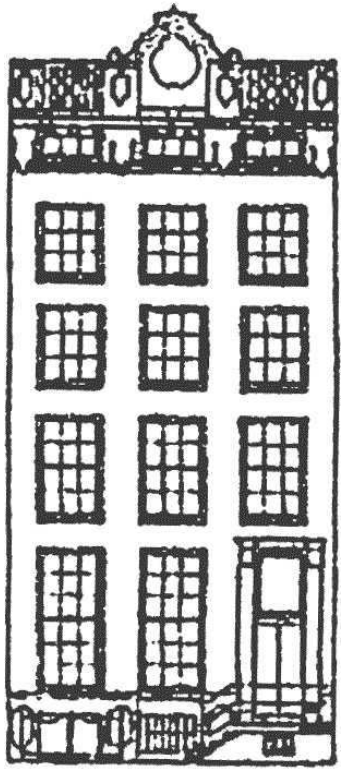
Vowel data of early speech development in several languages

Louis C.W. Pols

Elena Lyakso, Jeannette van der Stelt,
Ton G. Wempe & Krisztina Zajdó

Institute of Phonetic Sciences / ACLC
University of Amsterdam, The Netherlands



Multiling-2006
Stellenbosch,
April 10, 2006

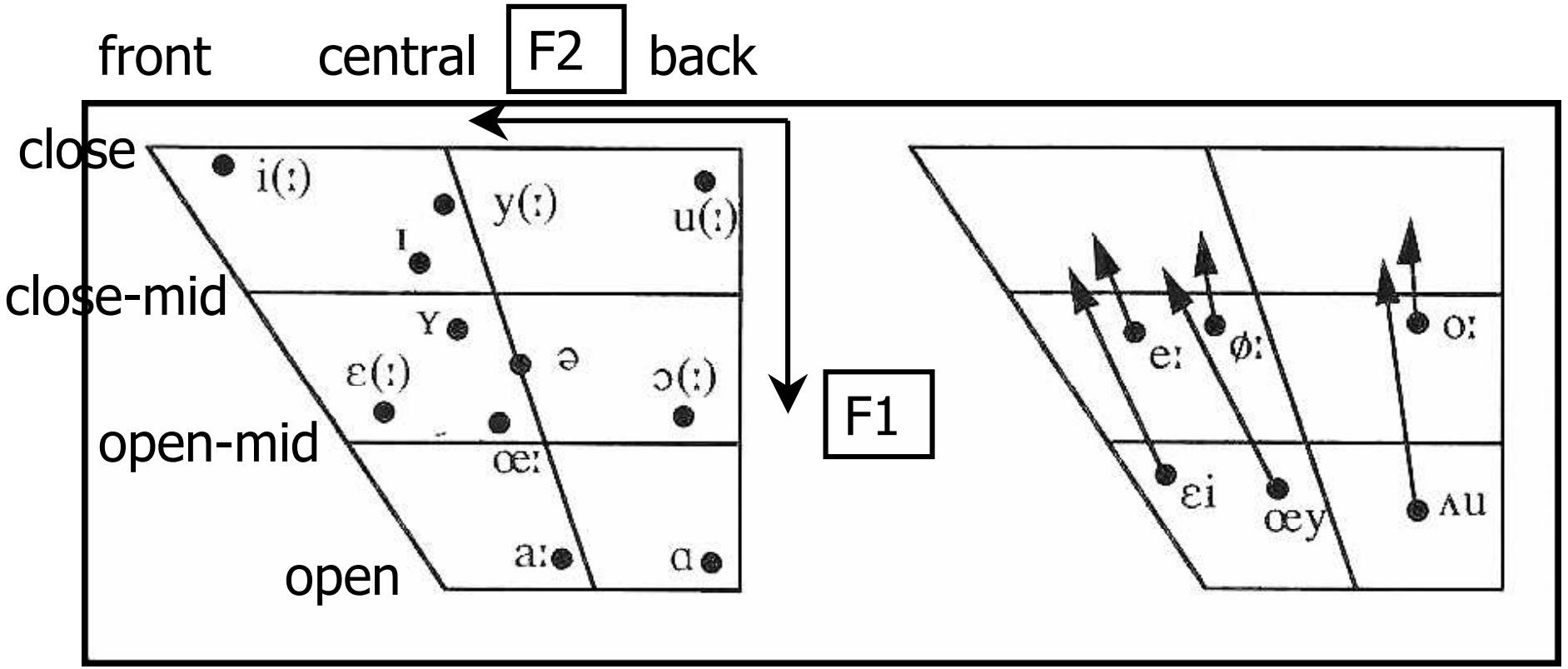
overview

- child speech and multilinguality
- size and form of acoustic vowel space
  - segments hard to measure, hard to label
  - universal vowel acquisition theory
  - detecting and treating disorders
- Dutch, Hungarian and Russian data (2 yrs.)
- analysis procedures (BF analysis, PCA)
