

Semester Thesis:

Effects of emphases on speech recognition

Background: In the Speech Processing Group we are developing an automated, personalised speech-to-speech translator. You say “I like cats” and it will say “Ich mag Katzen”. If you want to emphasise “cats” by saying “I like **CATS**”, it will say “Ich mag **KATZEN**”.

This automated translator consists of three modules: one for speech recognition (speech to text), one for translation (text to text) and one for speech synthesis (text to speech). Normally a speech recogniser is built on a huge number of recordings without such emphases. There has been little research on whether emphasised words are more difficult to recognise. As the first component of this automated translator, the speech recognition module needs to produce sufficiently accurate recognition results for the subsequent processing.



Task Description: We intend to investigate the effects of emphasised words on speech recognition results, though we do not anticipate substantially negative effects. Various languages, various spectral features, various types of emphases, etc will be involved in order to examine their correlation with the accuracy of speech recognition results.

Requirements: Programming skills and basic knowledge of speech processing would be of great help.

In case of interest, you may contact us by email and visit us afterwards. This web page provides you with more information about the entire Computer Engineering Group (TEC): <http://www.tec.ethz.ch/research.html>.

Contacts

- Hui LIANG (梁晖): liangh@tik.ee.ethz.ch, ETZ D97.4
- Beat Pfister: pfister@tik.ee.ethz.ch, ETZ D97.6