SSML for Urdu Speech Synthesis

Sarmad Hussain
Professor and Head
Center for Research in Urdu Language Processing
National University of Computer and Emerging Sciences, Lahore
Background

• Comments based on experience in developing a concatenative Urdu TTS system
  – inputs Urdu Unicode or UZT text and outputs Urdu speech
  – performs textual analysis, normalization and phonological processing to convert the text to speech
  – uses 5000+ diphone database
  – models duration and intonation based on learning through tagged corpus
  – complete framework developed by CRULP through funding by Ministry of IT, Govt. of Pakistan
Using SSML for Urdu Speech Synthesis

• SSML standards works well for Urdu
• SSML needs to be enhanced to handle some Urdu specific problems
  – Multilingual text
  – Digits
  – Date formats
  – Diacritics
  – Word Segmentation
Multilingual Text

• Need to identify one of the three strategies (for different contexts)
  – Switch to English TTS
  – Transliterate and Read
  – Spell-out
محمد الشباب، برزگان
کیلی تصویر کپسولات، ناشر کپسولات آکاکی کپسولات
هموئینت اختیاری کپسولات لاک مان

موجودات وقت جمعه 12 نم 2007 8:20 می. آپ کی آخری آمد سی یکی خطوط
کلی مراکز + UTC جو چطور کی جوانا کپسولات نیشان دادی کی کپسولات
This module contains the pre-installed package of the Urdu phpbb forum software and the category hierarchy mod. Using this package makes installation of category hierarchy mod much easier and introduces several other nice features.
Digits

• Should be able to read out Urdu digits
  – Issues related to normalization?
• Two strategies for English digits (to be defined by the website)
  – Read in English
  – Read in Urdu
Date Formats

• Cover other date formats in Urdu
  – A complete list can be provided
• Cover two calendars
  – Lunar based islamc calendar
  – Regular calendar
Diacritics

• Optionally, incompletely and (sometimes) incorrectly given

• اعلیً (almost), تقريباً (high (in quality or rank))

• بن یں بِن بن (ambiguous, knit (v), son of (particle), make (v))
Diacritics

• Lexical look-up
  – Should be ignored, and look up base form from the lexicon
  – MUST match given diacritics (even if incomplete)
    • Normally given where diacritic can cause ambiguity

• If lexical look-up fails, use a pronunciation guessing engine?
  • Use default mechanism/ No engine
  • Use an engine, at the given path: need to decide on
    – diacritics or pronunciation?
      » IPA, Custom Scheme, others?
    – Accuracy vs. performance?
      » POS tagger
      » WSD
Word Segmentation

• Urdu does not have the concept of SPACE
• SPACE used by typists to get the visual output
  – Spaces within words
    • بین و رسشی
  – No spaces between words, esp. after non-joiners, as words visually OK
    • آئندز رنر
• Should it point to word segmentation engine to enable correct output?
  – Use default mechanism/ No engine
  – Use an engine, at the given path: need to decide on
    • Accuracy vs. performance?
SSML for Urdu Speech Synthesis

Sarmad Hussain

Professor and Head
Center for Research in Urdu Language Processing
National University of Computer and Emerging Sciences, Lahore