Using Speech Technology to test Theories of Phonetic and Phonological Typology

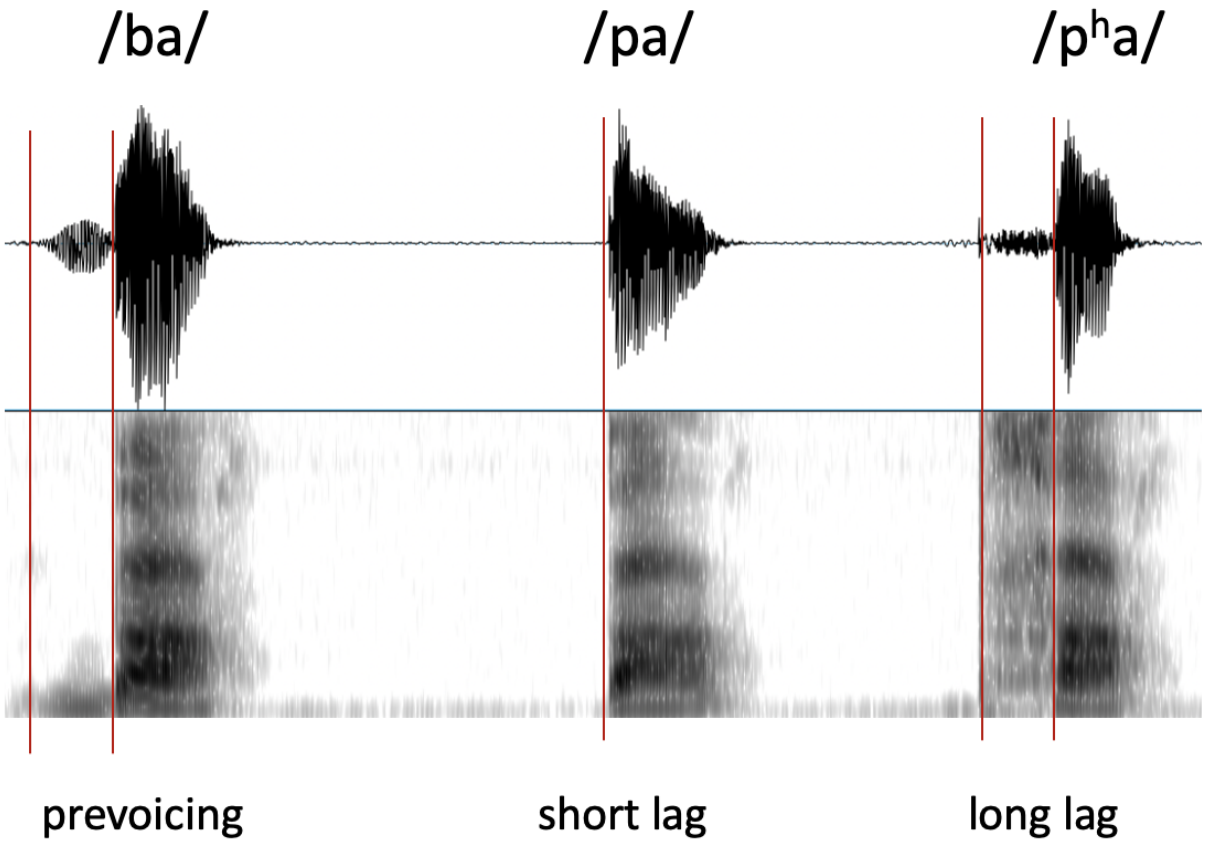
Anisia Popescu, Lori Lamel, Ioana Vasilescu

LREC-COLING - 20-25 May, 2024

Outline

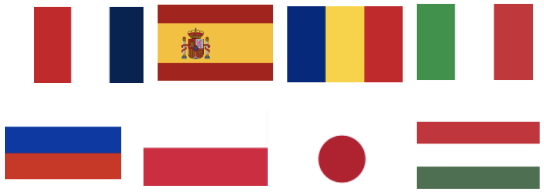
- Introduction : voicing contrast across languages
- Motivation behind the study
- Research questions
- Method
- Experiments
  - Experiment 1
  - Experiment 2
  - Experiment 3
- Discussion

Voicing Contrast

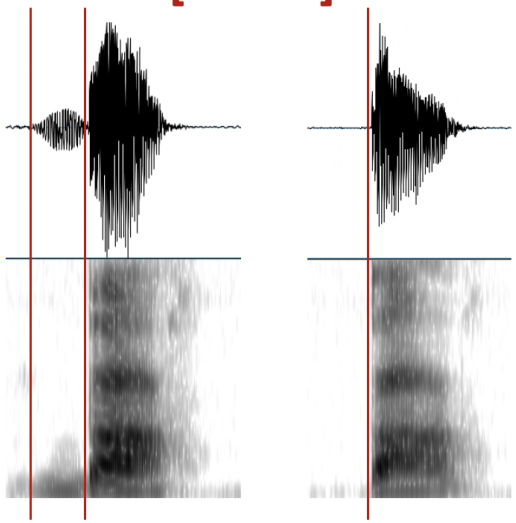


/ba/-/pa/ contrast : Voice Onset Time (VOT)