- some results

child speech and multilinguality

- Multiling-2006 mainly engineering
- still other aspects of multilinguality
- babies can learn to speak any language
- how does articulation develop?
  - normative behavior; remediation
  - over age; within and between utterances
  - per language
- how to do spectro-temporal analysis?
- keep in mind: no words in baby speech

adult vowel space, Dutch

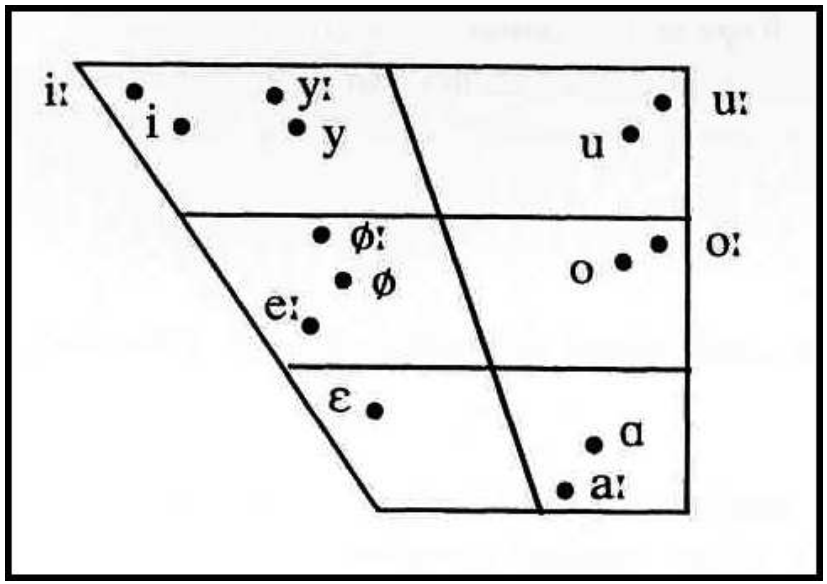


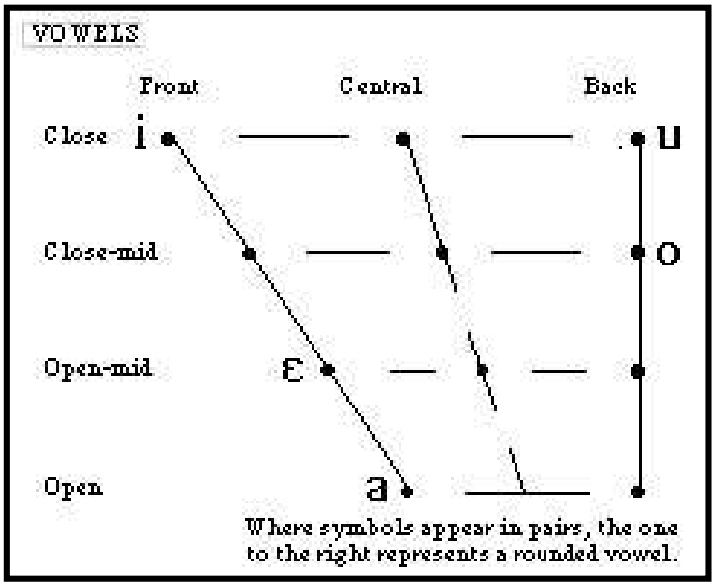
monophthongs

diphthongs and diphthongized vowels

(from IPA handbook)

adult vowel space, Hu & Ru





Hungarian

(from IPA handbook)

