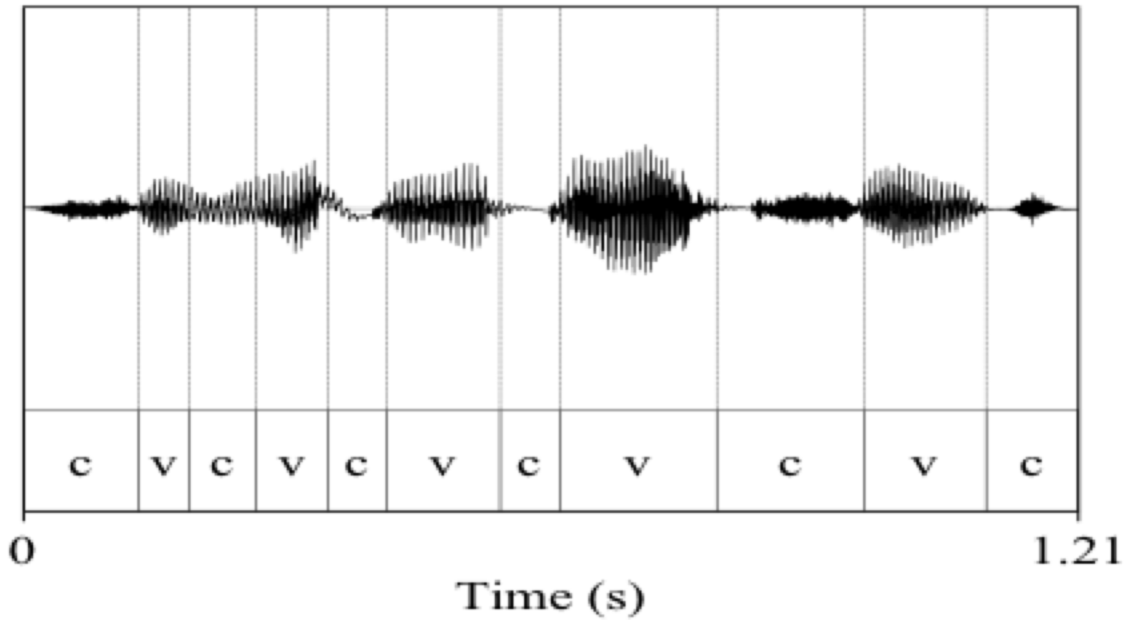
Rhythmic variability in Southeast Asian languages

Volker Dellwo & Peggy Mok

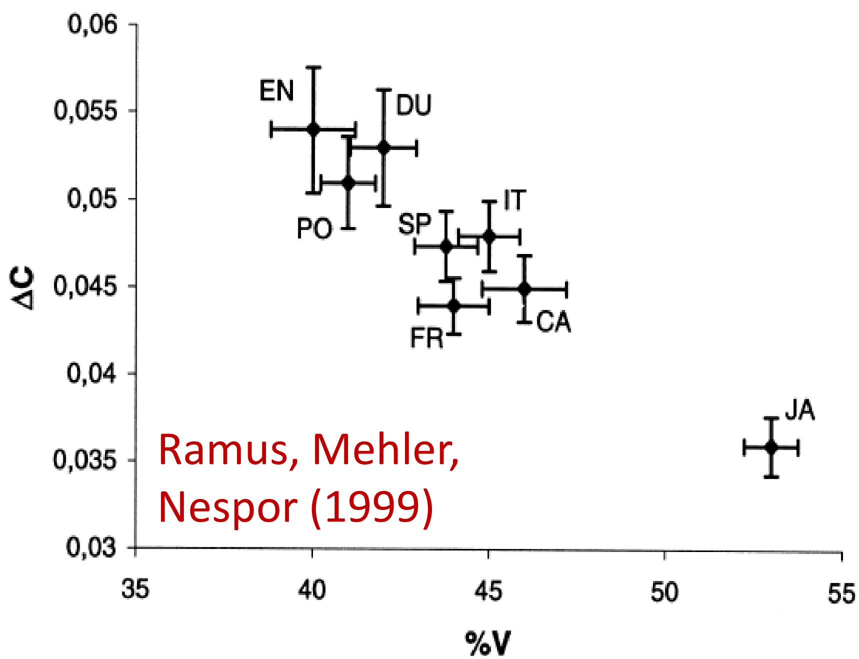
Correlates of speech rythm

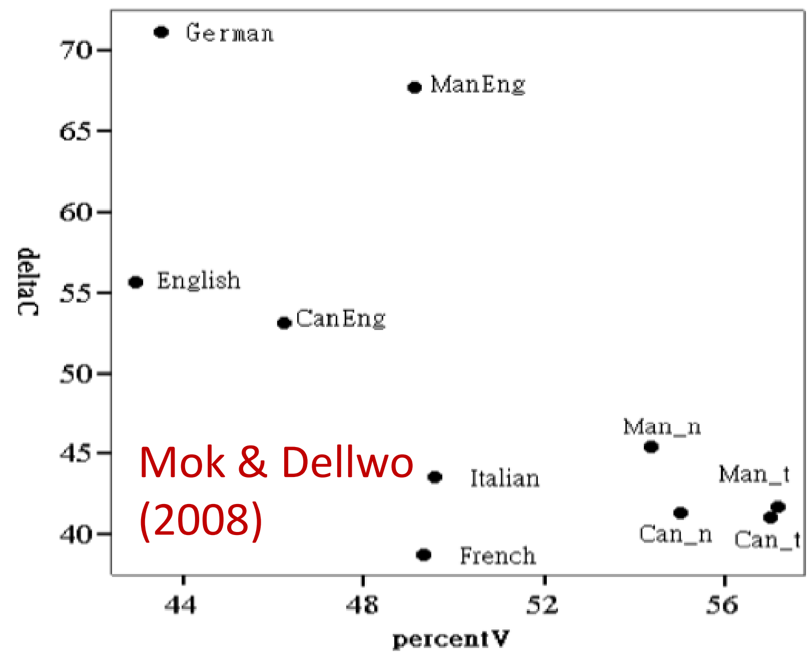


Ramus, Mehler, Nespor (1999-2005)
Nolan, Grabe & Low (2000-2010)
Barry et al. (2003-2007)
Mairano (2007-2012)
Dellwo et al. (2003-12)

e.g:
- (1) %V
- (2) deltaC
- (3) PVI
- ....
- (70) etc.