Voicing Contrast : Implementation mechanisms

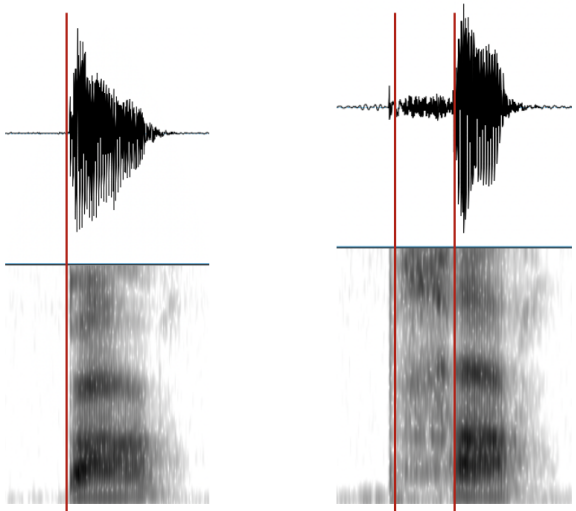


[VOICE]



prevoicing short lag

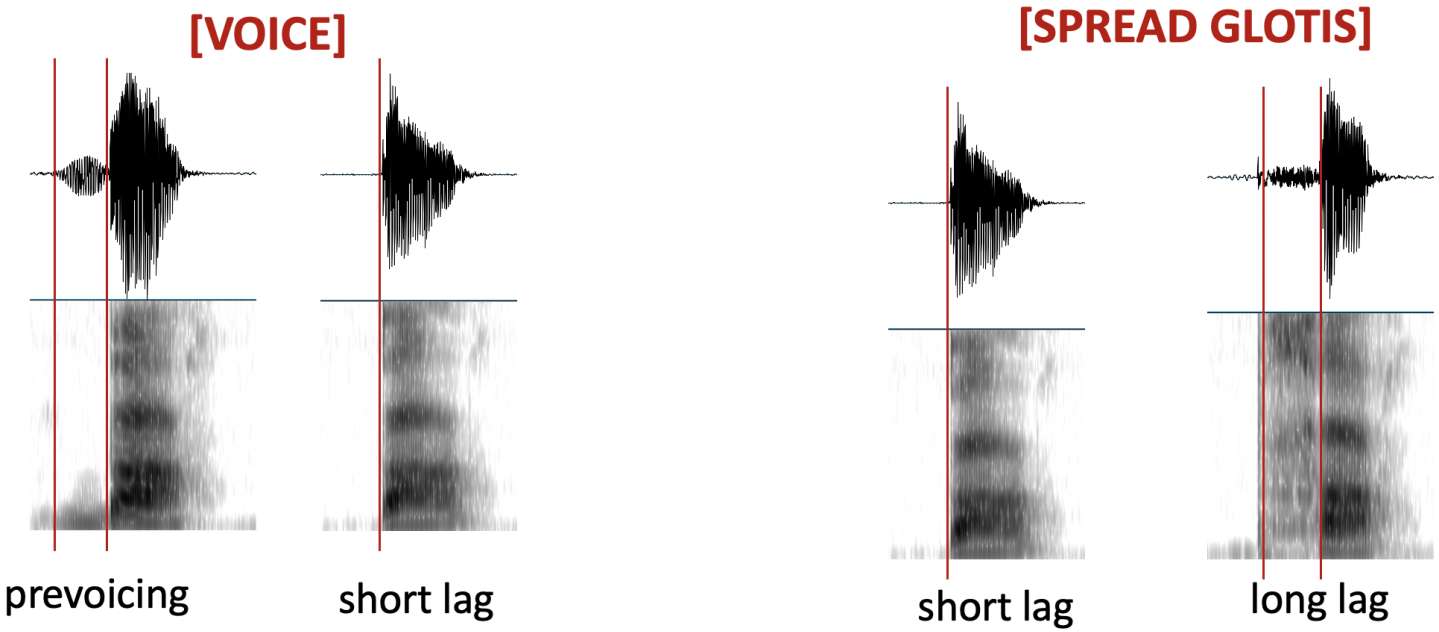
[SPREAD GLOTIS]



short lag long lag

Two main implementation mechanisms

What about European Portuguese



Two main implementation mechanisms

European Portuguese - the odd one out ?

- Traditionally described as a voicing language
- Tendency to not exhibit consistent robust voicing
  - High rates of devoicing - /b,d,g, v, z, ʒ/ realized as /p,t,k,f,s,ʃ/ (Lousada, Jesus & Hall, 2010 ; Jesus & Shadle, 2002 ; Hutin etal., 2021 ; Popescu etal., 2023)
  - EP voicing profiles are more similar to Germanic languages (Pape & Jesus, 2011, 2015 ; Shih & Möbius, 1999)
  - Decision tree-based classifications separate Portuguese from Spanish, French, Italian and Romanian when it comes to devoicing patterns (Wu etal., 2022)
- An alternative *hybrid voicing system* has been proposed for EP : [voice] for stops and [spread glottis] for fricatives (Ramsammy & Strycharczuk, 2016)

