Timing of word initial and word medial clusters in German

Jana Brunner, Christian Geng, Stavroula Sotiropoulou and Adamantios Gafos

Universität Potsdam, Department Linguistik

Phonetik und Phonologie 8, Jena 2012
Outline

1. Introduction
2. Methods
3. Results-Stability analysis
4. Methods-Computational modeling analysis
5. Results-Computational modeling analysis
6. Conclusion
Introduction

Complex onset languages
- Cluster consonants are timed with the vowel
  - English (Browman and Goldstein, 1988)
  - Georgian (Goldstein et al., 2007)
  - French (Kühnert et al., 2006)
  - Italian (Hermes et al., 2008)

Simplex onset languages
- Only prevocalic consonant is timed with the vowel
  - Berber (Hermes et al., 2011)
  - Moroccan Arabic (Shaw et al., 2009, Shaw et al., 2011)

Complex: C−Center stability
- C−Center stability
- LE stability
- RE stability

Simplex: RE stability
- C−Center stability
- LE stability
- RE stability
Introduction

Three analyses:

- Stability analysis (e.g. Browman and Goldstein, 1988)
- Lag analysis (e.g. Marin and Pouplier, 2010; Pouplier, 2012)
- Computational modeling analysis: how phonetic indices behave under change in various parameters such as variability and vowel duration (Shaw et al. 2009, 2011)
Introduction

Does German show properties of a complex onset language?
- stability analysis (Pouplier, 2012)
- identify parameters that vary with the stability indices
- computational modeling analysis
## Methods

### Corpus

<table>
<thead>
<tr>
<th>cluster</th>
<th>initial</th>
<th>medial</th>
<th>word boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>pl</td>
<td>plagen-lagen</td>
<td>geplagt-Gelage</td>
<td>knapp lagen</td>
</tr>
<tr>
<td></td>
<td>Plätze-Lätze</td>
<td>geplättet-verletzen</td>
<td>knapp Lätze</td>
</tr>
<tr>
<td></td>
<td>plauschen-lauschen</td>
<td>geplauscht-gelauscht</td>
<td>knapp lauschen</td>
</tr>
<tr>
<td>kn</td>
<td>knicken-nicken</td>
<td>geknickt-genickt</td>
<td>pack Nickel</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>gl</td>
<td>glauben-lauben</td>
<td>geglaubt-belaubt</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>kv</td>
<td>Quelle-Welle</td>
<td>gequellt-gewellt</td>
<td>pack Welle</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>sk</td>
<td>Sketche-Ketchup</td>
<td>gesketcht-gecatcht</td>
<td>lass Ketchup</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

- Accentuated: "Ich sah ’plagen’ an."
- Non-accentuated: "Ich sah mit Tom, nicht mit Anna ’plagen’ an."
Methods

- EMA (NDI Wave)
- 8 repetitions
- 7 speakers (3 analyzed)
Articulatory movements ("gestures") were labelled on the vertical velocity signal (20% threshold) using mtnew (P. Hoole)

- left edge to anchor interval
- C-center to anchor interval
- right edge to anchor interval

stability

- \( RSD = \frac{SD(\text{intervals})}{\text{MEAN}(\text{intervals})} \)
Results-Stability analysis

<table>
<thead>
<tr>
<th>cluster</th>
<th>initial</th>
<th></th>
<th>medial</th>
<th></th>
<th>word boundary</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RE</td>
<td>CC</td>
<td>RE</td>
<td>CC</td>
<td>RE</td>
<td>CC</td>
</tr>
<tr>
<td>gl</td>
<td>67</td>
<td>33</td>
<td>33</td>
<td>67</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>kn</td>
<td>50</td>
<td>50</td>
<td>33</td>
<td>67</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>kv</td>
<td>83</td>
<td>17</td>
<td>50</td>
<td>50</td>
<td>83</td>
<td>17</td>
</tr>
<tr>
<td>pl</td>
<td>94</td>
<td>6</td>
<td>67</td>
<td>33</td>
<td>78</td>
<td>22</td>
</tr>
<tr>
<td>sk</td>
<td>33</td>
<td>66</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>83</td>
</tr>
</tbody>
</table>

- results difficult to interpret
- initial: mostly RE stability (contra hypothesis of German as a complex onset language)
- medial: often CC stability
- word boundary: often CC stability (contra hypothesis)
- dependency on the cluster
Methods-Computational modeling analysis

Aim of this analysis
- study the full range of continuous manifestations of each syllable parse hypothesis as various parameters change
- fit these continuous manifestations to the experimental data
- find the syllable parse hypothesis with the better fit

Parameters that influence the stability indices and possibly hide an underlying CC stability
- cluster?
- word boundary?
- initial vs. medial clusters?
- vowel?
Methods-Computational modeling analysis

- CC-stability most likely for tense vowels
- CC-stability least likely for lax vowels
- Vowel has to be compressible for C-center stability to be low
Methods-Computational modeling analysis

Lag between C1 and C2 plateaus

- Clusters with small lags are very likely to have CC stability (/gl/, /sk/), clusters with large lags are not likely to have CC stability (/pl/, /kv/)
- Lag is larger after aspirated stops than after other sounds
- Smaller when these stops are followed by a word boundary
Methods - Computational modeling analysis

- Simulations
  - CCV and CV
  - Simplex onset parse
    - Vowel length
    - Lag duration
  - Complex onset parse

Surface production
  → eight CCV and eight CV productions → stability measures
Results-Computational modeling analysis

Simulations

- RSDs of CC-to-anchor interval affected by C1-C2-lag duration depending on lag duration both organizational patterns (simplex and complex onset) can give CC stability
- RSDs of RE-to-anchor interval not affected
Results-Computational modeling analysis

complex: small lags $\rightarrow$ CC stab.
complex: great lags $\rightarrow$ RE stab.
Results-Computational modeling analysis

- Complex: small lags $\rightarrow$ CC stab.
- Complex: great lags $\rightarrow$ RE stab.
- Simplex: very small lags $\rightarrow$ CC stab.
Vowel length effect

Simulations
- CC stability most likely for tense vowels, least likely for lax vowels
Simulations of word medial and word boundary clusters

**Input parameters**
- C1-C2-lag durations taken from the data
- Vowel duration (tense, lax, diphthong distinction)
- Simplex onset simulations and complex onset simulations

**Procedure**
- Per case 100 *simplex onset* and 100 *complex onset* simulations
- Percentage of cases with C-center stability calculated
- Difference between percentages in the simulations and percentages from the experimental data

**Hypothesis**
- For word medial clusters: \((data - simu_{complex}) < (data - simu_{simplex})\)
- For word boundary clusters: \((data - simu_{complex}) > (data - simu_{simplex})\)
Results-Computational modeling analysis

Squared differences between modelled and observed values are significantly smaller for

- the complex onset model for medial clusters (p=.042\textsuperscript{a})
- the simplex onset model for word boundary clusters (p=.028)

\textsuperscript{a}paired samples one tailed Wilcoxon test
Conclusion

- Stability analysis: results are difficult to interpret
- Computational modeling analysis: complex onset parse explains timing of medial cluster better than simplex onset parse; simplex onset parse explains timing of word boundary clusters better than complex onset parse
- C1-C2-lag duration has an important influence on the stability pattern
- Taking into account parameters of variation can uncover underlying organizational pattern
Bibliography


Acknowledgments

This work was carried out under the auspices of the Advanced ERC grant 249440.