Correlates of speech rythm





Perceptual evidence: sasasa-speech

(Compare work by Arvaniti [2010-2012] for a summary of criticism)

Journal of Experimental Psychology:
Human Perception and Performance
1998, Vol. 24, No. 3, 756–766

Language Discrimination by Newborns: Toward an Understanding of the Role of Rhythm

Thierry Nazzi, Josiane Bertoncini, and Jacques Mehler
Centre National de la Recherche Scientifique, Ecole des Hautes Études en Sciences Sociales

Three experiments investigated the ability of French newborns to discriminate between sets of sentences in different foreign languages. The sentences were low-pass filtered to reduce segmental information while sparing prosodic information. Infants discriminated between stress-timed English and mora-timed Japanese (Experiment 1) but failed to discriminate between stress-timed English and stress-timed Dutch (Experiment 2). In Experiment 3, infants heard different combinations of sentences from English, Dutch, Spanish, and Italian. Discrimination was observed only when English and Dutch sentences were contrasted wi... ...ish and Italian sentences. These results suggest that newborns use prosodic and, ...ally, rhythmic information to classify utterances into broad language classes d... global rhythmic properties. Implications of this for the acquisitio... ...s of the native language are discussed.

Language Discrimination by Human Newborns and by Cotton-Top Tamarin Monkeys

Franck Ramus,[1*†] Marc D. Hauser,[2] Cory Miller,[2] Dylan Morris,[2] Jacques Mehler[1]

Humans, but no other animal, make meaningful use of spoken language. What is unclear, however, is whether this capacity depends on a unique constellation of perceptual and neurobiological mechanisms or whether a subset of such mecha-... ...nisms is shared with other organisms. To explore this problem, parallel experiments ...and cotton-top tamarin monkeys to assess ...r languages. A habituation-dishabitu... ...an newborns and tamarins can dis... ...t not if the sentences are pla... ...are not present in bac... ...ning to certain pr... ...uditory system...

Effects of Backward Speech and Speaker Variability in Language Discrimination by Rats

Toro, Juan M.; Trobalon, Josep B.; Sebastián-Gallés, Núria (2005)

Journal of Experimental Psychology: Animal Behavior Processes **31** (1): 95–100.

COGNITION

Cognition 73 (1999) 265–292

www.elsevier.com/locate/cognit

Correlates of linguistic rhythm in the speech signal☆

Franck Ramus[a,*], Marina Nespor[b,c], Jacques Mehler[a]

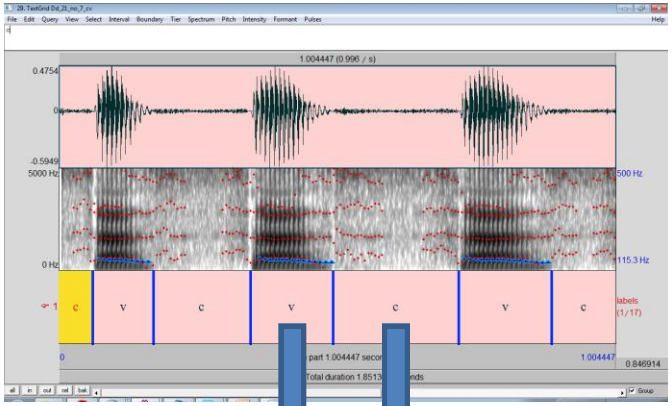
[a]Laboratoire de Sciences Cognitives et Psycholinguistique (EHESS/CNRS), 54 boulevard Raspail, 75006 Paris, France
[b]Holland Institute of Generative Linguistics, University of Amsterdam, Amsterdam, The Netherlands
[c]Facoltà di Lettere, Università di Ferrara, via Savonarola 27, 44100 Ferrara, Italy

Received 6 July 1998; accepted 14 September 1999

However, listeners might prefer syllabic cues when it comes to rhythmic information…

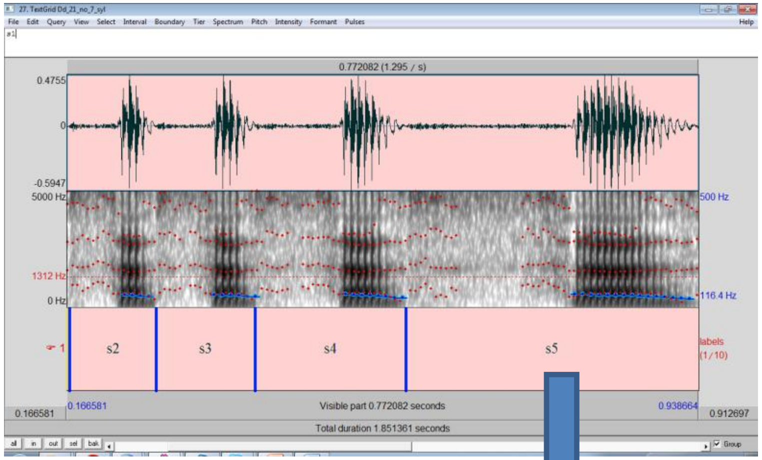
(a) CV sasasa

Dellwo (in print)