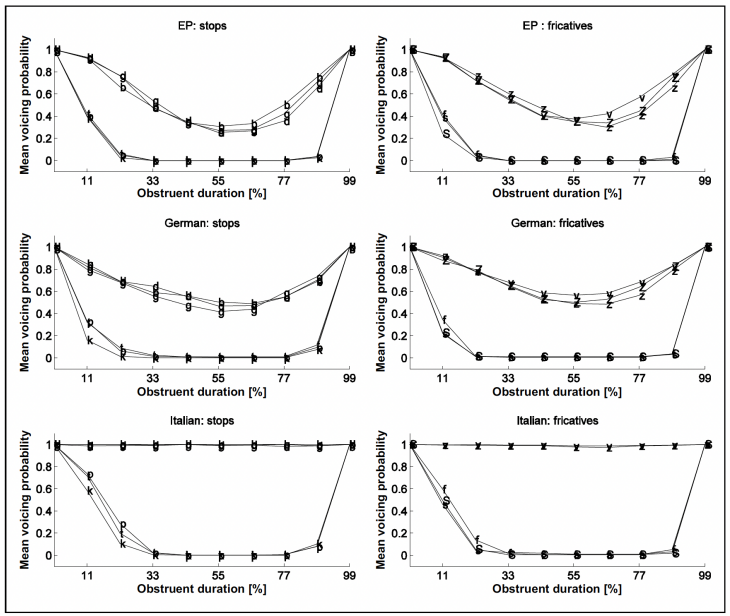
Research question

Is European Portuguese a true voicing language like its genetically related Romance languages, or did it shift towards an "aspirating" language voicing system ?

Background

Pape & Jesus, 2015

- Voicing profiles of EP obstruents are more similar to those of German than those of Italian
- Voicing probability drops after 30% of the obstruent in EP and German



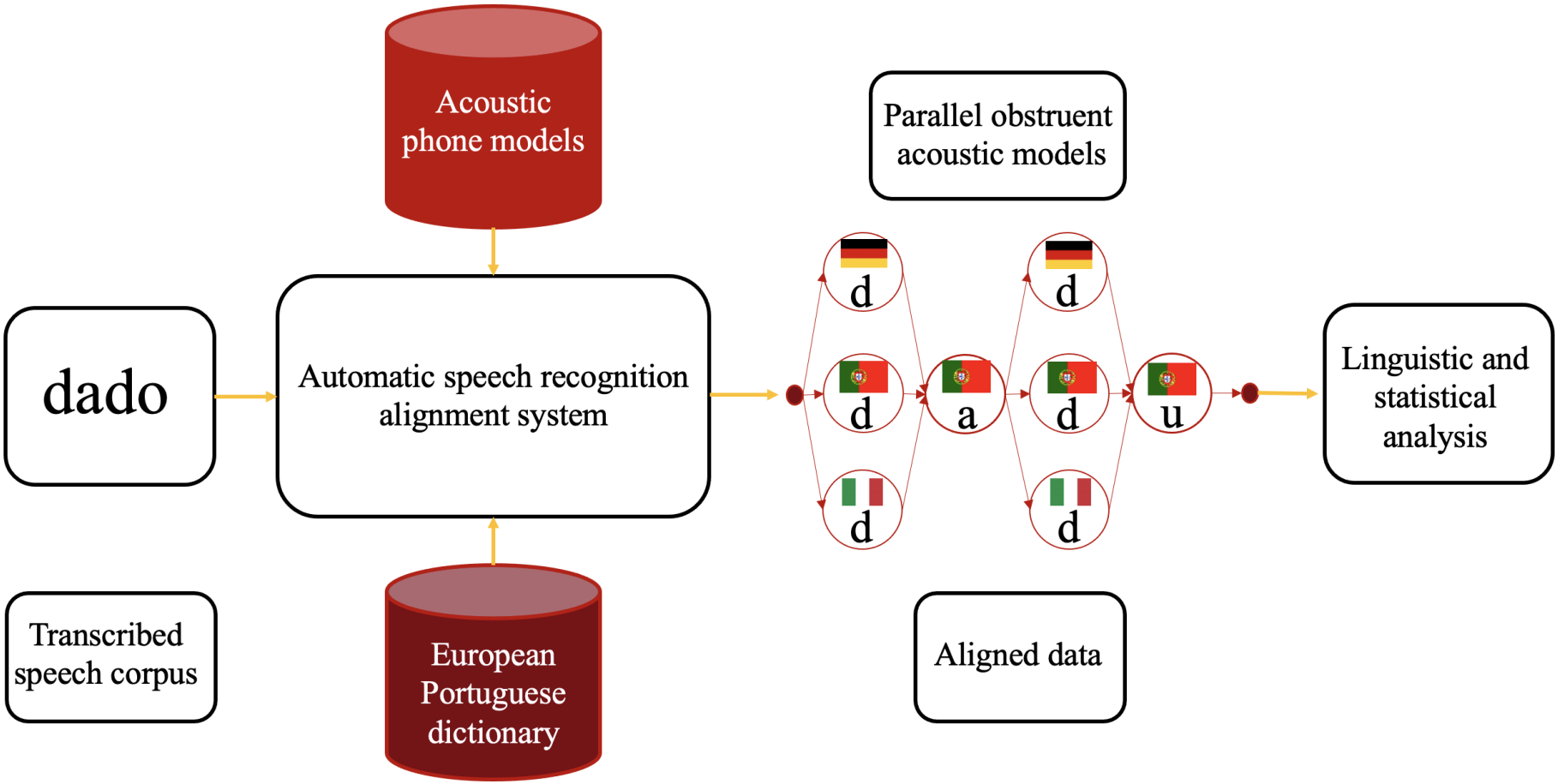
Source : Pape & Jesus, 2015 (p.232)

