



HAL
open science

On the automatic comparison and cloning of native and non-native speech prosody.

Daniel J. Hirst

► **To cite this version:**

Daniel J. Hirst. On the automatic comparison and cloning of native and non-native speech prosody.. Speech Prosody 2016, May 2016, Boston (MA), United States. ISCA, pp.1038-1042, 10.21437/SpeechProsody.2016-213 . hal-04537626

HAL Id: hal-04537626

<https://hal.science/hal-04537626>

Submitted on 8 Apr 2024

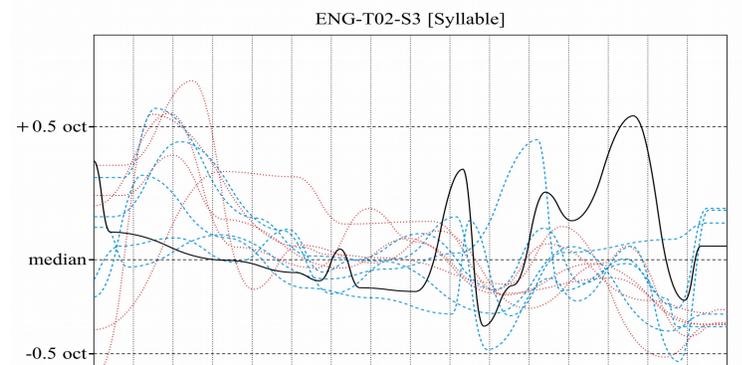
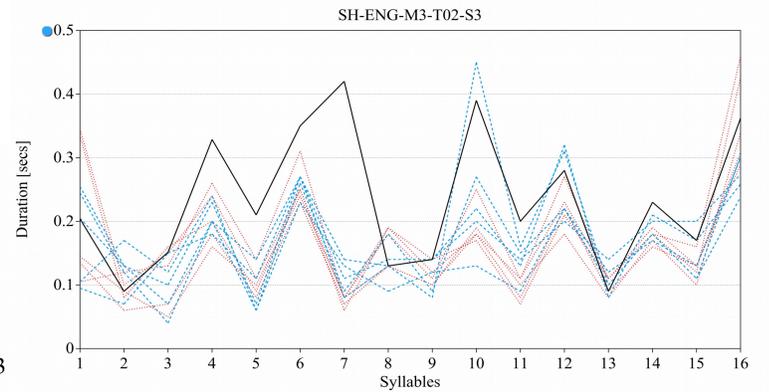
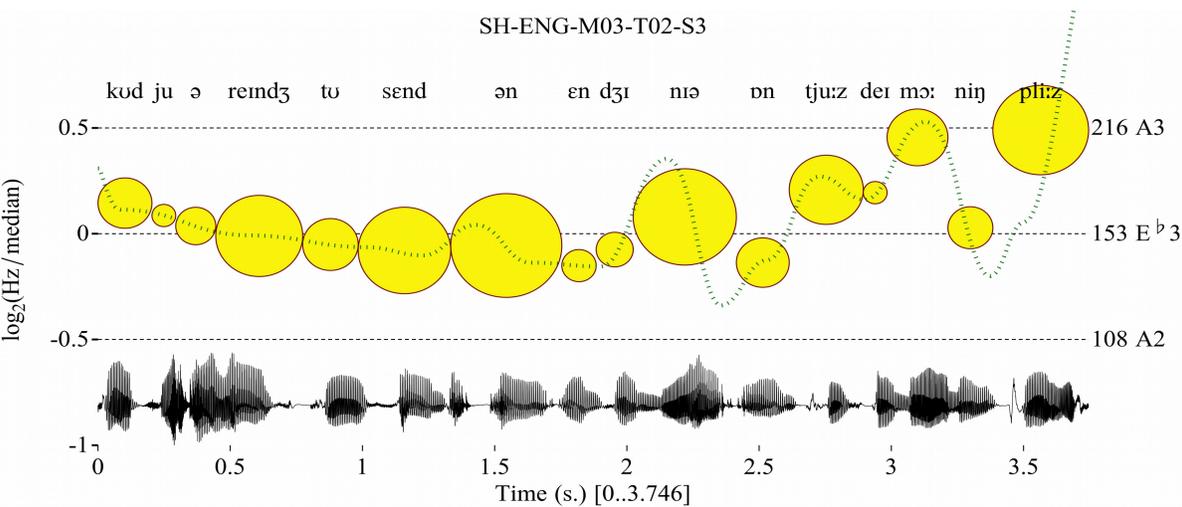
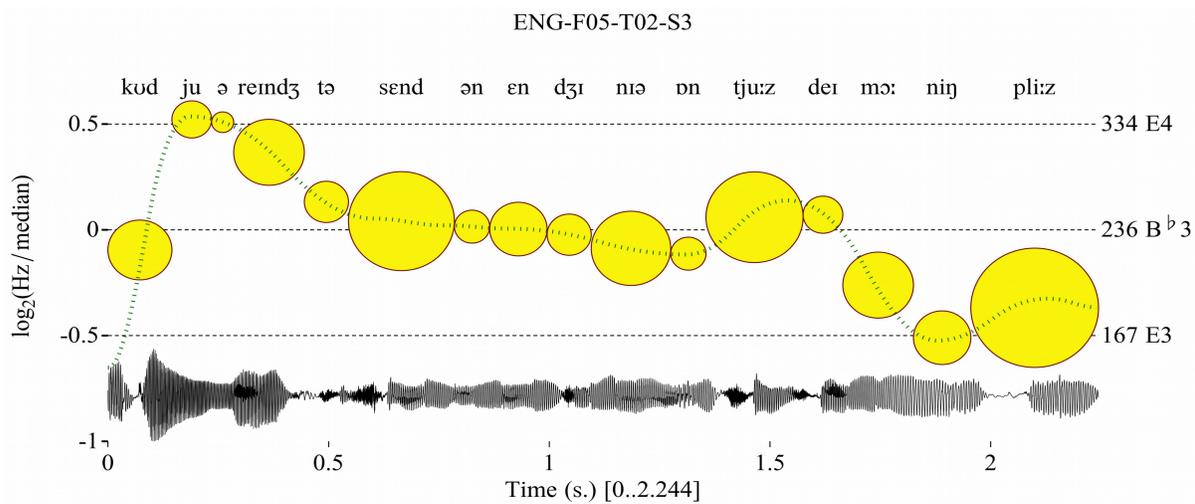
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