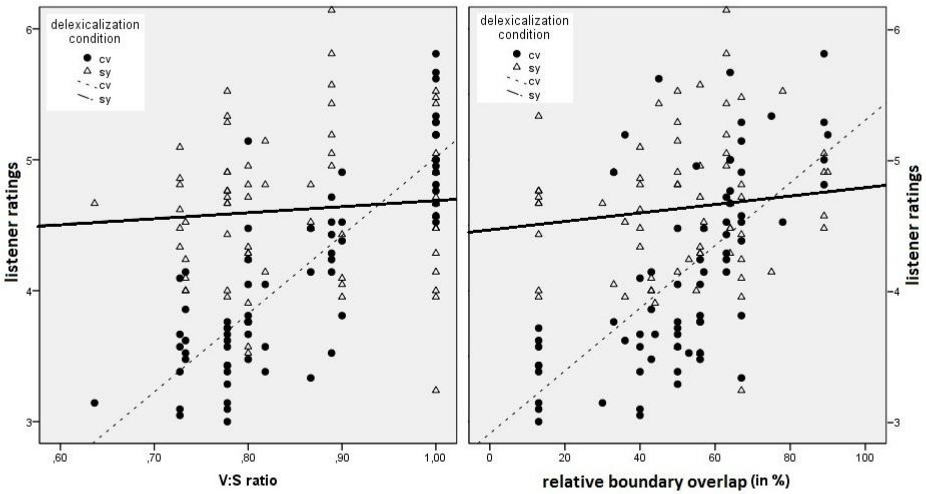


/s/ /a/

(b) syllabic sasasa



/sa/



We need correlates that are more perceptually salient in terms of speech rhythm…

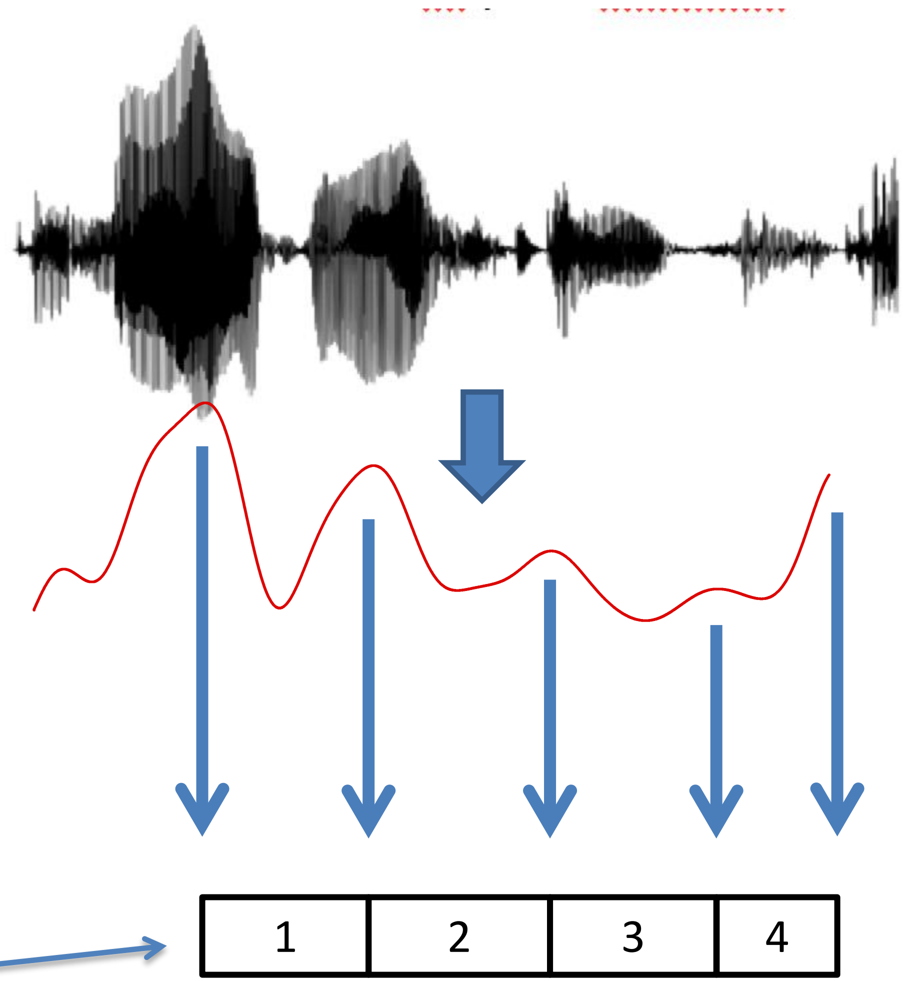
- **P-center work:**
  - Sophie Scott
  - Bernd Pompino-Marshall
- **Coupled oszillator work:**
  - Plinio Barbosa
  - Robert Port
  - Fred Cummins
- **Low-frequency fourier transform:**
  - Sam Tilsen
- **Voice source durational information:**
  - Volker Dellwo
  - Adrian Fourcin

AIM:

- Find way to extract low-frequency temporal (rhythmic) information from the speech signal that is easily and automatically applicable.
- Test whether there are between and within-language differences for Thai, Mandarin and Cantonese

Amplitude envelope

Peak-to-peak interval identification:



- Rectification of waveform
- Low-pass filtering (<10Hz)
- Peak-filter
- Flat peak deletion
- Empty interval detection
- etc.

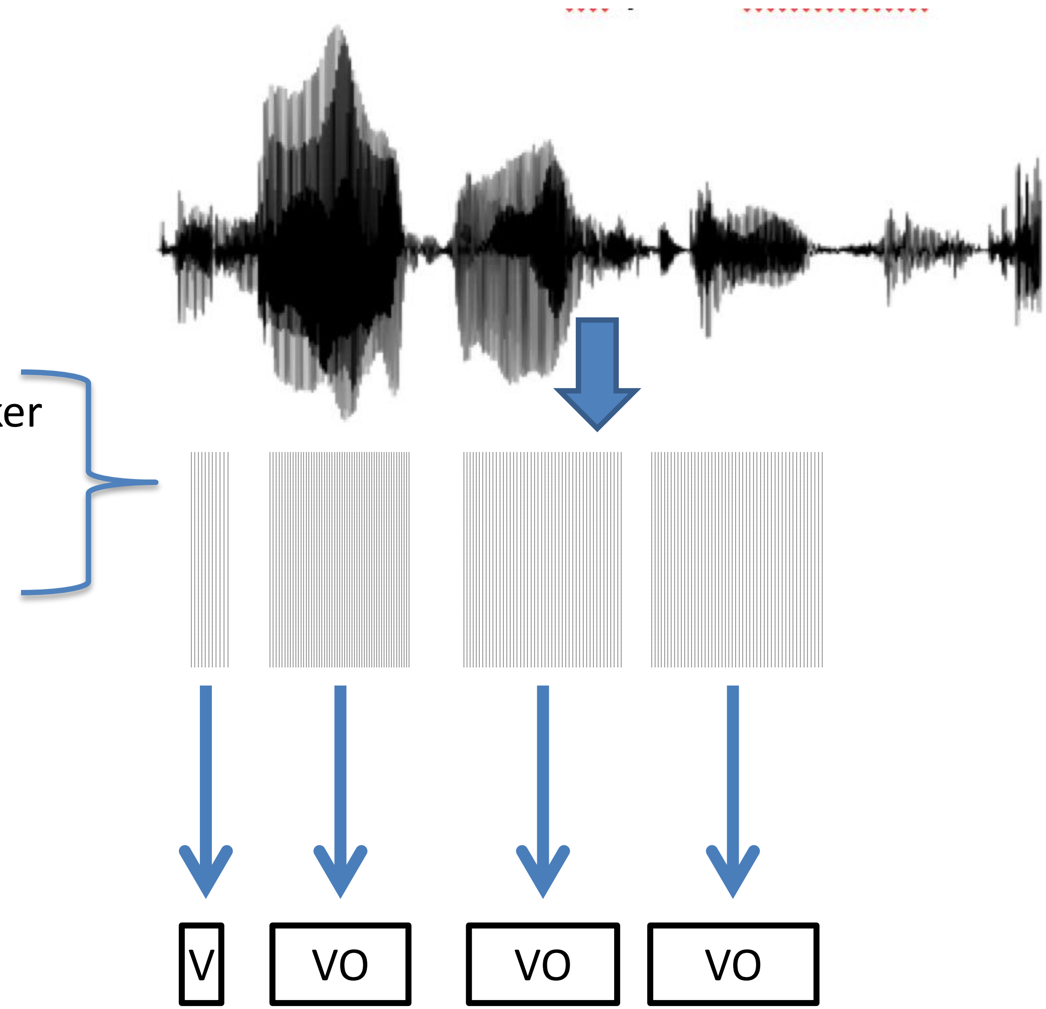
Approximation of p-center

Dellwo, Leemann & Kolly (2012)

Voice source

Voiced and unvoiced interval identification:

- Fundamental period marker
- Interval identification
- etc.

