

Database creation – experiences and requests

Hans Kruschke

TU Dresden

*Laboratory of Acoustics and Speech
Communication*

Hans.Kruschke@ias.et.tu-dresden.de

Creation of small diphone databases

- **Specification of desired units**
- **Specification of text material:**
 - Lexical words containing desired diphones
 - Several alternatives for each diphone
- **Controlled recording of monotone speech**

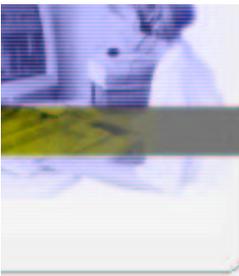
Creation of small diphone databases II

- Setup of data base: selection of best alternative of each diphone (semiautomatically)
- Automatic labeling with manuel correction
- Algorithm based marking of pitch periods
- German English



Problems with small diphone databases

- **Limited possibilities to simulate variations in prosody (e.g. with PSOLA), voice quality, etc.**
- **Disturbances at concatenation points**



Possible improvements

- **Variation of unit size**
- **Multiple alternatives for each unit**

Variation of unit size

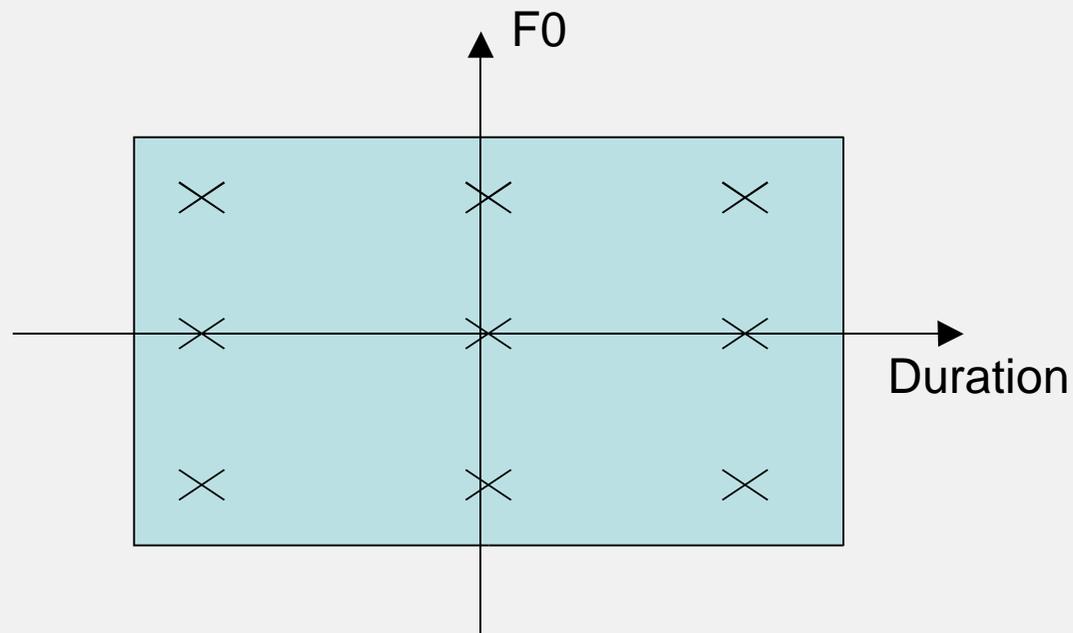
- **Example: Chinese (Mandarin)**
 - 420 pure syllables
 - 1200 syllables with consideration of 4 Tones
 - 3900 syllable based units with consideration of coarticulation
 - **German: 8000 – 10000 syllables**
- Which unit size for construction of text material ?**
- Diphones
 - Semisyllables

Reasons for acoustic- phonetic variations in natural speech

- Prosodic variation
- Varying precision of articulation
- Variation in speaking style
- Variation in emotional state

Proposal

Clustering of relevant variations to specify desired unit alternatives



Recording of variations by positioning

- Large amount of text material
- Each unit placed at different positions which will induce desired variations
- Possible positions are:
 - accented vs. unaccented
 - content vs. function word
 - initial, medial, final sentence position
 - ...

Recording of variations by speaker instruction

- Limited amount of text material
- Each unit recorded with different instructions / stimulations which will induce desired variations
- Possible instructions are to speak
 - Fast
 - Slow
 - High
 - Low
 - ...



Summary

- **Description of present method for database construction and its problems**
- **Variation of unit size for construction of text material**
- **Definition of requested alternatives**
- **Recording of variations by speaker instruction**