Experimental Research question

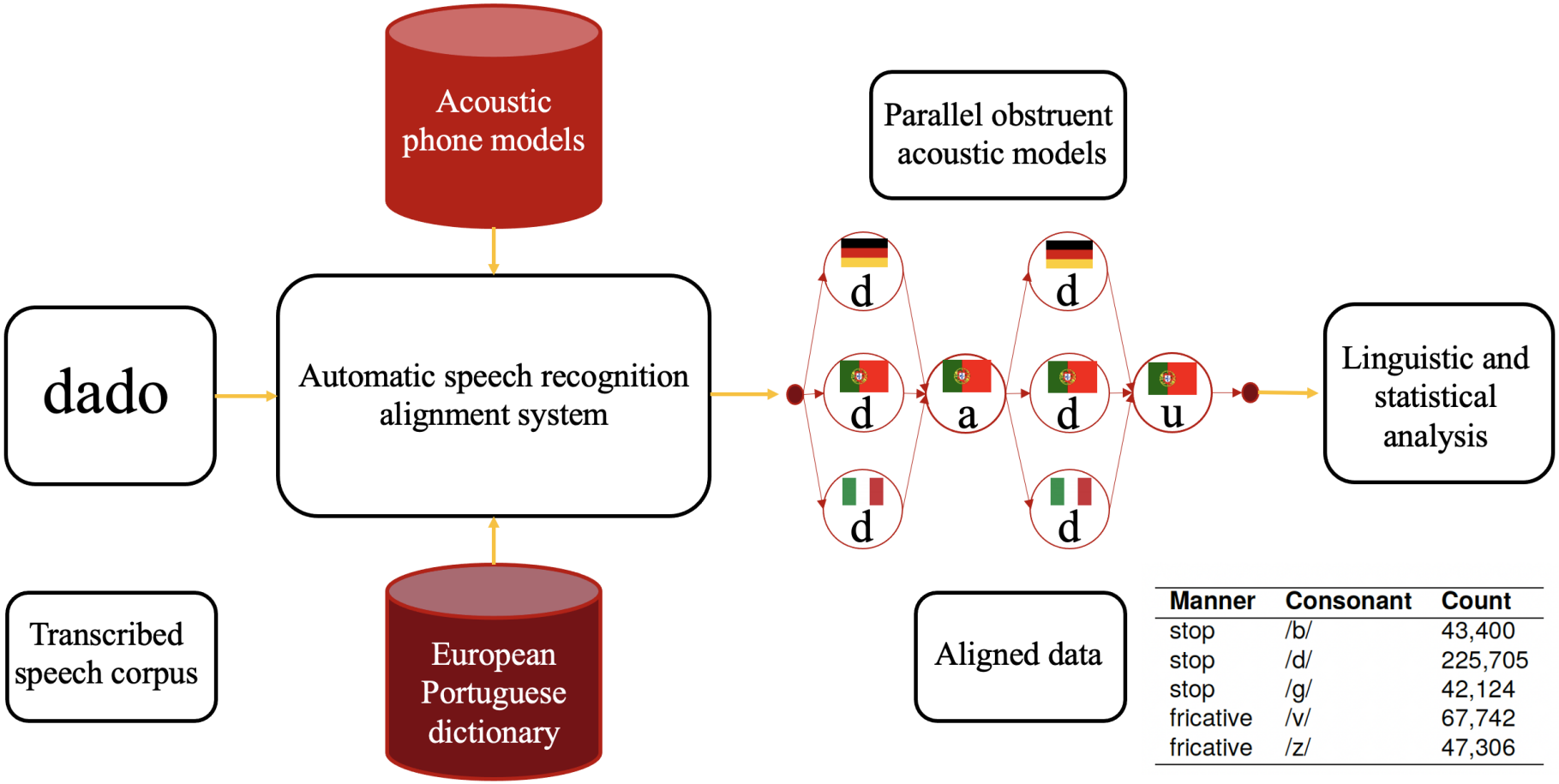
Are voicing profiles in European Portuguese obstruents more similar to German than to Italian ?

Is voicing in fricatives and stops implemented the same in European Portuguese ?