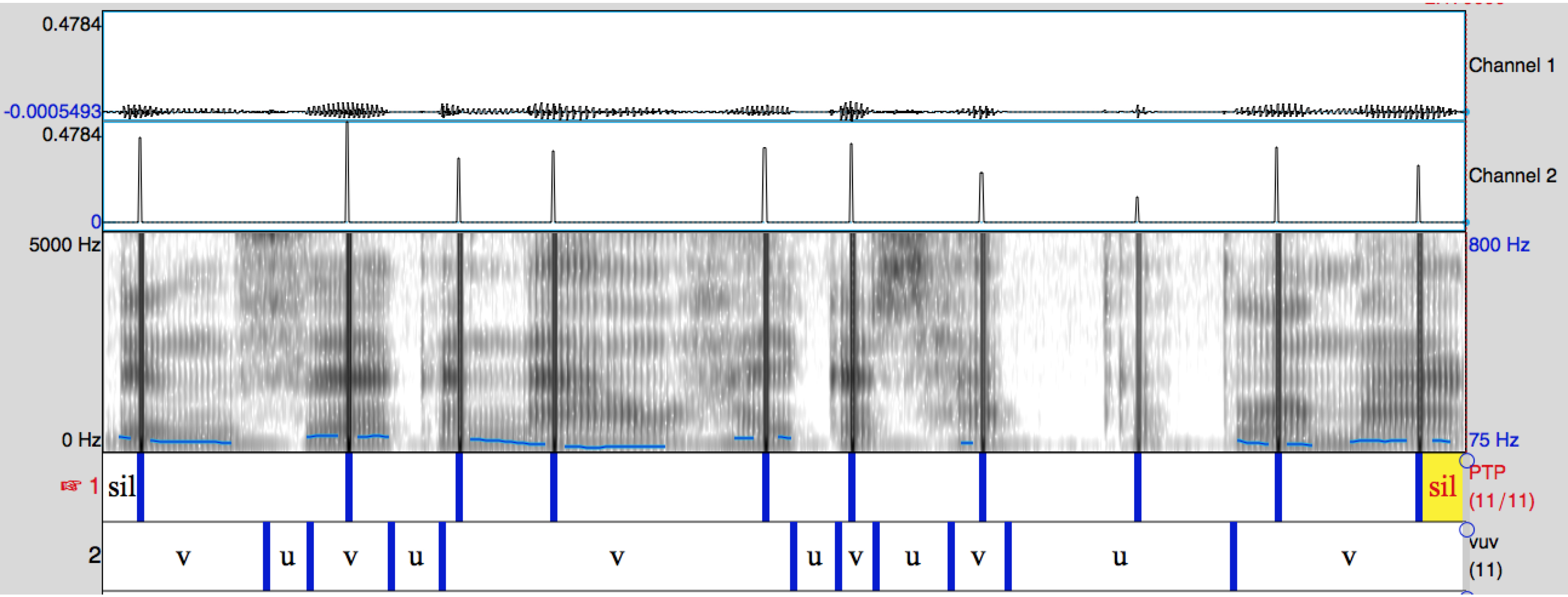


Dellwo, Fourcin & Abberton (2007)
Dellwo, Fourcin (in print)

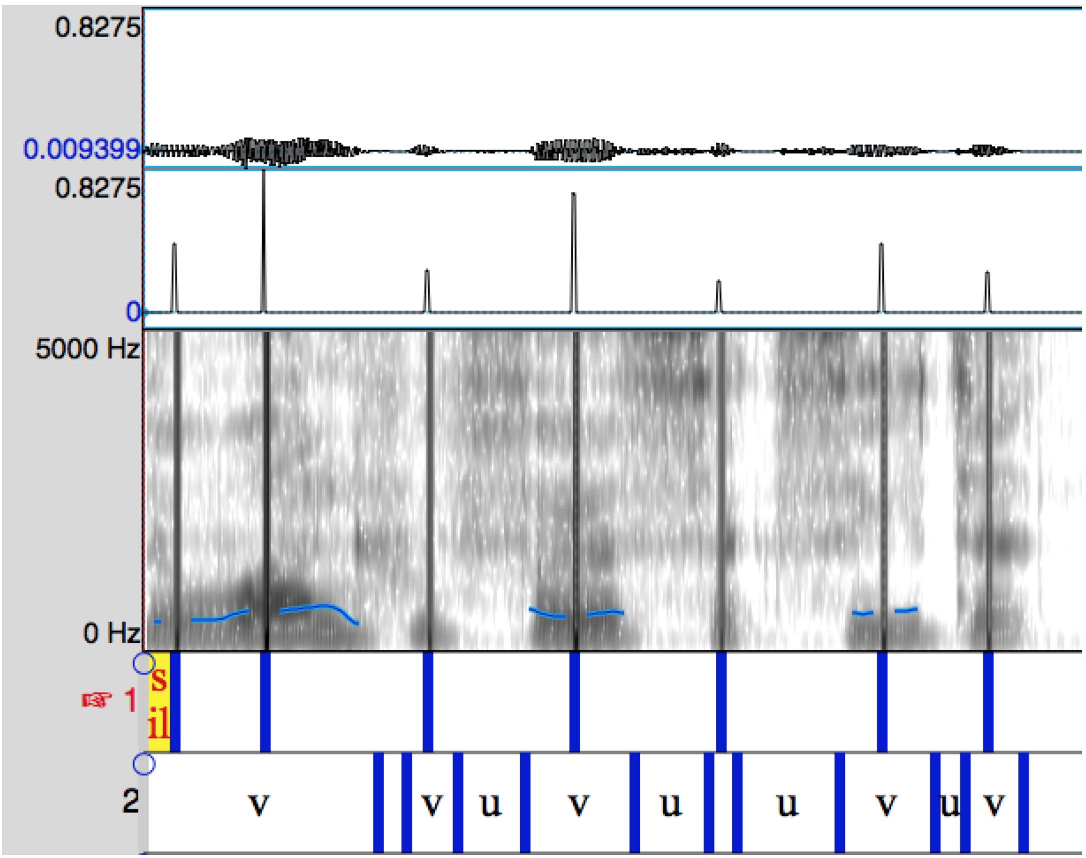
Languages

```
> summary(subset.seal.data$nativeLg)
cant mand thai
 142  136   73
> summary(subset.seal.data$discourseLg)
 l1  l2
189 162
> summary(subset.seal.data$speaker)
 1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19
22 26 16 16 20 20 22 14 13 20 19 16 23 11 20 16 15 21 21
```

