Development of a BOSS unit selection module for tone languages

Exemplification in Ibibio

Arne Bachmann & Stefan Breuer

University of Bonn

6th ISCA Workshop on Speech Synthesis

23rd August 2007

Ibibio corpus sample / synthesis

/ana ekop nsɔŋidem odo ɔjɔhɔ ɔjɔhɔ ke ifuuro uwem/

Original

Synthesized

Ibibio

Nigeria

Ibibio

State map

Ibibio

Tones

- High **H**
- Downstepped high **D** (!H)
- Low **L**
- Rising **R** (LH)
- Falling **F** (HL)

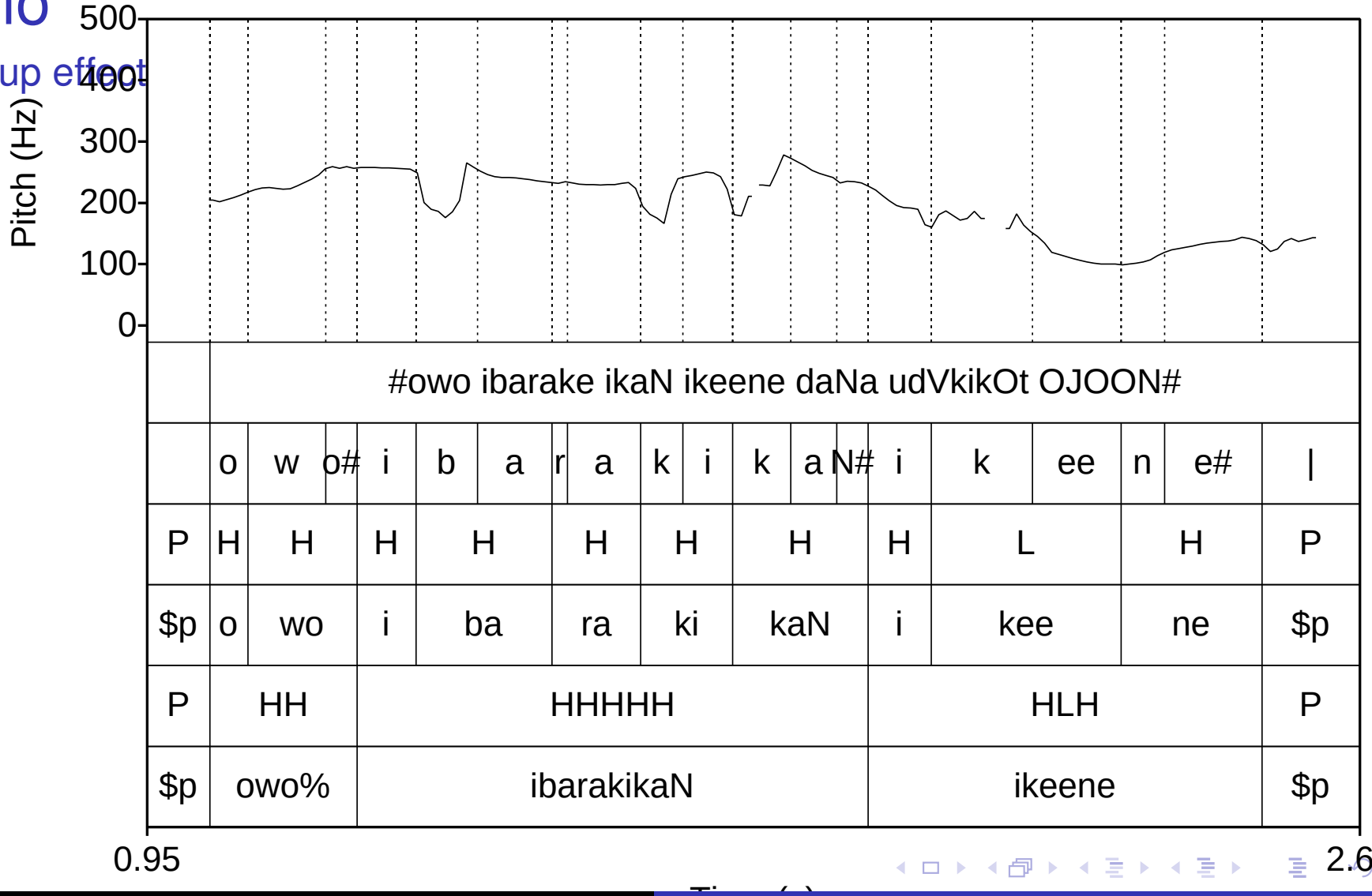
Ibibio

Tonal contrasts

óbóŋ	mosquito
óbòŋ	cane
sé	look
áà-sèè-hè	one who looks
áà-!ké-séé-hé	one who looked
áà-!dî-sé	one who will look

