Swedish-Danish word intelligibility

› Charlotte Gooskens (University of Groningen)
› Sebastian Kürschner (University of Groningen)

Sociolinguistics Symposium, Amsterdam 4 April 2008
Background

- Scandinavian languages (Danish, Norwegian, Swedish) are largely mutually intelligible
- Especially the Danish-Swedish communication presents some problems
- Previous investigations measured intelligibility of whole texts
- Results mainly explained by attitude and experience
- Little attention to linguistic phenomena
Present investigation

- Intelligibility of Swedish words by Danes
- Focus on the role of loan words versus inherited words
Swedish and Danish vocabularies

- Swedish and Danish share a large number of inherited words and loan words
- Loan words in Swedish from:
  - Low German (39%)
  - Latin (25%)
  - French (15%)
  - High German (14%)
  - Greek (5%)
  - English (2%)
  (from Gooskens et al in press)
Loan words in Swedish probably have more transparent phonetic correspondences with their counterparts in Danish than inherited words.

Inherited words:
Swedish *lag* [laːɡ] vs. Danish *lov* [lʌw] ‘law’

Loan words:
Swedish *linje* [liːnje] vs. Danish *linje* [linje] ‘line’
Hypotheses

1. Intelligibility is higher for loan words than for inherited words.

2. The phonetic distances between Danish and Swedish loan words are smaller than between inherited words.

3. There is an inverse relationship between intelligibility scores and phonetic distance.
Test words:

- 355 cognate nouns
- 29 non-cognate nouns
- randomly selected from a database with parallel lists of 2575 frequent spoken words
- concepts known to secondary school pupils between 15 and 18 years of age
- read aloud by a male speaker of standard Swedish
- phonetic transcriptions
Subjects:

› 29 native speakers of Danish
› pre-university education
› between 16 and 19 years
› not from border region
› little contact with Swedish (tested by means of non-cognates)
Task:

- translation of spoken Swedish test words into Danish
- via internet
Design:

- 4 word blocks of 96 words
- 3 to 17 subjects per block
Phonetic distance

Levenshtein-algorithm:

› Heeringa (2004)
› Measures the phonetic distances between related language varieties
› Compares the sounds of cognate word pairs
› Counts how many sounds must be substituted, added or removed in order to change the sounds of one word into the sounds of another word
**Example Levenshtein-distance:**

Danish *butik* vs. Swedish *butik* ‘shop’:

<table>
<thead>
<tr>
<th>Alignments</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swedish</td>
<td>b</td>
<td>u</td>
<td>t</td>
<td>i:</td>
<td>k</td>
</tr>
<tr>
<td>Danish</td>
<td>b</td>
<td>u</td>
<td>t</td>
<td>i</td>
<td>g</td>
</tr>
<tr>
<td>Costs</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[(1+1+1=3)/5 = 60\% \text{ difference}\]
Results

[Bar chart showing % correct answers and phonetic distance for loan words and inherited words]

- % correct answers
- Phonetic distance

_loan words inherited words_
Results

% correct  phonetic distance

- inherited (N=157)
- English (N=13)
- Latin (N=64)
- Low German (N=54)
- Frans (N=37)
- Greek (N=14)
- High German (N=9)
Results

- Correlation between % correctly translated words and phonetic distances is significant: $r = -0.36$, $p = 0.000$
Conclusions

1. It is easier for Danish listeners to understand Swedish cognate loan words than inherited words

2. The phonetic distances between Danish and Swedish loan words are smaller than between inherited words

3. Phonetically similar words are easier to recognize than phonetically dissimilar words
   BUT
   phonetic similarity only explains a small part of the variance in word intelligibility
Possible explaining factors

Communicatively relevant differences between sounds should be incorporated into the Levenshtein algorithm, for example:

- consonants versus vowels
- position of sound differences in words
- degree of differences between sounds
- word length
Possible explaining factors

Other explaining factors:

- knowledge of other languages
- word frequency
- orthography
- neighborhood density