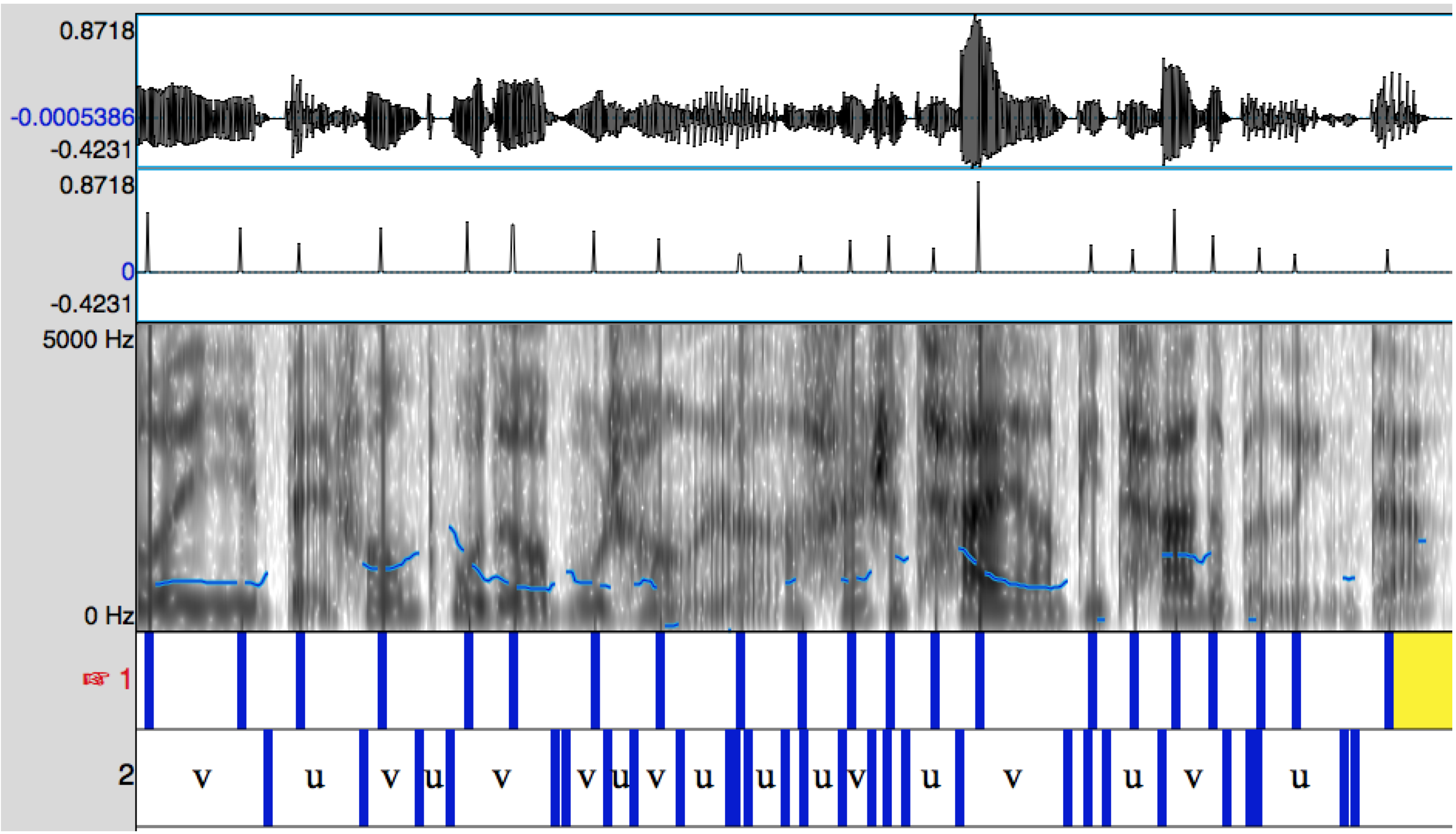
Thai


0.4784
-0.0005493
Channel 1
0.4784
0
Channel 2
5000 Hz
0 Hz
800 Hz
75 Hz
1
sil
sil
PTP
(11/11)
2
v
u
v
u
v
u
v
u
v
u
v
vuv
(11)