- Notoriously difficult to evaluate L2 prosody objectively - little consensus as to what constitutes a correct prosody for a given utterance.
- Automatic procedure: compare non-native speaker's production with 10 instances of the same utterance, taken from the OMProDat database, and read by native speakers.
- Pitch and relative syllable duration normalised and compared - version from native speaker most closely correlated with that of the non-native speaker chosen as a model.

•OMProDat Database

- 40 five sentence continuous passages read by 10 speakers
- English, French, German, Finnish, Korean, Mandarin, Cantonese
- L2 English-French, French-English, Mandarin-English,



- Normalised pitch and syllable durations of native speaker's recording cloned and transferred to the L2 utterance.
- The original and re-synthesised versions of the learner's utterance can then be used to provide both visual and auditory feedback to the language learner.

- All of the steps described here can now be carried out fully automatically with freely available tools.
- The tool is necessarily limited to read speech, because of the need for pre-recorded models. It seems likely, though, that this type of feedback could be extremely beneficial to language learners at all levels..

future work

- Apply the technique systematically to larger samples of L2 speech
- quantify the automatic evaluation of an L2 speaker's prosody and look at how well this automatic evaluation corresponds to the subjective evaluation by language teachers.
- Investigate the effectiveness of a tool such as this in providing visual and auditory feedback to a language learner.