Russian

(from A. Lyovin (1997)

problems in analyzing child speech

- vocal tract under development
- big tongue
- highly variable articulation
- voicing irregular
- high pitch (400 Hz and more)
- formants poorly defined (few harmonics)
- vowel quality or words hard to identify

pitch-synchronous BF analysis

- formants are production-oriented
- we choose for perception-orientation
- bandfilter analysis on selected items
  - max. 10 per 'utterance', at proper pitch (< 425 Hz) and level (neither clipped (-0.5 dB) nor too low (-10 dB))
- recycle one pitch period up to 50 ms
- Kaiser window to reduce 20-Hz ripple
- apply pre-emphasis
- swept Gaussian bandpass filter analysis
  - 40 filters (0-7000Hz), step 175 Hz, BW 1.1 x 425 Hz
- level normalization

principal components analysis

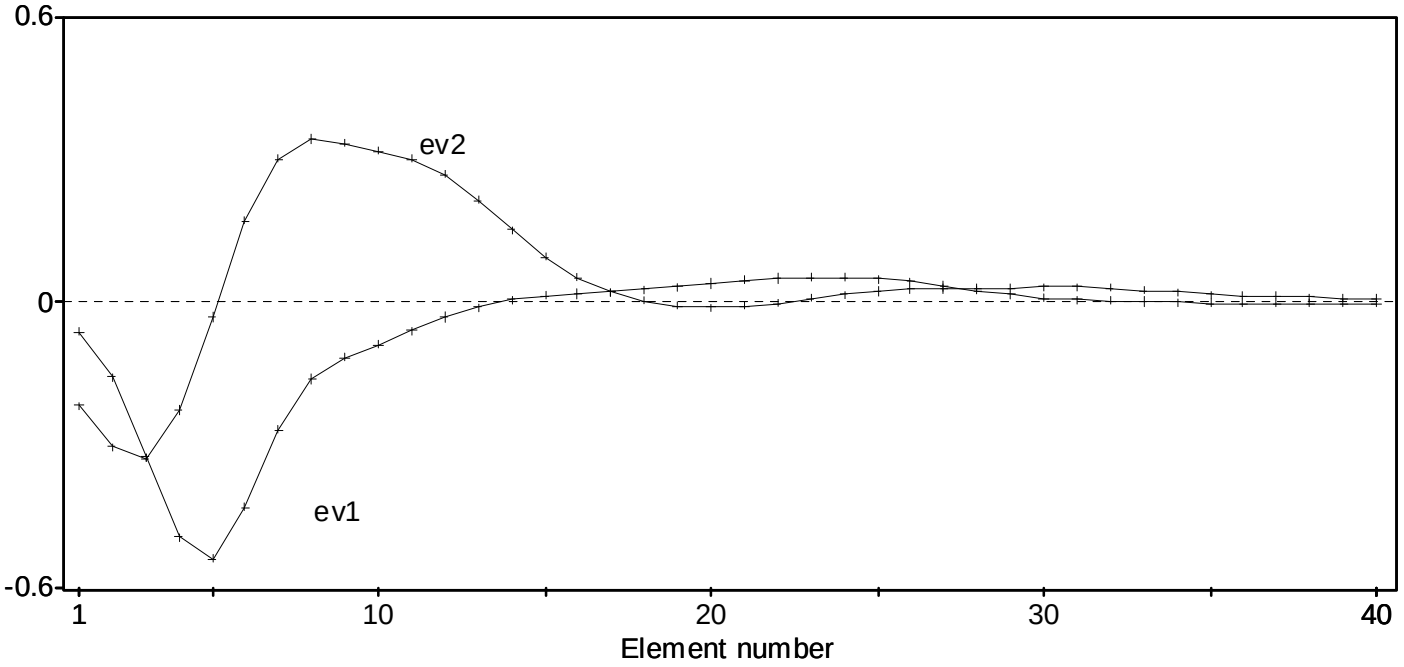
- each spectrum is a point in 40-dim. space
- we can use unlabeled data!
- determine joint Du-Hu-Ru representation
  - use equal numbers of points per language
  - find 2-3 dim. that explain most variance
- plot language-specific data
- find ways to specify size and structure
- language-universal and language-specific characteristics

