German [u] vs [ʊ]: checking automatic labelling using various formant measurements

jane.wottawa@univ-lemans.fr, martine.adda-decker@sorbonne-nouvelle.fr

Introduction

- Production variation in German [u] and [ʊ]
- Differences in duration and/or timber [1, 7]
- Preliminary study
- Grapheme <u> in German: [u] or [ʊ] depending on context [2]
- Example: *zusammen* [tsuˈzɑmən] (DUDEN canonical) or [tsʊˈzɑmən]

Which vowel do speakers produce?

Research Questions and Hypotheses

- Can automatically aligned "canonical" [u] and [ʊ] segments be separated using F1/F2 acoustic measurements?
- Can back vowel formants be reliably detected automatically (using Praat)?
- To what extent do acoustic measurements help separate German [u] and [ʊ] in fluent speech?

H1 Vowel duration is not a strong cue to separate [u] and [ʊ] in spontaneous speech.

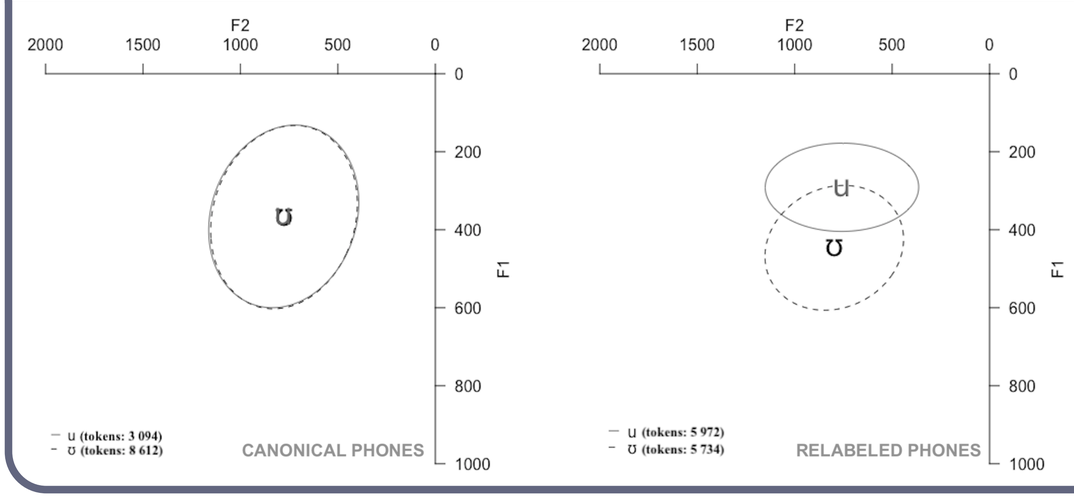
H2 F1 and F2 values are stronger cues than duration separating German [u] and [ʊ].

H3 After relabelling, automatic classification of [u] and [ʊ] is more successful.

Speech resource

- German broadcast ARTE TV documents
- late 1990s
- subset 35h of shows
- **Transcription:** manual, German orthography
- **Alignement:** automatic, German Limsi system [4]
  - rule-based dictionary
  - manual checking
  - DUDEN pron. dictionary as a reference

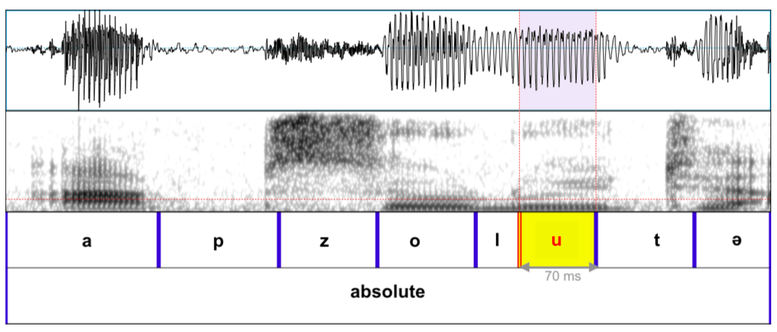
Classification of canonical and new labels (RWeka)



Classification using RWeka (11 706 tokens)