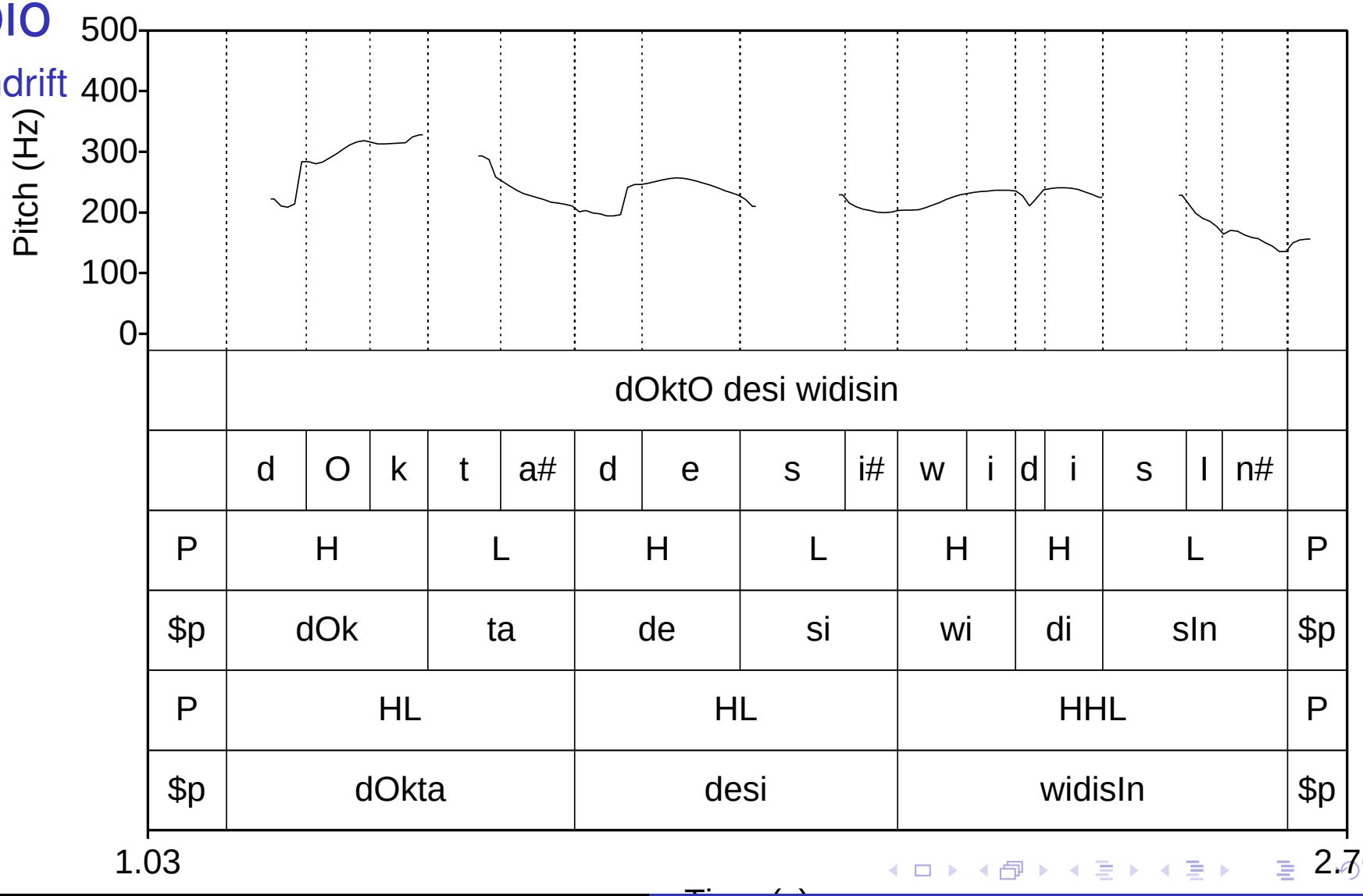
Ibibio

Start-up effect



Ibibio

Downdrift



BOSS

Bonn Open Synthesis System

- BOSS source code:
  `sourceforge.net/projects/boss-synth`

- BOSS website:
  `www.ikp.uni-bonn.de/boss/`

Goals

- Adaptability
- Extensibility