English by Thai native

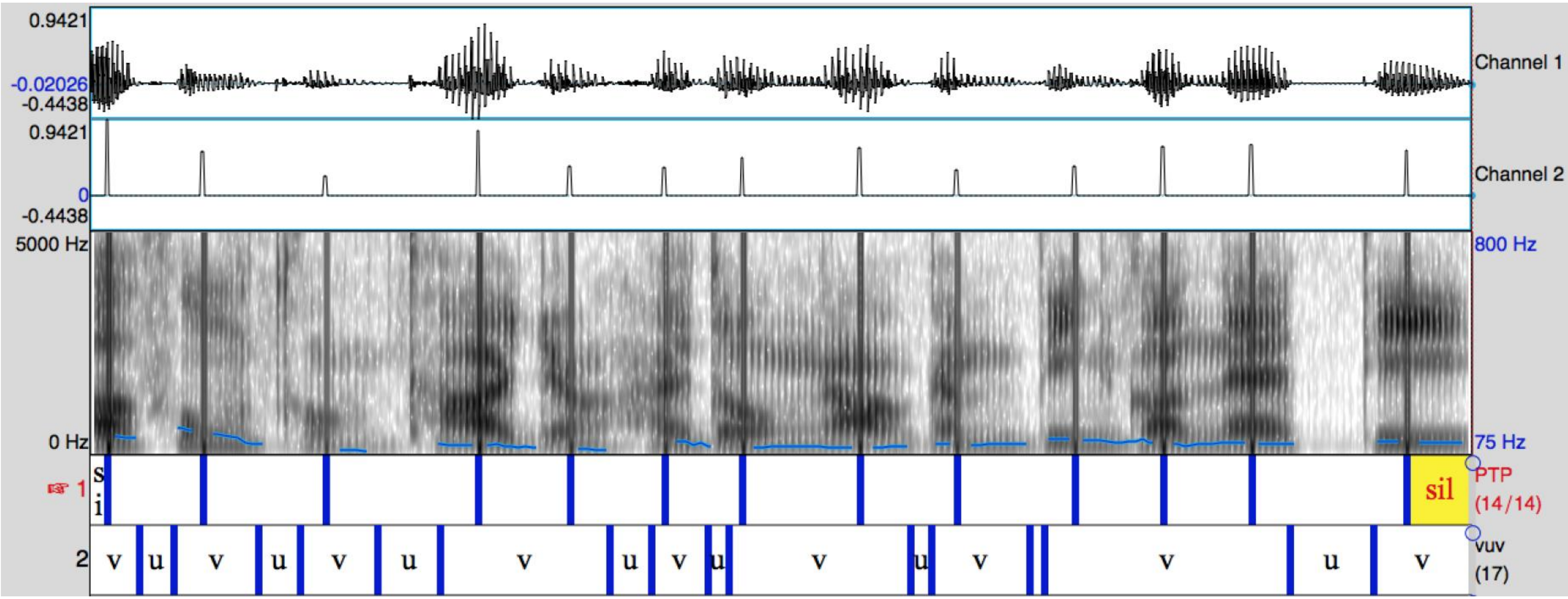


Mandarin



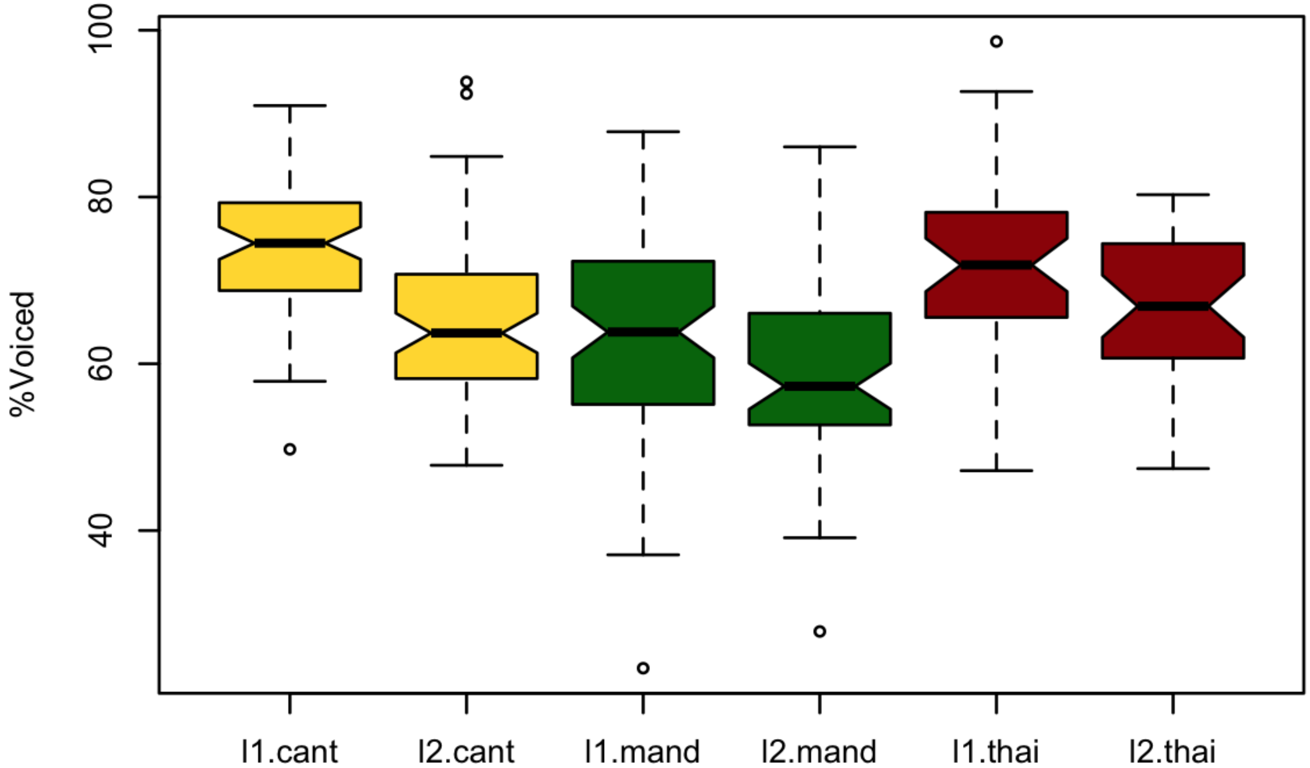
0.8718
-0.0005386
-0.4231
0.8718
0
-0.4231
5000 Hz
0 Hz
1
2
v u v u v v u v u u u v u v u v u

Cantonese



0.9421
-0.02026
-0.4438
Channel 1
0.9421
0
-0.4438
Channel 2
5000 Hz
0 Hz
800 Hz
75 Hz
1
si
sil
PTP
(14/14)
2
v u v u v u v u v u v u v v u v
vuv
(17)

