Experiment on word intelligibility

Internet-based experiment including 7 test-languages
testing the intelligibility of isolated words

Research questions

- Do intelligibility scores differ for written vs. spoken words?
- Do intelligibility scores differ between cognates and non-cognates?
- Do intelligibility scores differ between inherited words and loan words (of different origin)?
- How does linguistic distance affect word intelligibility?
- Do specific sounds make it easier or more difficult to understand words?
- Do subjects make specific mistakes frequently?
- Do intelligibility scores differ with respect to subject groups, e.g. age-groups, geographic origin...
- etc.

Languages

- Dutch
- Frisian
- Standard High German
- Low German
- Danish
- Swedish
- Norwegian

Materials

- Database
  - 2575 words which are highly frequent in spoken language
    - Formal speech: Monologues in the European parliament
    - Informal speech: Dialogues from the Corpus Gesproken Nederlands (CGN, Corpus of Spoken Dutch)

Translated to all languages
Database

Annotation
- word class
- pronunciation (phonetic transcription)
- origin (inherited word or loan word)
- for loan words: direct origin, initial origin, and point of time when the word entered the language
- cognate or non-cognate

Test words

384 nouns
- randomly selected from the database
- recorded from male speakers of the 7 languages

Experiment

Subjects
- high school pupils
- 15-19 years
- (in principle) no prior knowledge of the test language (except for German)
- native speakers
- in total more than 1000 subjects
- http://www.let.rug.nl/lrs
- login: germanic
- password: guest

Design

<table>
<thead>
<tr>
<th>Test language</th>
<th>Dutch</th>
<th>Norwegean</th>
<th>Swedish</th>
<th>Dutch</th>
<th>Frisian</th>
<th>High German</th>
<th>Low German</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dutch</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Norwegian</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swedish</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frisian</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High German</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low German</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Reading task

Results by now

Extracting data from the LRS (Language Research System)

› Sebastian Kürschner, s.kurschner@rug.nl

LRS: Two interfaces

› Participant interface
  • URL: http://www.let.rug.nl/lrs/client/index/

› Administrative interface
  • http://www.let.rug.nl/lrs/admincp/login/
  • Login on your user’s manual

Exercise 1 and 2

› Exercise 1: Login and read the first page of the user’s manual. Try which information you can access when you click on the links in the left menu.

› Exercise 2: The page “Deviating translations” has been added later on and is therefore yet undocumented in the user’s manual. Open the site and select a combination of languages which you find interesting. Only encounter the first five columns. What information do you get? Which tasks would you likely need this information for?

Export: Exercise 3

› Exercise 3: Read the manual for the Export-site. Then try the “example querie” (under 3. in the manual).

ATTENTION: Do not use the filters “score_listen” and “score_read” at the moment, because an unresolved programming mistake has manipulated these numbers. For the moment, use the average values in the export-file instead.
Groups

When doing the following exercises, keep in mind that for some language combinations only a few subjects have participated yet. Since a lot of Danish subjects have participated, it is always a good choice to look for their results.

Build groups of two persons each for the following tasks.