Combilex RP

A high-quality Received English pronunciation lexicon for speech synthesis and speech recognition systems, which contains c.145,000 entries, including the 20,000 most frequent words.

lexicon for the Received Pronunciation accent of English
Combilex RP is a high-quality pronunciation lexicon for the British Received Pronunciation accent of English.

The lexicon contains c.145,000 entries, including the 20,000 most frequent words and contains a variety of linguistic information alongside detailed pronunciations.

Combilex RP is an ASCII text file, one entry-per-line, which is easily adaptable for use in text-to-speech synthesis (voice-building or run-time synthesis) and in speech recognition systems.

Full manually notated orthographic-phonemic correspondences are included, allowing derivation of accurate grapheme-to-phoneme rules.

The University of Edinburgh is seeking interest from commercial organisations to license this technology on a non-exclusive basis for internal product development purposes within all identified applications for speech synthesis.

Authors

Dr Ed Maughfling (Admin)

References

Richmond, Robust LTS rules with the Combilex speech technology lexicon, Interspeech 2009, Brighton

https://licensing.edinburgh-innovations.ed.ac.uk/product/combilex-rp