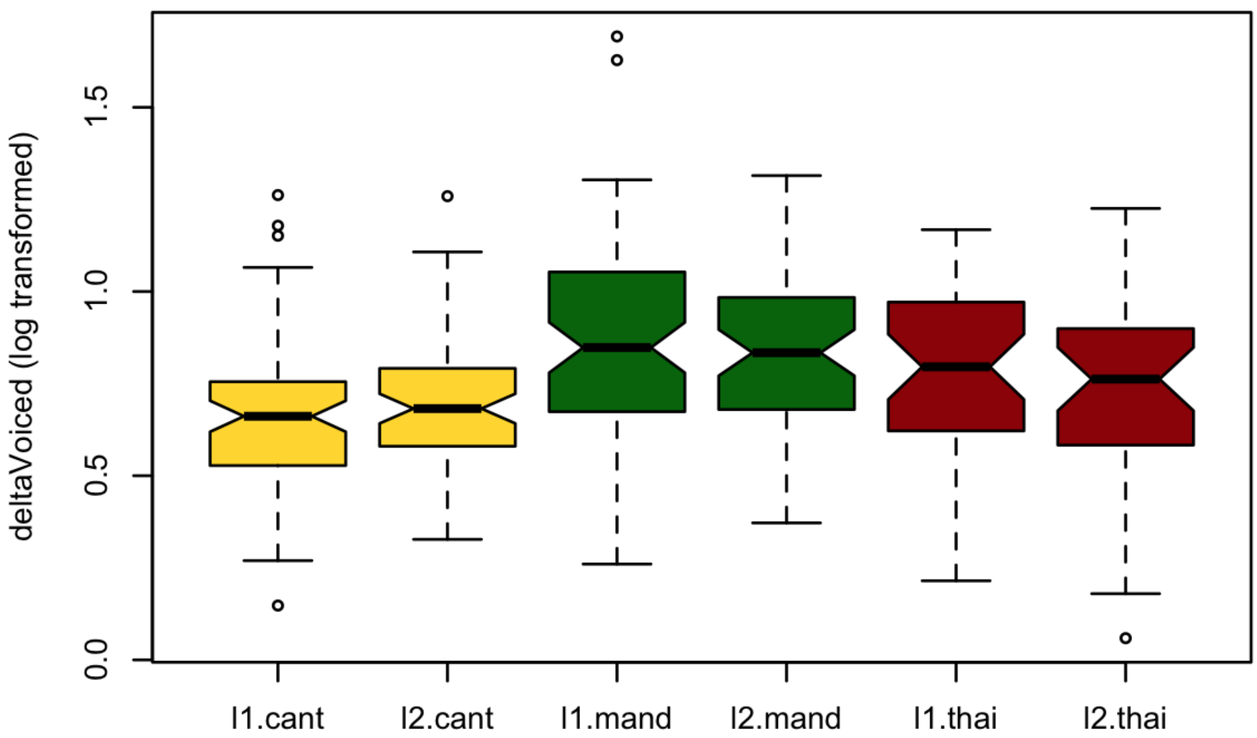
Voiced-voiceless ratio



	numDF	denDF	F-value	p-value
nativeLg	2	16	7.313	0.0055
discourseLg	1	329	40.556	<.0001
nativeLg:discourseLg	2	329	3.115	0.0457

Between-language variability: **yes**
Between-accent variability: **yes**

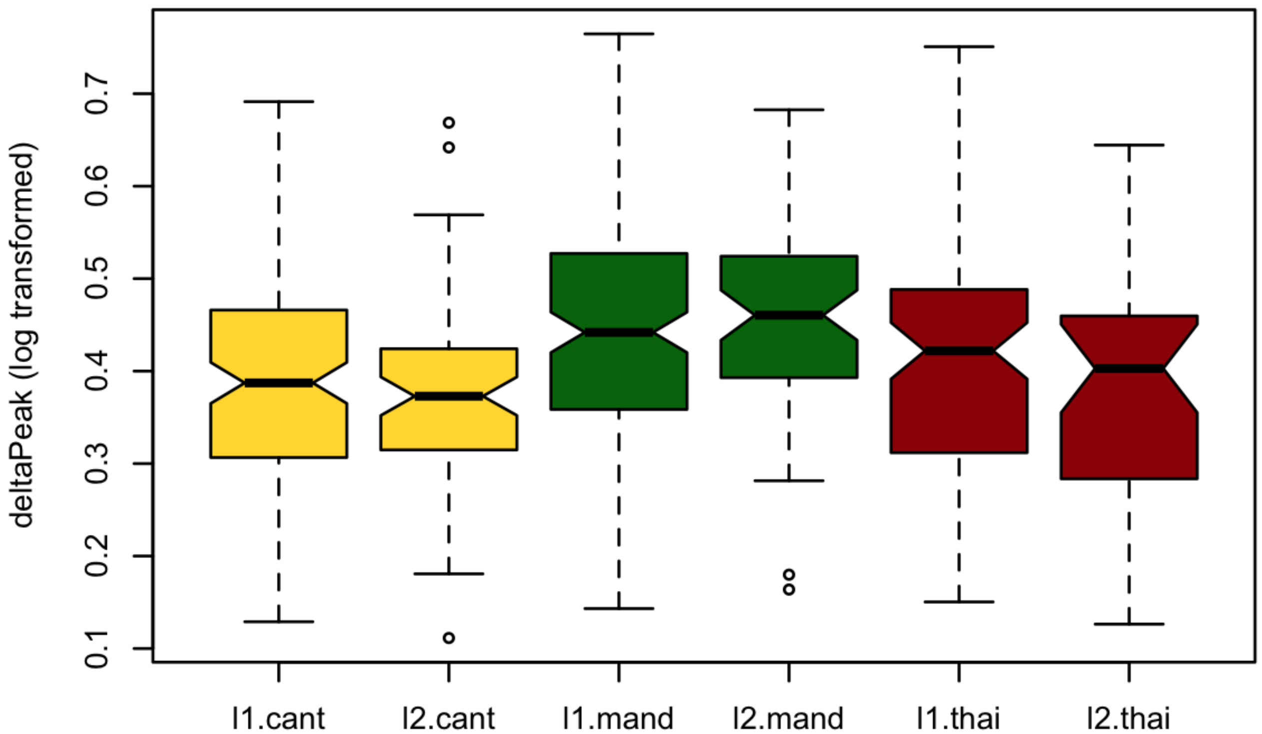
Durational variability of voiced intervals



	numDF	denDF	F-value	p-value
nativeLg	2	16	13.4332	0.0004
discourseLg	1	329	0.0207	0.8856
nativeLg:discourseLg	2	329	1.7981	0.1672