available child speech recordings

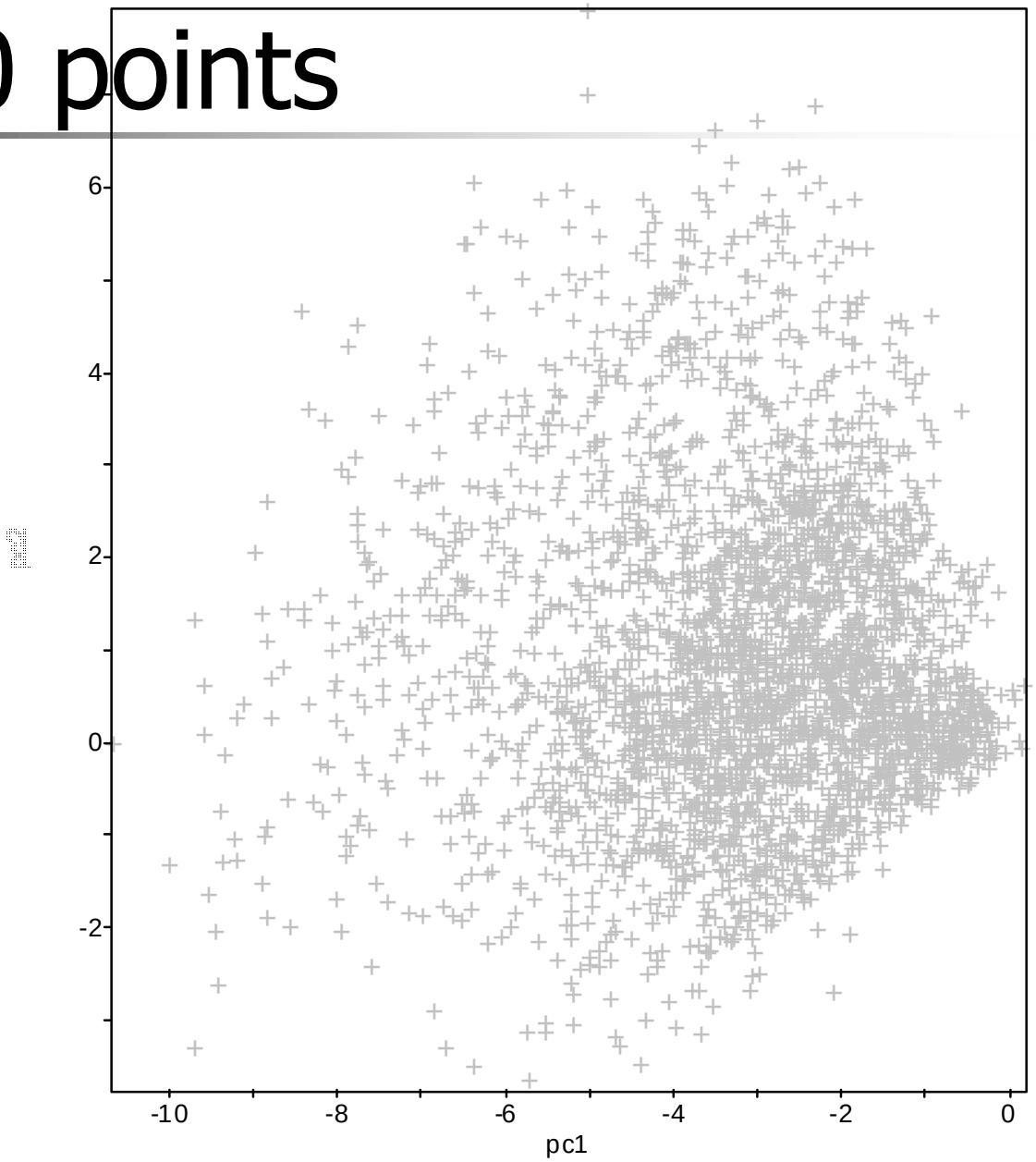
language	*elicitation*	*children (2 yrs)*	*# utter.*	*# segm.*	*# segm. used*
Dutch	free comm. in home situation	5 boys	5 x 50	1392	980
Hungarian	controlled interaction in the lab; puppets /pi:pi:/ /tu:tu:/	8 boys	763 —> 229	988	980
Russian	free comm. in home situation	5 boys	5 x 50	980	980

Du-Ru-Hu reference space

- 980 level normal. vowel spectra per language
- ev1 (27.95%), ev2 (25.62%), ev3 (13.37%)



