Do rhythm measures separate languages or speakers?

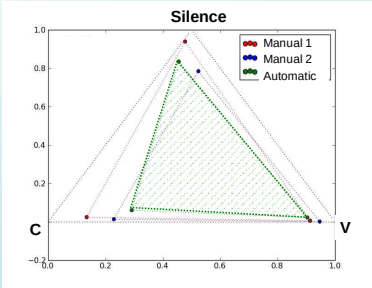
Anastassia Loukina[1], Greg Kochanski[1], Elinor Keane[1], Burton Rosner[1], Chilin Shih[2]
[1]Oxford University Phonetics Laboratory, UK
[2]EALC/Linguistics, University of Illinois, Urbana-Champaign, USA

Data

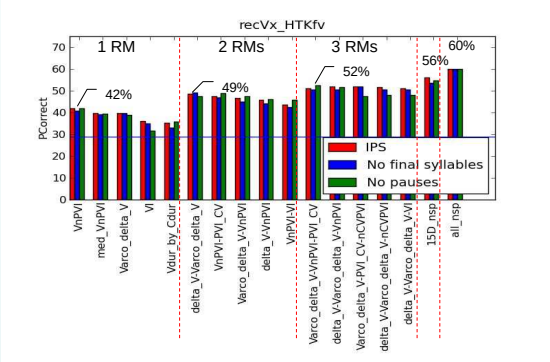
- **5 languages**: British English (Southern England), Modern Greek (Athens), Russian (Moscow/St. Petersburg), French (Paris), Taiwanese Mandarin (Taipei)
- **45 short texts** in each language (extracts from "Harry Potter", Aesop's fables, Cinderella)
- **50 speakers**: 20-28 years old. All non-English speakers lived outside their home country <4 years.
- **Total size of the corpus: 2265 texts**

Automatic segmentation

- Language independent
- A specialized speech recognition system, based on the HTK toolkit using acoustic description vector
- Produces sequence of C, V and S
- In the training data, all vowels and sonorants were mapped into V and all other phonemes into C

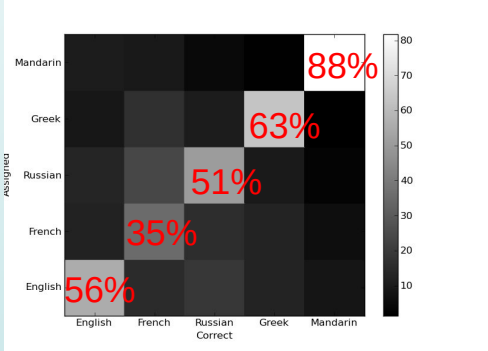


Separating languages



- **45** classifiers based on single RM
- **315** classifiers based on pairs of RM
- **1365** classifiers based on three measures
- **3** classifiers based on 15 RM
- **1** classifier based on 45 RM

- 45D classifier
- Chance=29%
- %Correct=60%
- K*=0.44



*K=(%correct-chance)/(perfection-chance)

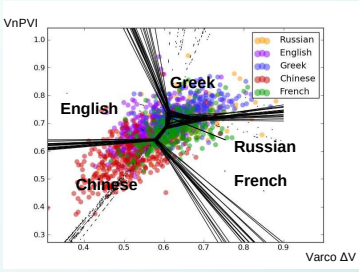
Conclusions

- Rhythm is a multidimensional phenomenon
- Substantial variation within languages makes it impossible to reliably separate languages based on the rhythm of a single paragraph
- The differences between speakers of the same language can be more consistent than the differences between different languages

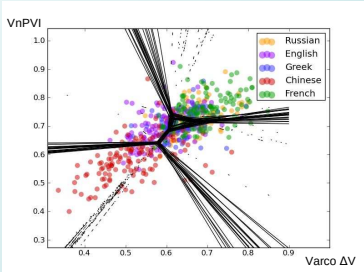
Classifier

- 12 splits into training set/test set
- 20 classifiers for each split
- **Data from one speaker always in the same set**