Between-language variability: **yes**
Between-accent variability: **no**

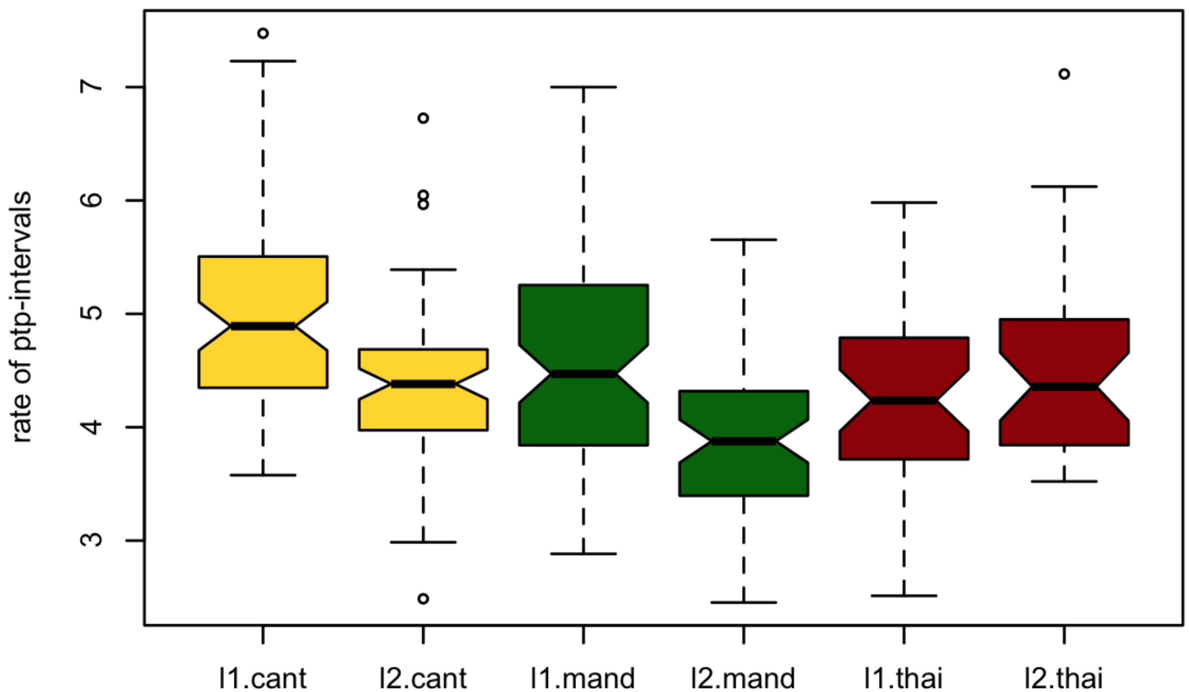
Peak-to-peak interval variability



	numDF	denDF	F-value	p-value
nativeLg	2	16	8.9036	0.0025
discourseLg	1	483	0.1587	0.6905
nativeLg:discourseLg	2	483	1.0551	0.3490

Between-language variability: **yes**
Between-accent variability: **no**

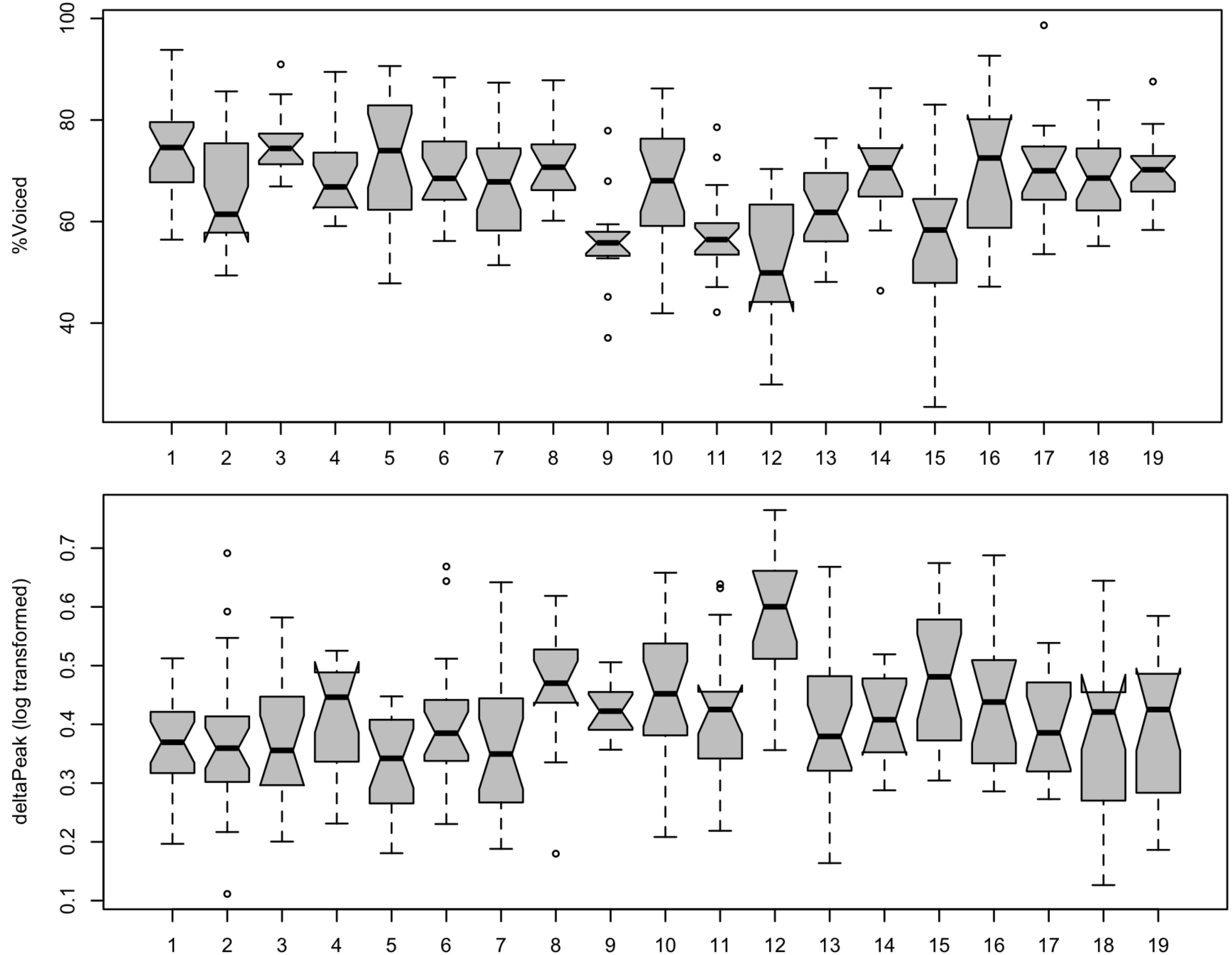
Peak-to-peak interval rate



	numDF	denDF	F-value	p-value
nativeLg	2	16	2.5283	0.1111
discourseLg	1	329	41.9046	<.0001
nativeLg:discourseLg	2	329	11.9854	<.0001

Between-language variability: **no**
Between-accent variability: **yes**
But: simple effects significant!

Strong variability between individuals…



cor(%Voiced, deltaPeakLn): -0.372

Conclusions

Summary:

- Automatically retrieved low-frequency temporal information from the speech signal might be useful in obtaining differences…
  - between languages (Thai, Cantonese, Mandarin)
  - accent varieties of the same language (L1, L2)
  - variability between speakers within the language
- Different measures may reveal different types of information.

Further work:

- Include other variables (in particular f0!) for p-center detection.
- Test whether obtained acoustic variability is perceptually salient
- This is a signal based approach. Does it tell us anything meaningful about the linguistic structure of these languages?