- Simplicity
- Universality

Goals

- Adaptability
- Extensibility
- Simplicity
- Universality

Goals

- Adaptability
- Extensibility
- Simplicity
- Universality

Goals

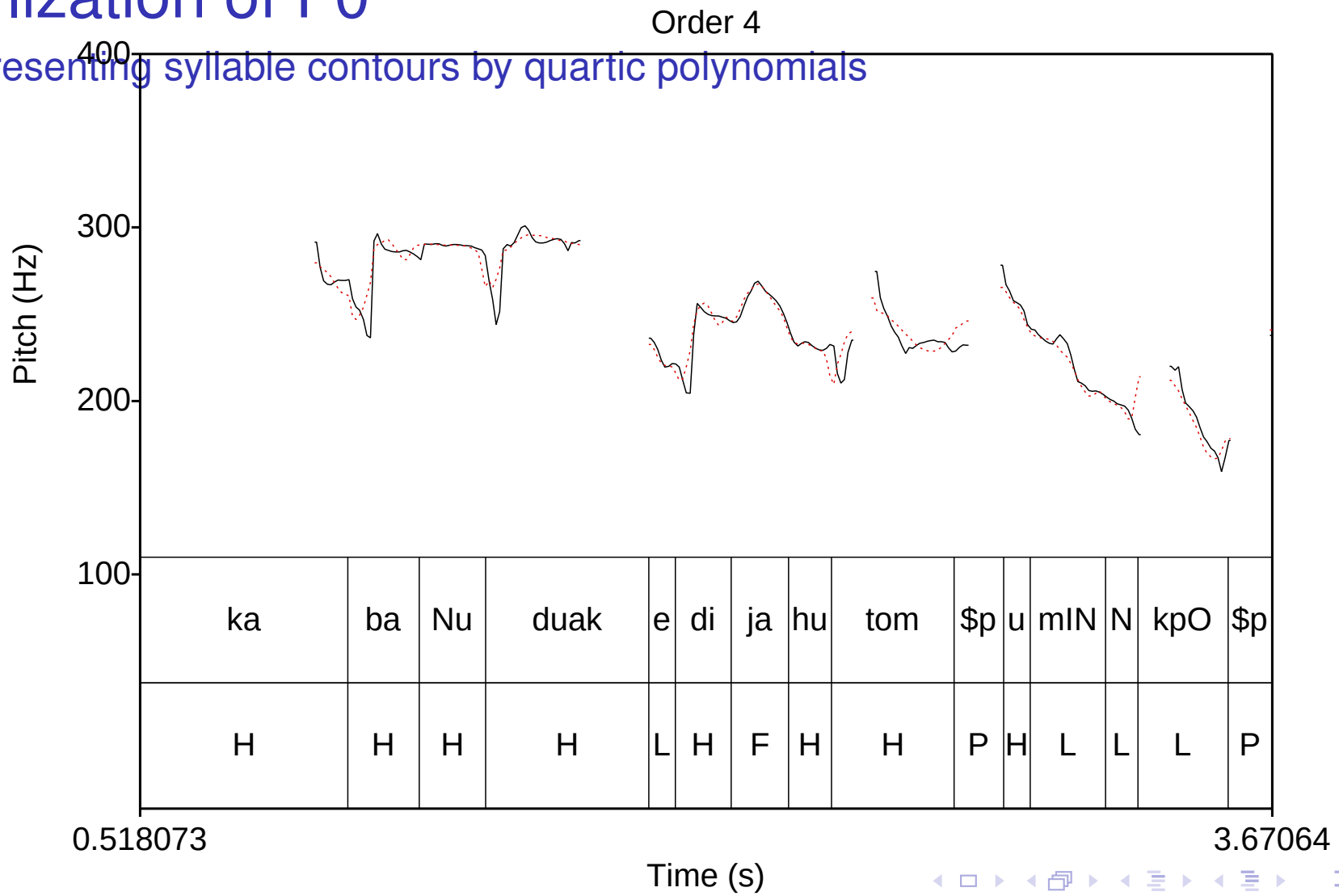
- Adaptability
- Extensibility
- Simplicity
- Universality

