Automatic classification of lexical stress errors for German CAPT

Anjana Vakil and Jürgen Trouvain

UNIVERSITÄT
DES
SAARLANDES

Department of Computational Linguistics and Phonetics
Saarland University, Saarbrücken, Germany

SLaTE 2015, Leipzig
4 September 2015

Outline

Lexical stress [TODO (LS)] in German

Accentuation/prominence of syllable(s) in a word

In German:

- Variable placement, contrastive function

  um·FAHR·en vs. UM·fahr·en
  to drive around *to run over*

- Reflected by duration, fundamental frequency (F0), intensity[1]
- Impacts intelligibility of non-native (L2) speech[2]

[1] G. Dogil and B. Williams. "The phonetic manifestation of word stress". In: *Word Prosodic Systems in the Languages of Europe*. Ed. by H. van der Hulst. Walter de Gruyter, 1999. Chap. 5, pp. 273–334.

[2] U. Hirschfeld. *Untersuchungen zur phonetischen Verständlichkeit Deutschlernender*. Vol. 57. Forum Phoneticum. 1994.

CAPT for lexical stress errors [TODO (LSEs)]

UNIVERSITÄT DES SAARLANDES

- Contrastive LS notoriously difficult for French speakers[1]
- CAPT offers huge potential for individualized instruction

- Automatic detection of LS errors in L2 German unexplored
- Recent work shows promising results using machine learning for classification of English stress patterns[2]

Our goal: classification-based detection of lexical stress errors by French learners of German

---

[1] E. Dupoux et al. "A Destressing 'Deafness' in French?" In: *Journal of Memory and Language* 36.3 (Apr. 1997), pp. 406–421.

[2] Y.-J. Kim and M. C. Beutnagel. "Automatic assessment of American English lexical stress using machine learning algorithms". In: *SLaTE*. 2011, pp. 93–96.

Data

IFCASL corpus of French-German speech[1]

- Phonetically diverse German sentences
- Read aloud by L1 and L2 (L1 French) speakers
  - Adults (>18) and children (15-16)
  - Levels[2] A2, B1, B2, C1 (children all A2/B1)
- Automatic (forced-alignment) segmentation at phone, syllable, and word level
- No lexical stress error annotation

[1] C. Fauth et al. "Designing a Bilingual Speech Corpus for French and German Language Learners: A Two-Step Process". In: *9th Language Resources and Evaluation Conference (LREC)*. Reykjavik, Iceland, 2014, pp. 1477–1482.

[2] Common European Framework of Reference, www.coe.int/lang-CEFR

Data annotation [TODO figure?]

Manual annotation of LSEs in subset of corpus:

- 12 bisyllabic, initial-stress words (word types)
- Word utterances (tokens) extracted automatically
- 668 tokens from ∼55 French speakers

Each token assigned a class label:

3 stress classes:	2 error classes:
[correct]	[bad_nsylls]
[incorrect]	[bad_audio]
[none]	

Data annotation

15 annotators (12 native), each token labeled by $\geq$2

Pairwise agreement calculated using:

- Percentage agreement: *N* agreed / *N* both annotated
- Cohen's Kappa (κ): accounts for chance agreement

	Mean	Maximum	Median	Minimum
% Agreement	54.92%	83.93%	55.36%	23.21%
Cohen's κ	0.23	0.61	0.26	-0.01

Single gold-standard label selected for each token

Data annotation results