

## 3.6 Speech Synthesis

### Keywords

Speech synthesis; text-to-speech systems;  
Markov models; speech signal processing

### Key contact researcher(s)

Dr. Philip N. Garner  
[Phil.Garner@idiap.ch](mailto:Phil.Garner@idiap.ch)  
Tel.: +41 27 721 77 68

### Technology Transfer Office

Dr. Florent Monay  
Dr. Hugues Salamin  
[tto@idiap.ch](mailto:tto@idiap.ch)  
Tel.: +41 27 721 77 72

### Corporate Sponsorship Program

See Section 4 of the present document

### File reference & version number:

Software disclosure 5879  
Software disclosure 8700  
Software disclosure 11457

## Functional description

Idiap's speech synthesis technology is based on years of experience in speech recognition, and in cross fertilisation of the two technologies. We are able to provide specific voices in several languages, including English and the main Swiss languages. We can also provide average voice based TTS, that is, generic voices that can be adapted quickly to more personal voices. We are active in further merging the technologies for recognition and synthesis allowing advances in either domain to benefit the other.

## Innovative aspects

- Convergence of recognition and synthesis technology
- Cross-lingual adaptation

## Commercial application examples

- Speech to speech translation
- Hands free feedback
- Prosthetic voices
- Assistance for blind and partially sighted

## More information

“Combining Vocal Tract Length Normalization with Hierarchical Linear Transformations”, Lakshmi Saheer, Junichi Yamagishi, Philip N. Garner and John Dines, in: *IEEE Journal of Selected Topics in Signal Processing - Special Issue on Statistical Parametric Speech Synthesis*, 8(2):262–272, 2014.

## Software & IPR status

Most of the underlying software is open source. Some databases associated with certain languages may require separate agreements.

- VTLN: <http://www.idiap.ch/software/hts-vtln/>
- SSP: <https://github.com/idiap/ssp>
- HTS: <http://hts.sp.nitech.ac.jp/>
- Kaldi: <http://kaldi.sourceforge.net/about.html>