Stylization of F0

Linear model

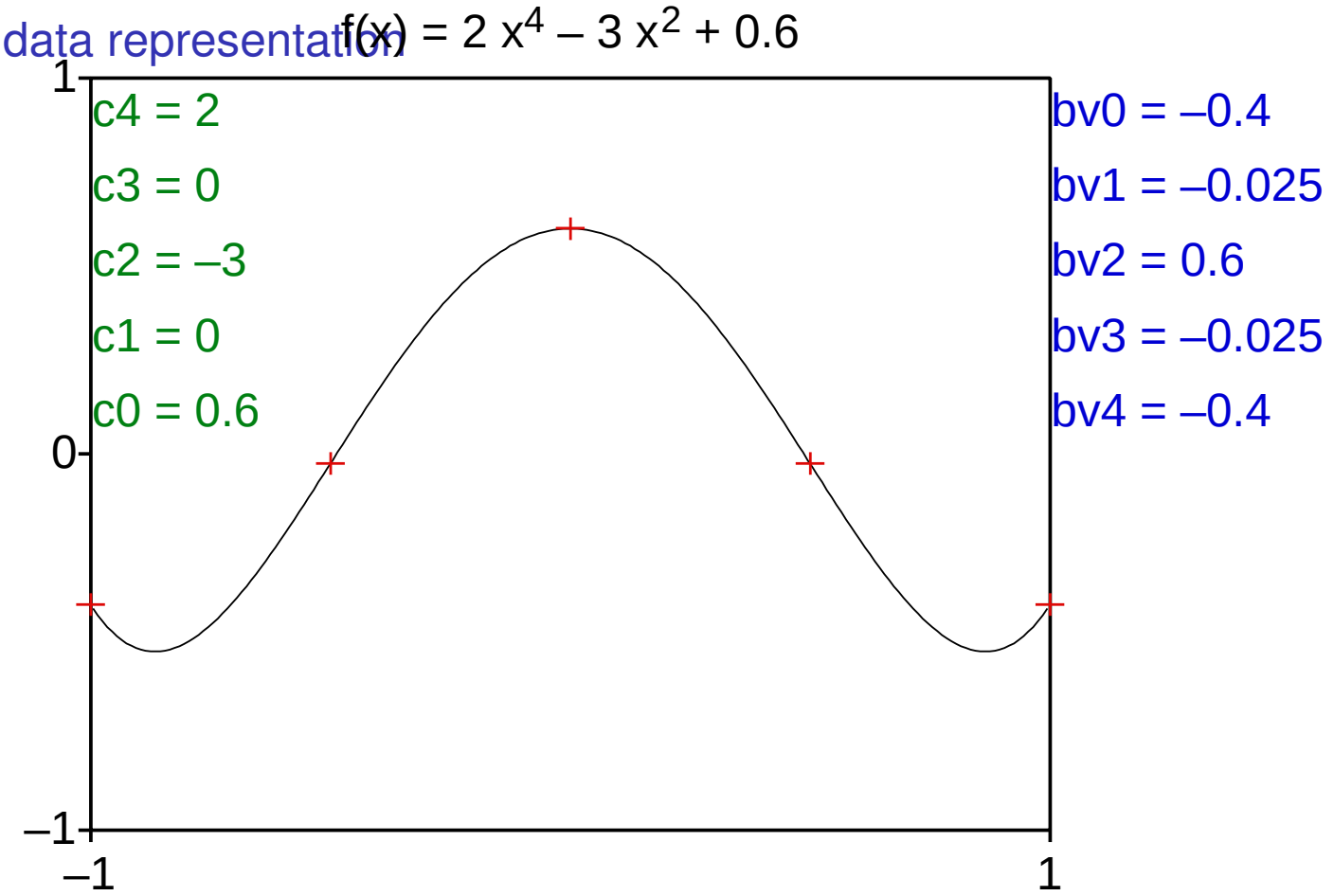
Stylization of F0

Representing syllable contours by quartic polynomials

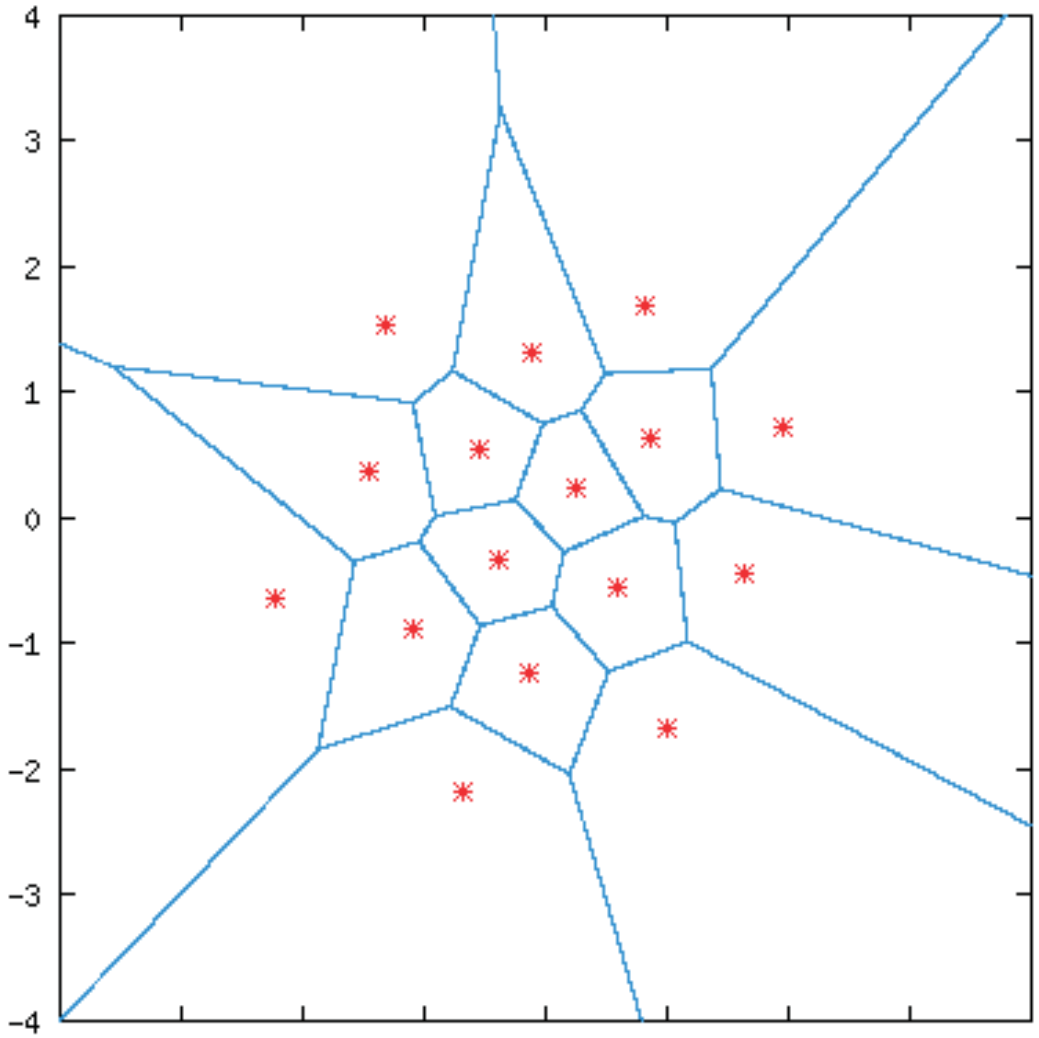


Stylization of F0

Basic value data representation

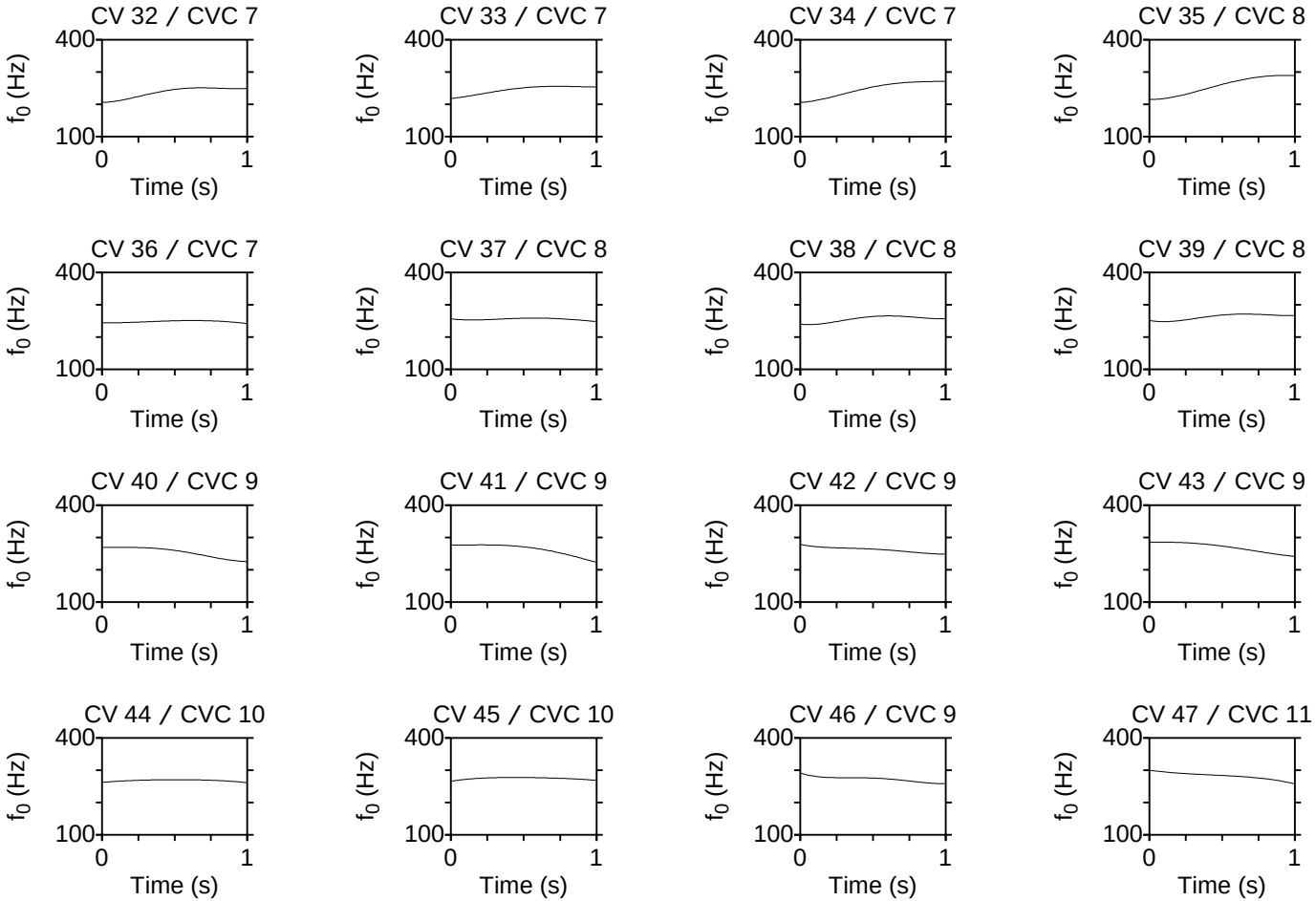


Vector quantization



Vector quantization

Code vectors 32 - 47 out of 64



Prediction

Requirements

Prediction

Most important parameters of a test tree

- Position of the syllable in the phrase
- Position of the word in the phrase
- Number of phones in the syllable
- Syllable structure (e. g. C, V, N)
- 4^{th} left neighbouring tone

- Remaining

Prediction

Most important parameters of a test tree

- Position of the syllable in the phrase