3 x 980 points



labeled data

- for 2-yr olds, labeling is possible
- (if available) 15 correctly produced corner vowels /i:/, /u:/, /a:/, according to judgment of adult native listener

	Dutch	Hungarian	Russian
identified /a/, /i/, /u/	76/54/68	98/110/127	6/29/19
displayed /a/, /i/, /u/	15/15/15	15/15/15	6/15/15

980 Dutch spectra in joint space

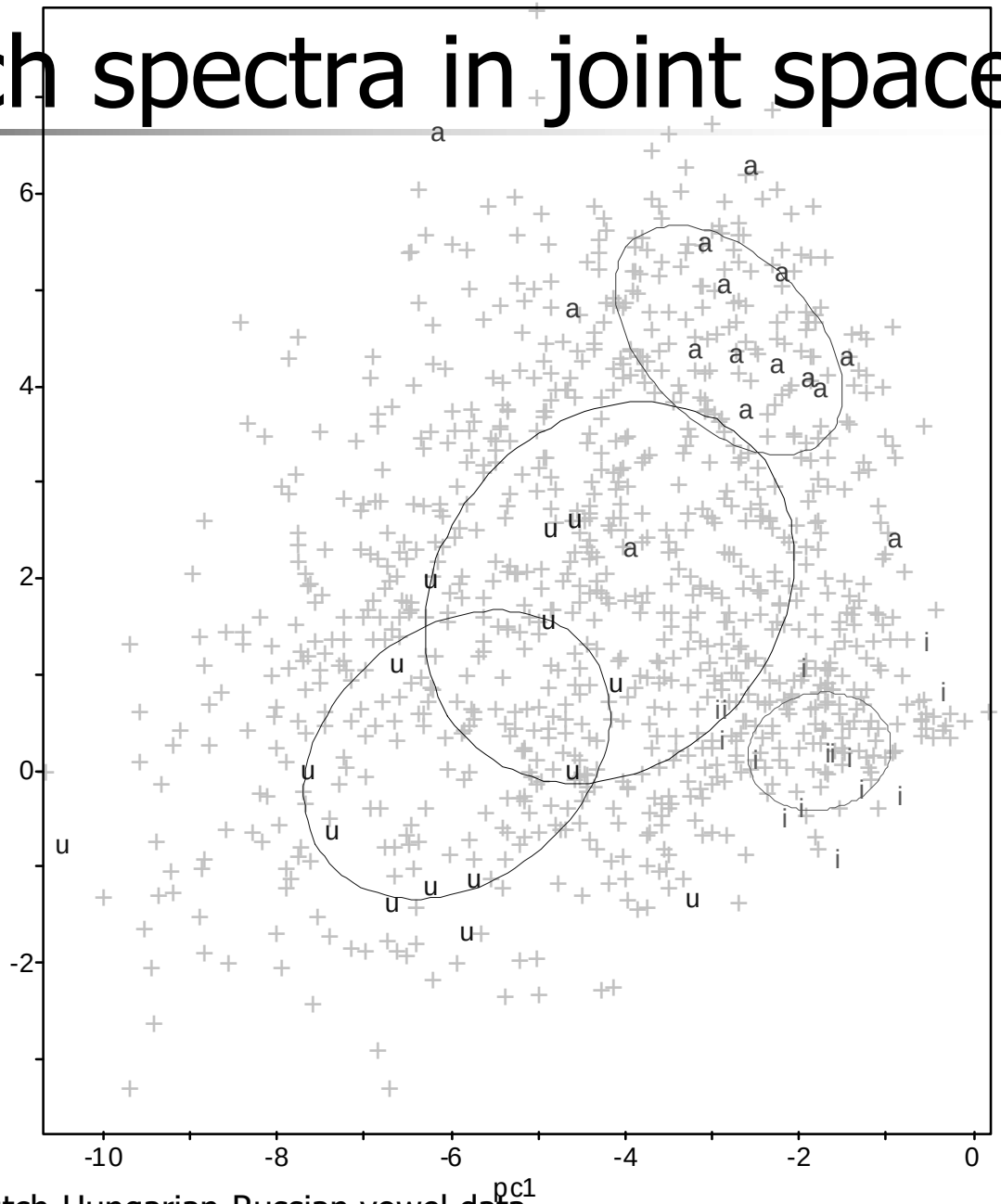
black: 1-sd all 980

labeled vowels:

blue: 1-sd 15 /u/

red: 1-sd 15 /a/

green: 1-sd 15 /I/



980 Hungarian spectra

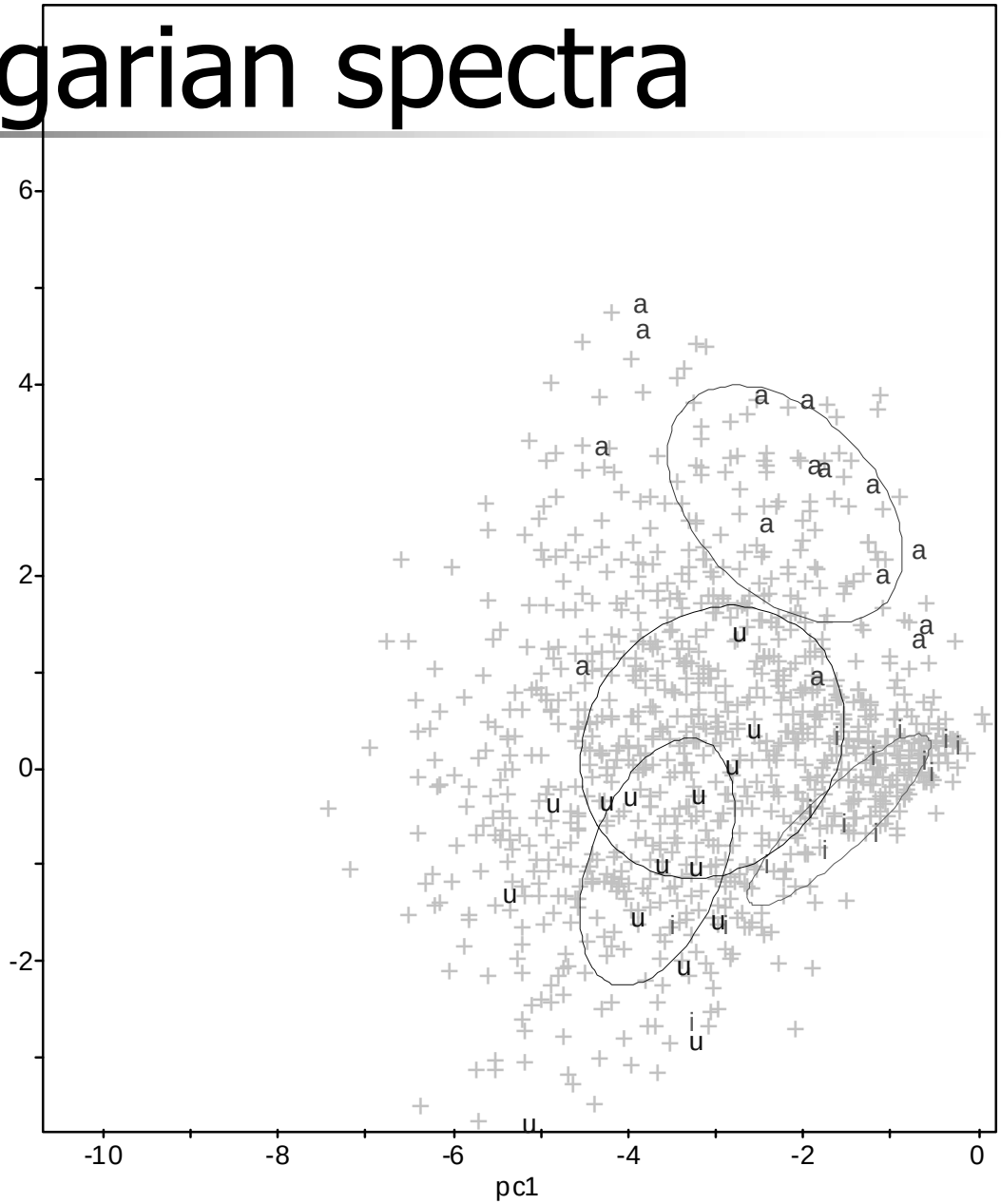
black: 1-sd all 980

labeled vowels:

blue: 1-sd 15 /u/

red: 1-sd 15 /a/

green: 1-sd 15 /I/



980 Russian spectra

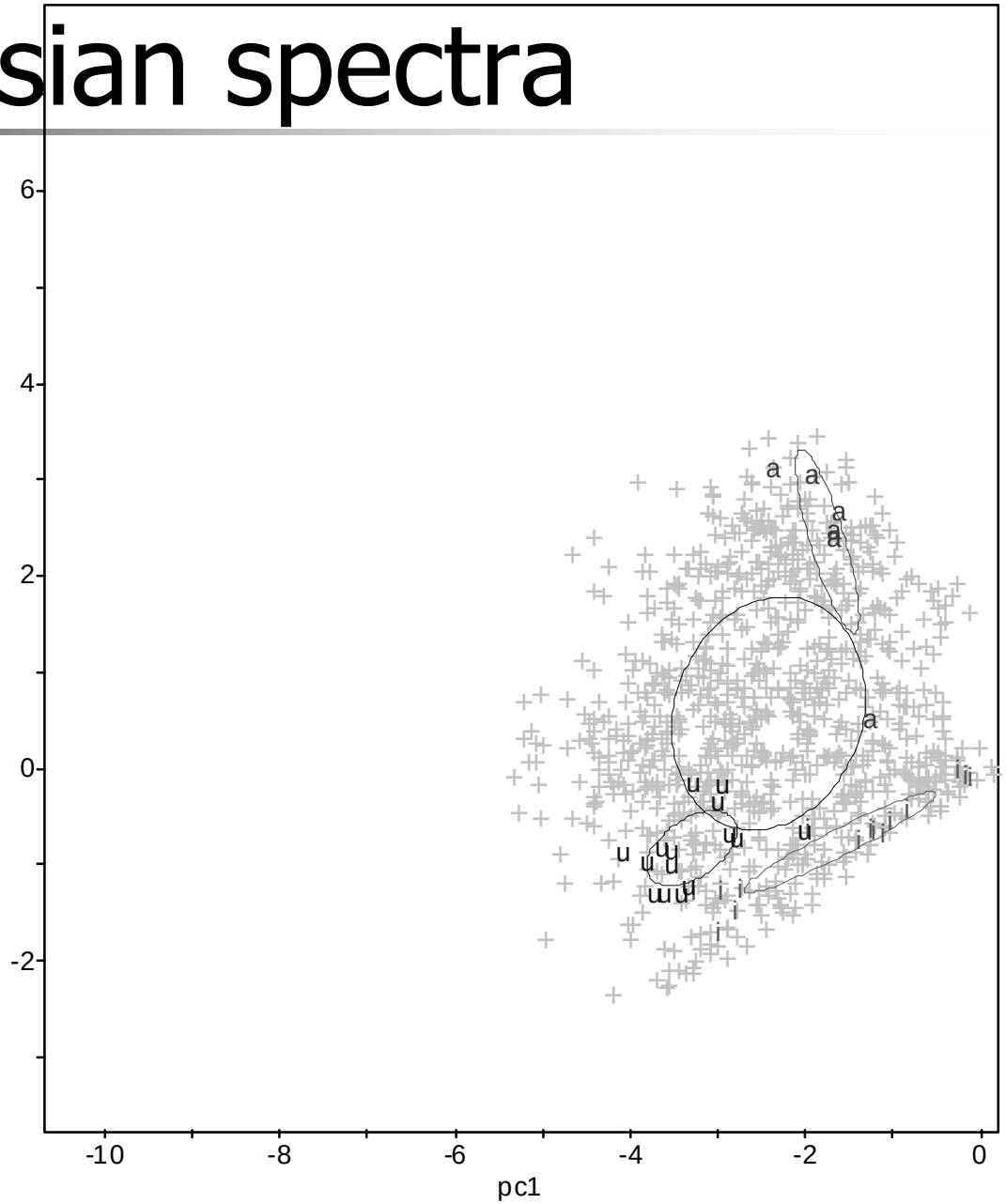
black: 1-sd all 980

labeled vowels:

blue: 1-sd 15 /u/

red: 1-sd 6 /a/

green: 1-sd 15 /I/



language-specific phenomena

- Dutch data most spread
- Russian data most clustered, still corner vowels clearly separated
- Hungarian and Russian more alike
- lip rounding *vs.* degree of closeness high V
- However: be aware of potential artefacts!
  - labeling strict, or based on whole utterance
  - highly variable quality recordings; effect of noise
  - also elicitation varied

how to characterize?