Method


Acoustic phone models
Parallel obstruent acoustic models
dado
Automatic speech recognition alignment system
d
d
d
a
d
d
u
d
Linguistic and statistical analysis
Transcribed speech corpus
European Portuguese dictionary
Aligned data

Method



Manner	Consonant	Count
stop	/b/	43,400
stop	/d/	225,705
stop	/g/	42,124
fricative	/v/	67,742
fricative	/z/	47,306

Acoustic models

- trained on similar types of data :
  - $\approx$ 100h of transcribed broadcast news data
- using same acoustic analysis
  - ceptstral : Perceptual Linear Prediction (PLP)
  - pitch : F0
- using same type of models
  - context-, speaker-, word-position independent monophone models
  - 3-state left-to-right continuous density HMM with Gaussian mixtures with up to 32 Gaussians per state

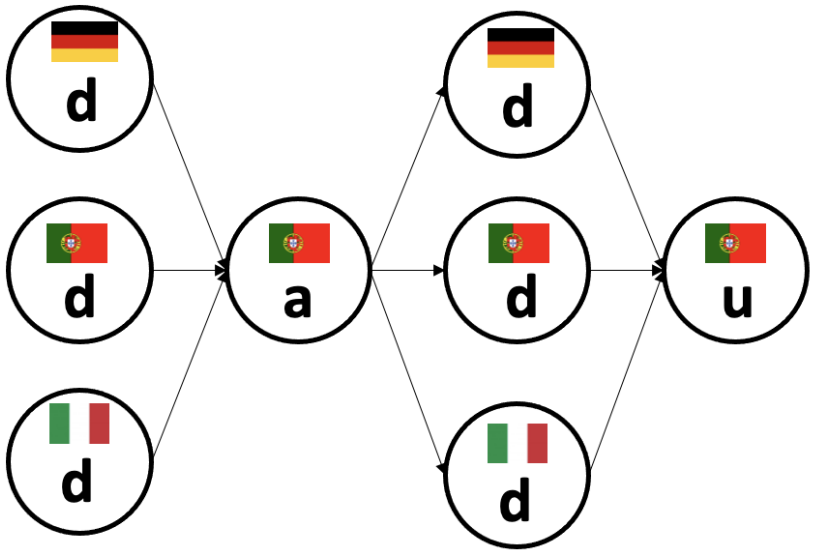
	Portuguese	**Italian**	**German**
Tokens	1.1 million	1.8 million	1.8 million
Word types	46k	58.8k	90k

Predictions

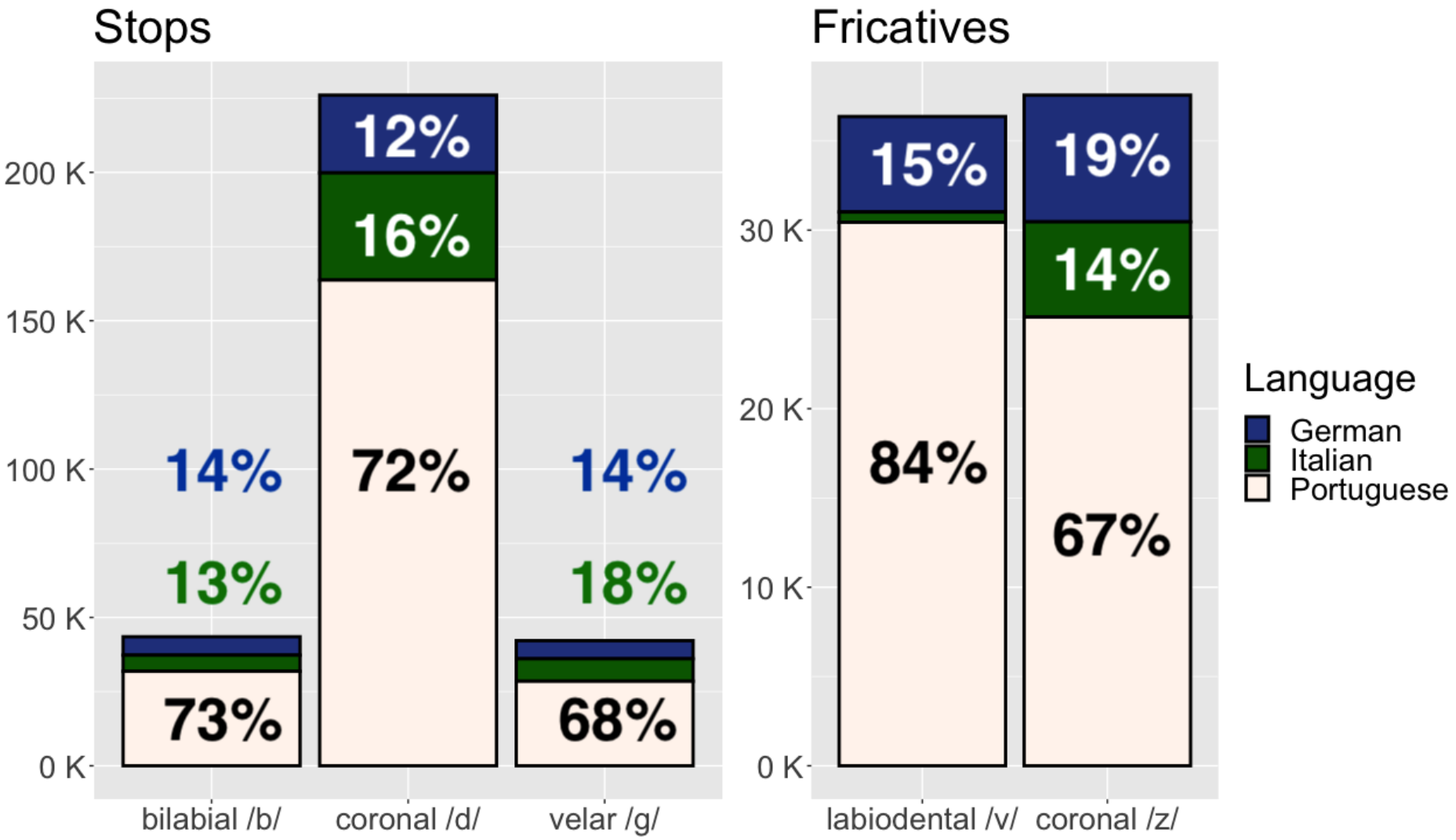
If voicing in EP obstruents is more similar to German we would expect the recognition system to choose the German model to a higher degree than the Italian one

