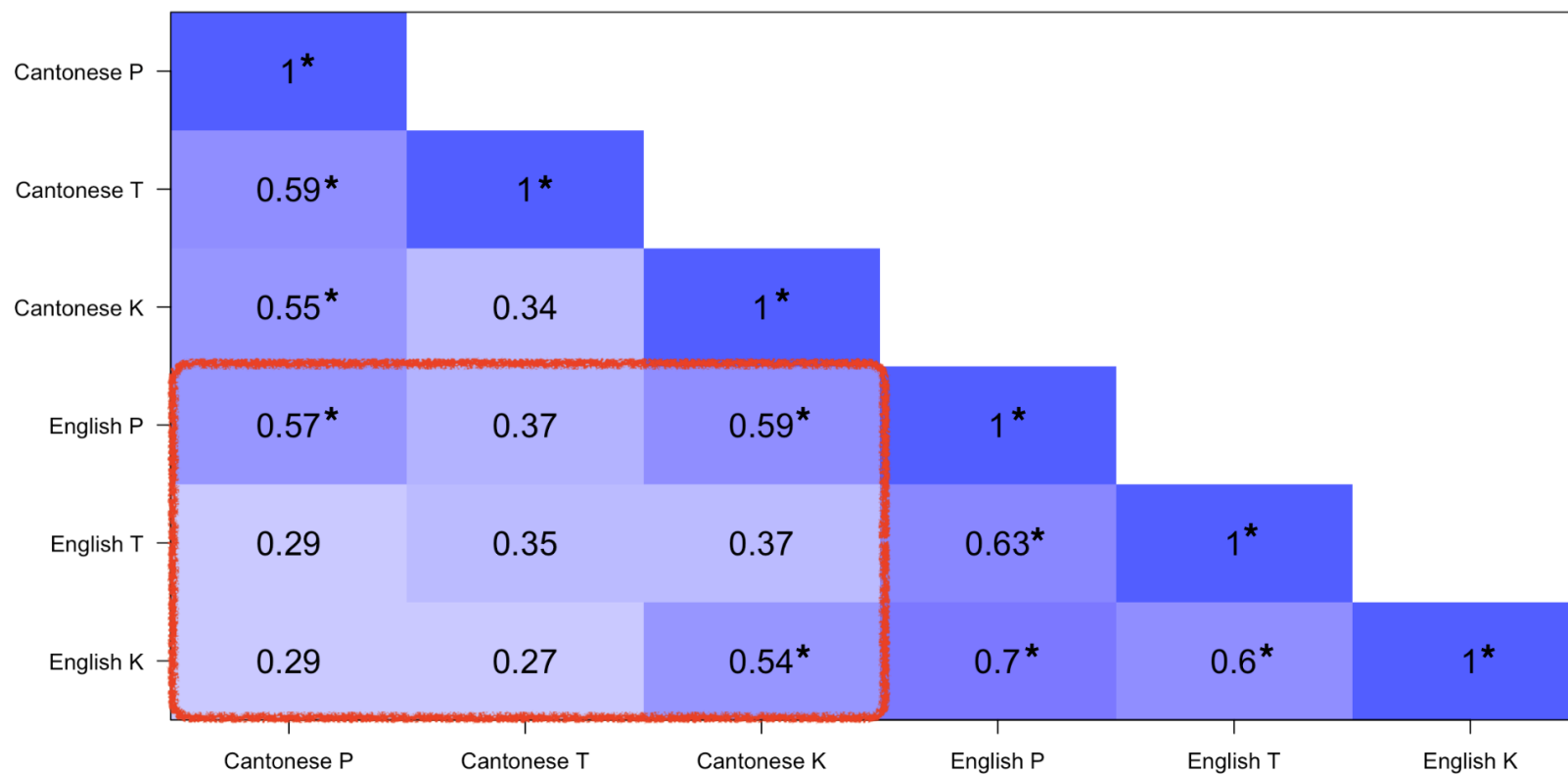
## Relatively few talkers show expected ordinal relationship for mean VOT

Language	$P < T$	$T < K$	$P < K$
Cantonese	0.24 ❌	0.61 🙄	0.49 🙄
English	0.82 ✅	0.06 ❌	0.47 🙄

# Within language correlations show *some* uniformity



# Across language correlations show *very little* uniformity



# The takeaway? It's murky.

- Less structure to VOT variation than anticipated
- Mixed effect model indicates English T/K have consistently longer VOT
- Implications for multilingual representation