AhoTTS Conversion System for the Basque Language

Inma Hernández, Eva Navas, Iñaki Gamindo, Borja Ebexbarri, Imanol Madariaga

University of the Basque Country

GENERAL SETTINGS:
- Language: Basque or Spanish
- Dictionary name
- Intonation model: Peaks & valleys or Fujisaki’s
- Duration model: Rules based or statistical
- Pitch Mean & StdDev
- Speaking rate
- Volume
- Speaker
- Synthesis Method: MROLA, FSOLA, Formants
- Time labels: Unit/Phone/OLA level
- Spectral Smoothing Y/N
- Power Smoothing Y/N

GENERAL FEATURES:
- Modular architecture - allows the use as a developing tool & as an API
- Information flow is unidirectional: each module adds information to the data stream
- Multilingual architecture - presently developed mainly for Basque
- Multiprocessor: Windows/VC6.0 & Linux/GCC
- Multi-thread

INPUT TEXT:
- ISO 8859-1 Latin
- Acronyms, numbers, dates, times accepted
- Escape sequences for manual control currently breaks, in the future other parameters

DICTIONARY:
- Morpho-syntactic information
  - List of verbs
  - List of suffixes and prefixes
  - Acronyms’ expansion
  - Information about breaks
  - Stress information
- Up to 14 different groups of properties for every entry
- Currently contains 10000 entries
- Two data formats:
  - Binary & fast looks
  - ASCII for customer data

SYNTHESIS UNITS:
- Di/Tri/Tetra/phones, & pseudo-phones
- Wave or Multiband-Re-synthesized unit
- Sony’s database - 900 units

Text processing

Linguistic processing

Dictionary

Unit selection

Synthesis engine

AhoTTS

- Syntactic info
- Words to characters
- Grapheme to phoneme
- Syllable boundary
- Stress
- Segmental duration
- Pitch
- Power

- Selects units based on unit length
- Moves prosodic information from phones to units
- OLA synthesis with prosody modifications
- Allows spectral smoothing at the concatenation points

Synthetic waveform

Power

Spectrogram

Synthetic phone labels

Synthetic unit labels

Text