If voicing in EP behaves similar to Italian, we would expect the Italian acoustic model to be preferred

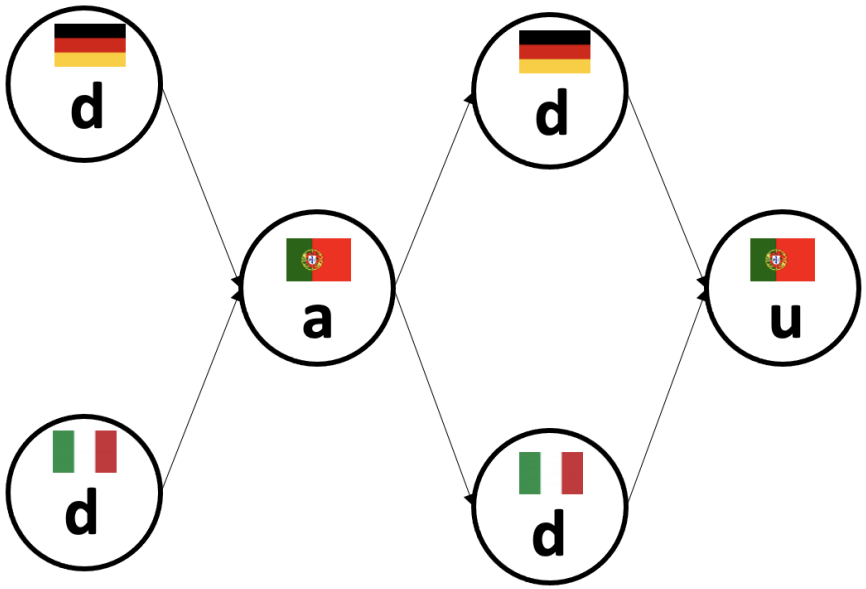
Experiment 1 : Three-way choice of acoustic models



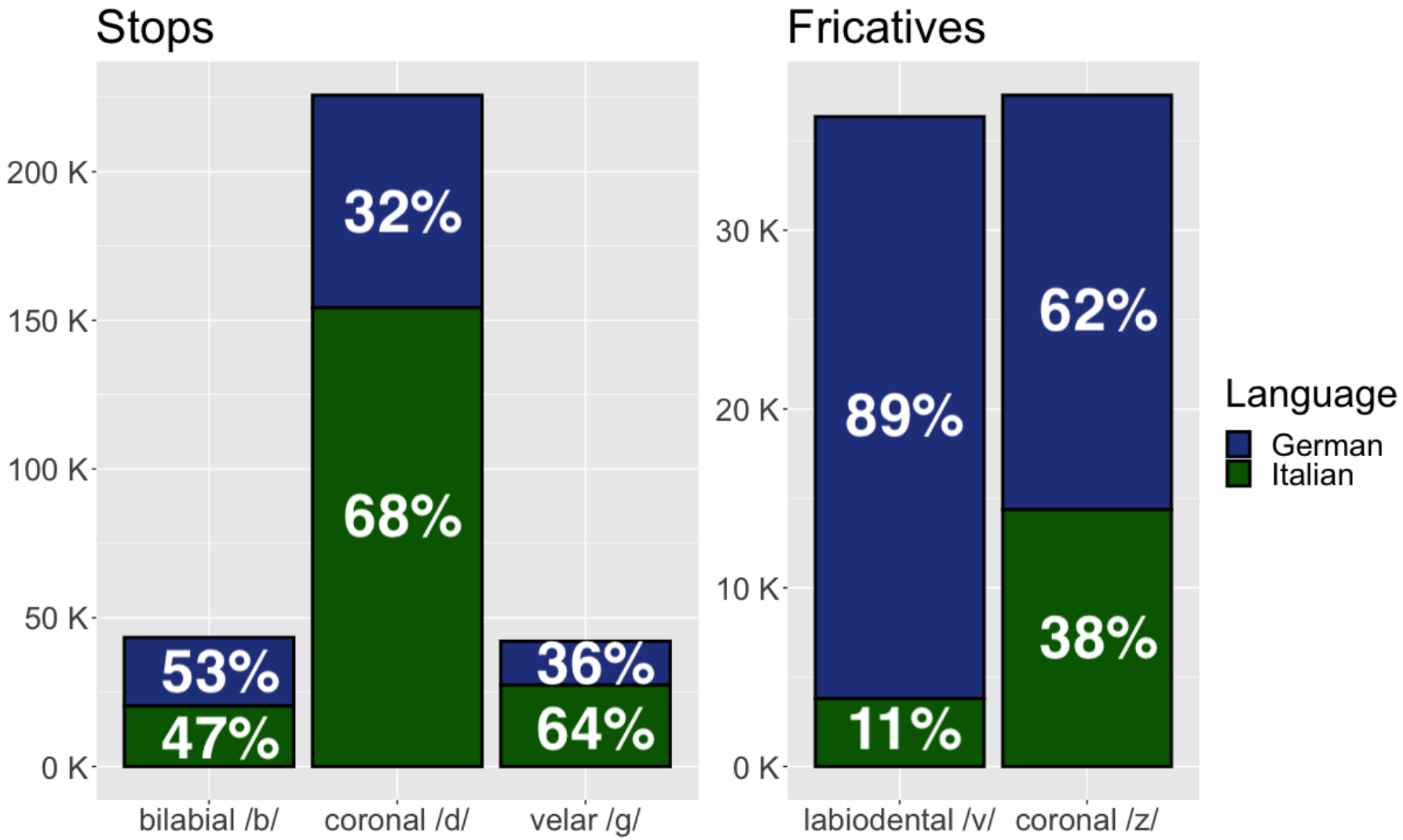
Experiment 1 : Three-way choice of acoustic models