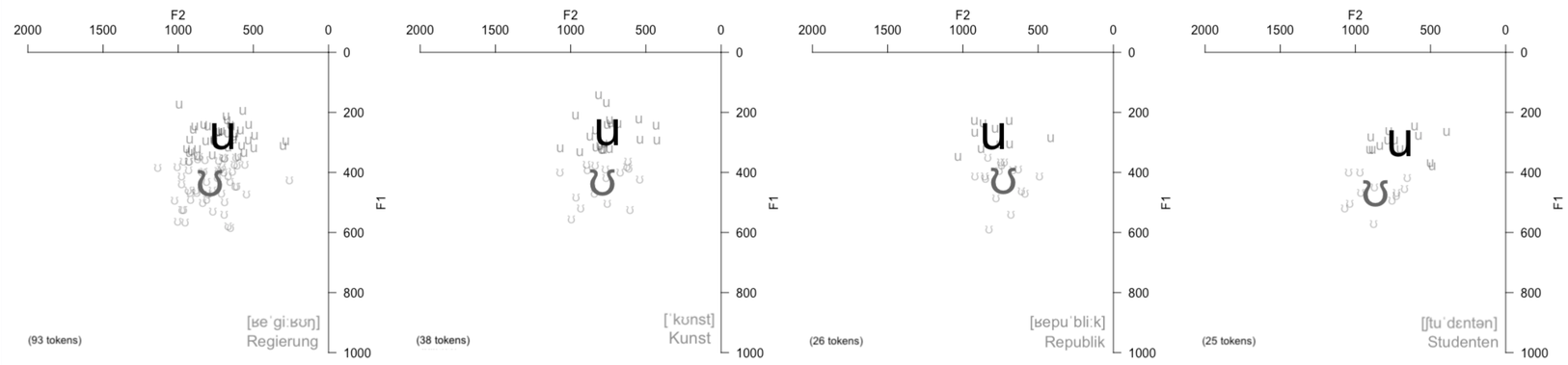
- **Canonical phone labels:**
  - Accuracy: 73.6%
  - Recall [u]: 0.0%
  - Recall [ʊ]: 100.0%
- **Relabeled phones:**
  - Accuracy: 95.4%
  - Recall [u]: 91.0%
  - Recall [ʊ]: 100.0%

Aligned spectrogram



- Example: *absolute* [apzoˈluːtə]
- Formant detection challenging because of small (F2-F1) distance.

Words with either [u] or [ʊ]



Dur. [u]: 55.4 ms	Dur. [u]: 68.1 ms	Dur. [u]: 65.0 ms	Dur. [u]: 55.4 ms
Dur. [ʊ]: 58.5 ms	Dur. [ʊ]: 75.3 ms	Dur. [ʊ]: 63.1 ms	Dur. [ʊ]: 60.8 ms

Protocol

- Extraction of F1 and F2 in a 0-1200Hz range of automatically aligned [u] and [ʊ]

Canonical Phone labels

- Comparison of the duration of [u] and [ʊ] → not conclusive ($M_{[u]}$=70.2 ms, $M_{[ʊ]}$=64.1 ms)
- Automatic [u] *vs.* [ʊ] classification using RWeka with F1, F2, (F2-F1) and duration → no relevant features found

Relabeling of Phones

- F1>4.0 Bark = label [ʊ] [5, 6]
- F1<3.5 Bark = label [u] [5, 6]
- 3.5 Bark<F1<4.0 Bark = label unchanged
- Comparison of the vowel duration → not conclusive ($M_{[u]}$=65.9 ms, $M_{[ʊ]}$=65.4 ms)
- Automatic classification using RWeka [3] → relevant feature: F1: [u]<346 Hz<[ʊ]
- Number of word types (3,324 in total)
  - without variants: 2,520 word types
  - with variants: 804 word types

  → function or frequent content words
  → coarticulation of [ʊ] with [n]

Discussion & Conclusions

✓ **H1** Duration is not a strong cue to separate [u] and [ʊ] in fluent speech.
➢ Spontaneous speech is more rapid, overall articulation rate might impact duration differences .

✓ **H2** Formant values are strong cues to separate [u] and [ʊ].
➢ Only F1 was used, a combination of F1 and F2 values might increase precision.

✓ **H3** The classification of the relabeled data predicts [u] and [ʊ] based on the F1 values.
➢ Neither F2 or F1-F2-distances seem to be strong cues in our data.

Can [u] and [ʊ] be considered free variants in German unstressed syllables?

- Validity of the analyses should be tested → Human perception test.

References

[1] C. Gendrot and M. Adda-Decker. Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. In *Interspeech*, 2005.

[2] T A. Hall. *Phonologie: Eine Einführung*. Walter de Gruyter, 2011.

[3] K. Hornik, C. Buchta, and A. Zeileis. Open-source machine learning: R meets Weka. *Computational Statistics*, 24(2):225–232, 2009.

[4] K. McTait and M. Adda-Decker. The 300k LIMSI German broadcast news transcription system. In *Interspeech*, 2003.

[5] W. Strange, O.-S. Bohn, K. Nishi, and S. A Trent. Contextual variation in the acoustic and perceptual similarity of North German and American English vowels. *JASA*, 118(3):1751–1762, 2005.

[6] W. Strange, O.-S. Bohn, S. A Trent, and K. Nishi. Acoustic and perceptual similarity of North German and American English vowels. *JASA*, 115(4):1791–1807, 2004.

[7] J. Wottawa and M. Adda-Decker. Quand les voyelles longues et brèves ne tiennent pas en place: la qualité vocalique en allemand L2. In *JEP 2018*, pages 64–71, 2018.

Le Mans Université · lium · Laboratoire de Phonétique et de Phonologie · Université Sorbonne Nouvelle Paris 3