- evolving vowel system
- here, at 2 years of age
- similarity with F1-F2
- characterize overall size
- characterize local distributions of clusters
- so far: only one-standard-deviation ellipses
- potentially: development over age
  dynamicity within utterances

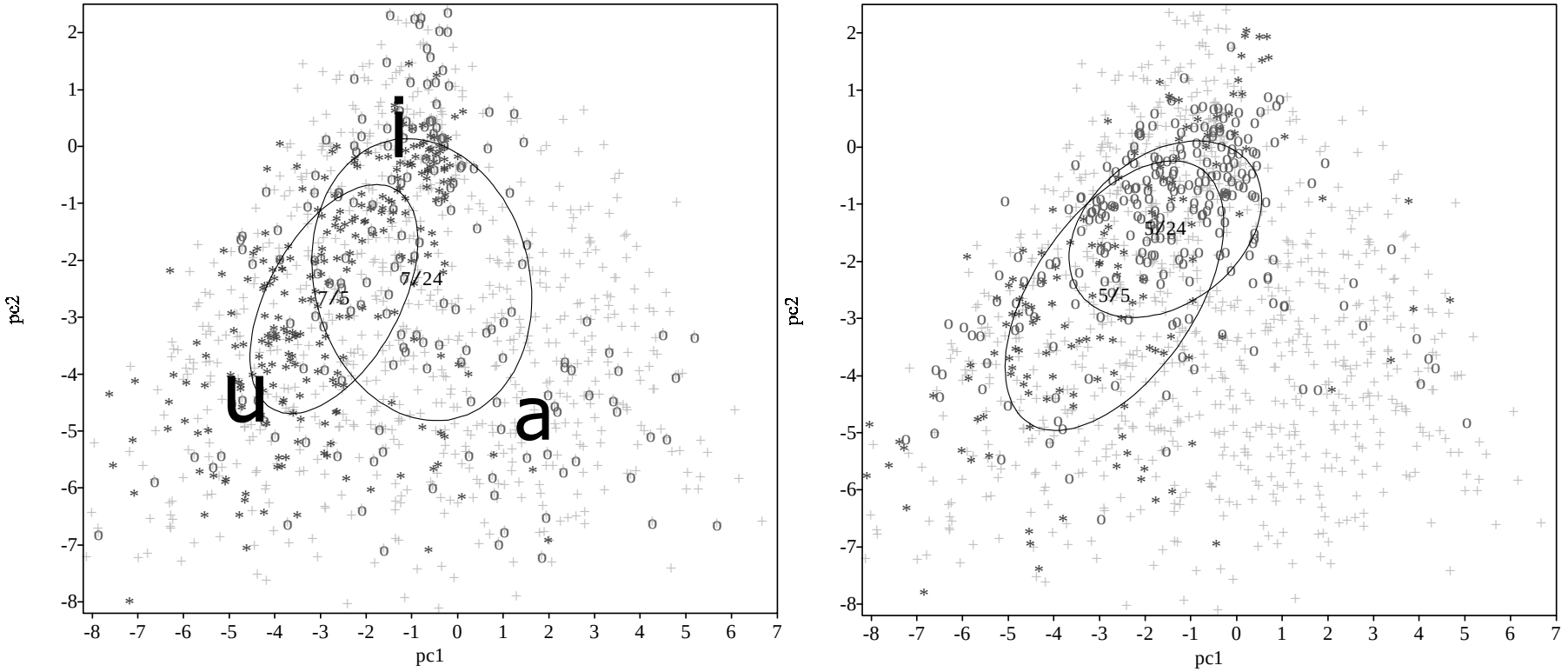
we need help from sp. techn.

- for characterizing overall size and local distribution of (unlabeled) clusters
- do better than one-sd ellipses
- could vector quantization and/or clustering be helpful?

future work

- normal *vs.* pathological speech
  - vocalizations of deaf children
- longitudinal development
  - 6 □□ 12 □ 18 □ 24 months
- spectro-temporal characteristics
  - within vocalizations / utterances
- more languages

normal vs. pathological speech



normal hearing child 5 & 24 mo.

hearing-impaired child 5 & 24 mo.

van der Stelt, Wempe & Pols (2003)