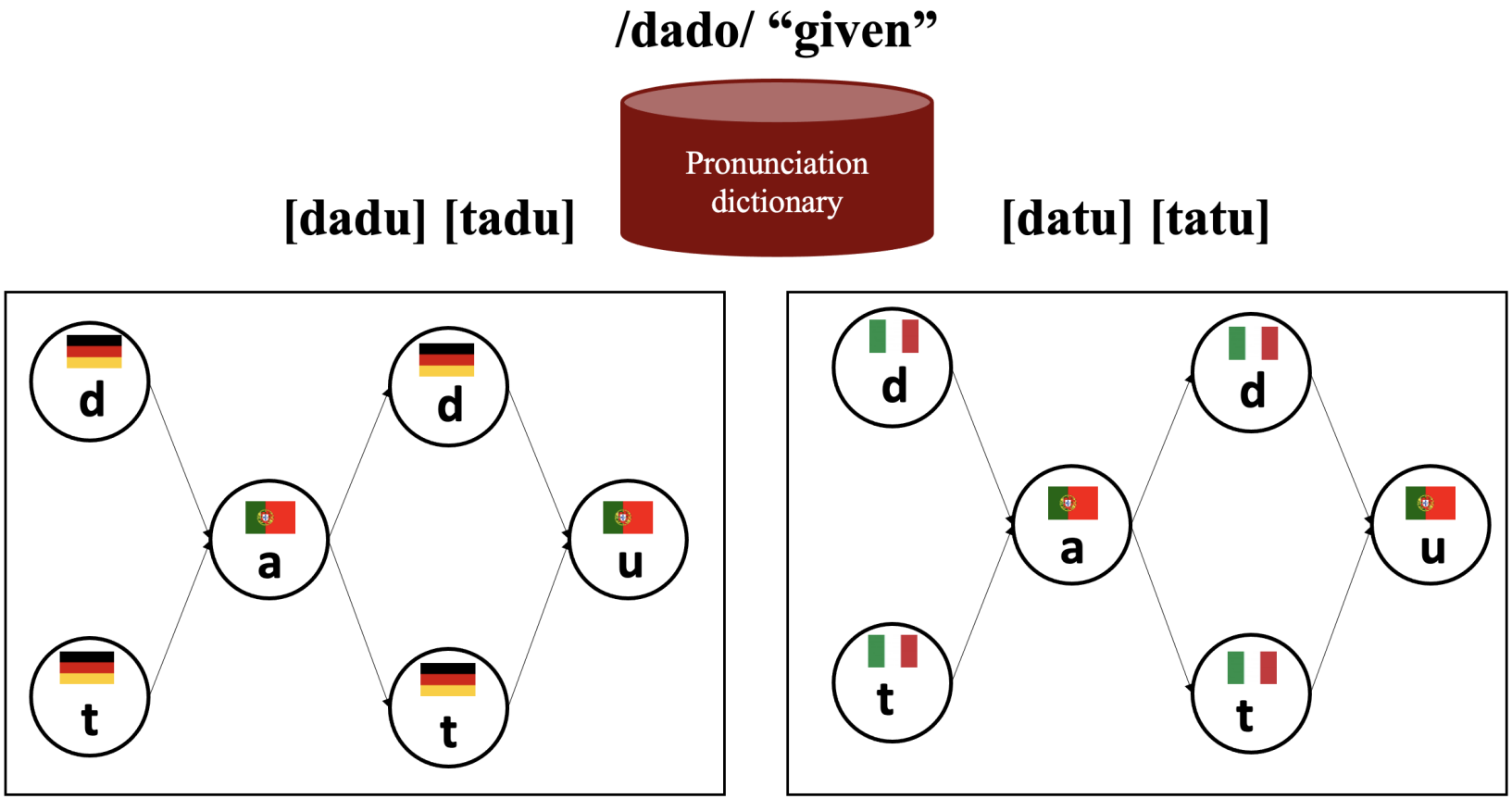
Experiment 2 : Two-way choice of acoustic models