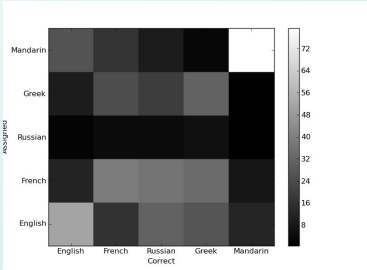
All classifiers for one training set



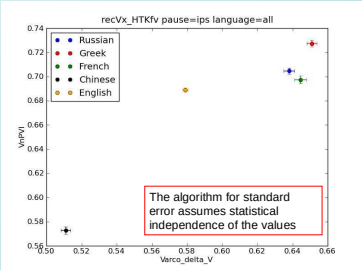
All classifiers for one test set



Confusion matrix: 47% correct



Mean values and standard errors

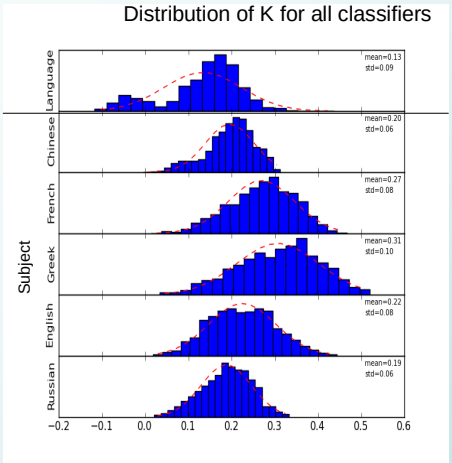


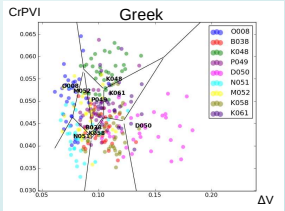
Varco ΔV- VnPVI

Separating people

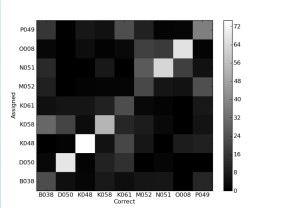
How well can a linear classifier identify the speaker of an utterance?

Distribution of K for all classifiers





Chance=12%, Correct=47%, K=0.40



Maximum % correct achieved for each language*

Language	Russian	English	Greek	French	Chinese
Chance	10%	9%	12%	11%	10%
Max %	40%	49%	58%	53%	38%

*For 1D, 2D and 3D classifiers

How do ΔV vs. CrPVI separate speakers in other languages?

French: 38% (11%), K=0.30
English: 31% (9%), K=0.24
Chinese: 30% (10%), K=0.22
Russian: 27% (10%), K=0.19

And how do they separate languages?

27%= chance, K=0

Rhythm measures

Global	Local	CV based
%V	CrPVI	PVI-CV
ΔV	VnPVI	VI
ΔC	CnPVI	YARD
Vdur/Cdur	med CrPVI	nCVPVI
VarcoΔV	med VnPVI	
VarcoΔC		

References

W. Barry, B. Andreeva, M. Russo, S. Dimitrova, and T. Kostadinova, "Do rhythm measures tell us anything about language type?," in *Proceedings of the 15th ICPhS*, M. J. Solé and J. Romero, Eds. Barcelona, 2003, pp. 2693-2696.
F. Ramus, M. Nespor, and J. Mehler, "Correlates of linguistic rhythm in the speech signal," *Cognition*, vol. 73, p. 28, 1999.
E. Grabe and E. L. Low, "Durational Variability in Speech and the Rhythm Class Hypothesis," in *Laboratory Phonology, 7*, C. Gussenhoven and N. Warner, Eds. : Mouton de Gruyter, Berlin, Germany, 2002, pp. 515-46.
E. Ferragne and F. Pellegrino, "A comparative account of the suprasegmental and rhythmic features of British English dialects," in *Modelisations pour l'Identification des Langues*. Paris, 2004.
P. Wagner and V. Dellwo, "Introducing YARD (yet another rhythm determination) and re-introducing isochrony to rhythm research," in *Speech Prosody 2004*, Nara, Japan, 2004, pp. 227-230.
E. L. Asu and F. Nolan, "Estonian rhythm and the pairwise variability index," in *FONETIK 2005*, Göteburg University, 2005, pp. 29-32.
V. Dellwo, "Rhythm and speech rate: a variation coefficient for deltaC. ," in *Language and language-processing: Proceedings of the 38th Linguistics Colloquium. Piliscsaba 2003.* , P. Karnowski and I. Szigeti, Eds. Frankfurt am Main: Peter Lang Publishing Group, 2006, pp. 231-241.

Acknowledgements

This project is supported by the Economic and Social Research Council (UK) via RES-062-23-1323.
The support for Dr Shih was provided by NSF IIS-0623805 and NSF IIS-0534133