- Position of the word in the phrase
- Number of phones in the syllable
- Syllable structure (e. g. C, V, N)
- 4^{th} left neighbouring tone
- Remaining

- 62.96 %
- .
- .
- .
- 67.82 %
- 71.89 %

Prediction

Example from CART

```
((sylphrase < 15.7)
 ((sylphrase < 8)
  ((sylphrase < 3.4)
   ((ltone2 is P)
    (ltone1 is P)
     ((phonessyl < 1.3)
      (50)
      (53)
     (50)
    ((f < 0.1)
     ((wordphrase < 1.2)
      (45)
      (catsylword is i)
       (45)
       (50)
```

Prediction

Tone template classification results

On 10 different test sets (10 % held-out data):

- Correct code vectors (out of 64):
  - 38.55 % - 59.04 %
- Correct code vector classes (out of 16 in second codebook):
  - 42.77 % - 68.07 %

Unit selection

Conclusion

Conclusion

To do

- Get hold of a bigger corpus
- Perform subjective listening tests
- Adapt the module to another tone language
- Try the approach for an accent language

Conclusion

To do

- Get hold of a bigger corpus
- Perform subjective listening tests
- Adapt the module to another tone language
- Try the approach for an accent language

Conclusion

To do

- Get hold of a bigger corpus
- Perform subjective listening tests
- Adapt the module to another tone language
- Try the approach for an accent language

Conclusion

To do

- Get hold of a bigger corpus
- Perform subjective listening tests
- Adapt the module to another tone language
- Try the approach for an accent language

Thank you!