Experiment 2 : Two-way choice of acoustic models

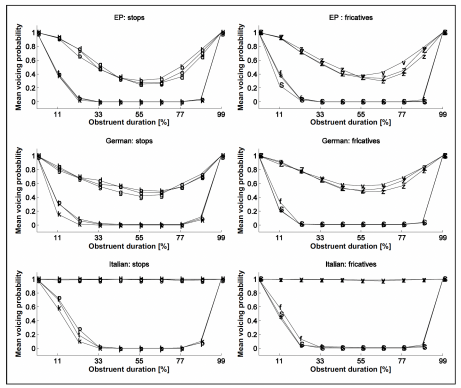


Experiment 3 : Single acoustic models with pronunciation variants



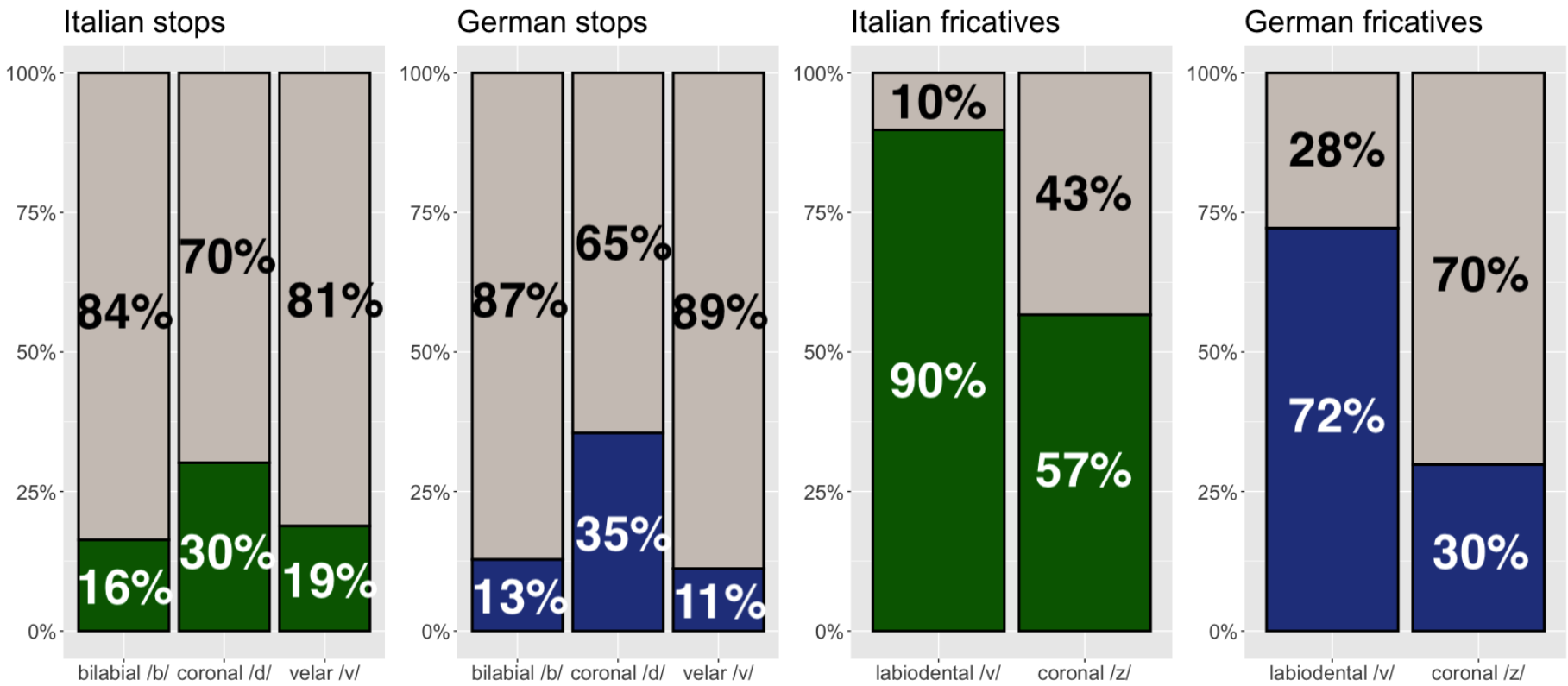
Experiment 3 :Single acoustic models with pronunciation variants

- The probability of voicing throughout the obstruent stays close to 1 for Italian, and drops after 30% for both EP and German
- If EP voicing profiles are more similar to German than Italian
  - Higher percentages of voiceless variants when using the Italian acoustic model



Source : Pape & Jesus, 2015

Experiment 3 : Single acoustic models with pronunciation variants



Conclusion

- Experiments 1 & 2 :
  - EP voicing profiles of fricatives (/v,z/) are closer to German than Italian
  - EP voicing profiles of stops (/b,d,g/) are closer to Italian than German
- Experiment 3 :
  - using a different method relying on pronunciation variants instead of letting the system chose its preferred acoustic model shows the same pattern

Discussion

- Results show that the voicing patterns in EP are shifting towards those seen for "aspirating" languages
- The difference between fricatives and stops points towards a hybrid system (see Ramsammy & Strycharczuk, 2016)
- The difference between consonant type among stop consonants suggest there is an ongoing change withing this subset of obstruents

THANK YOU FOR YOUR ATTENTION