INFORMATIONAL LOAD AS A TRIGGER FOR DISFLUENCIES IN INTERPRETING

Koen Plevoets
Intro

Interpreting: The rendition of utterances in another language
- Consecutive
- Simultaneous
- ...

‘Online’ translation
Intro

‘Interpreting is a cognitively demanding activity’
Intro

‘Interpreting is a cognitively demanding activity’

• Multitasking: Division of attention to different concurring tasks
• ‘Tightrope hypothesis’: Interpreters work at the limits of their processing capacities
  (Gile 1999)
Intro

Effort Model
(Gile 1985; 1997)

Interpreting = L + P + M + C

- L: Listening effort
- P: Production effort
- M: Memory effort
- C: Coordination effort
Figure 1: A model of the simultaneous interpretation process.
Moser (1978)
### Intro

Research into cognitive load in interpreting:

<table>
<thead>
<tr>
<th>Source text</th>
<th>Temporal characteristics</th>
<th>Textual characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Speech rate</td>
<td>Complexity</td>
</tr>
<tr>
<td>Target text</td>
<td>Ear-Voice Span</td>
<td>Disfluencies</td>
</tr>
</tbody>
</table>
Intro

Temporal characteristics:

• Speech rate: 120 words/minute is comfortable (Gerver 1969)

• Ear-Voice Span/“décalage”: 2-3 seconds on average (Treisman 1965; Anderson 1994)
Intro

Textual characteristics:

• Complexity: Cognitive load increases with
  • Lexical content (Gile 1995)
  • Numbers (Gile 1995)
  • Syntactic embedding (Dillinger 1994; Tommola & Helevä 1998)
• Disfluencies: e.g. silent/filled pauses: $uh(m)$,…
  (Tissi 2000; Mead 2002; Bakti 2009; Tóth 2011)
### Intro

**Setton (1999: 247)**

<table>
<thead>
<tr>
<th></th>
<th><strong>Attention to input</strong></th>
<th><strong>Attention to formulation</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Long silent pause</td>
<td>High</td>
<td>-</td>
</tr>
<tr>
<td>Short pausing</td>
<td>Normal listening</td>
<td>Routine planning</td>
</tr>
<tr>
<td>Filled pause</td>
<td>Normal listening</td>
<td>Routine planning</td>
</tr>
<tr>
<td>Mixed: Short &amp; filled pauses &amp; voice effects</td>
<td>Normal listening</td>
<td>Routine planning</td>
</tr>
<tr>
<td>Long filled pause</td>
<td>Relaxed or off</td>
<td>Planning/Searching</td>
</tr>
<tr>
<td>Fluent unmodulated string</td>
<td>Relaxed or off</td>
<td>Off</td>
</tr>
</tbody>
</table>
Intro

‘Disfluencies are a window on cognitive planning’

(Arnold et al. 2003; Bortfeld et al. 2001; Clark & Fox Tree 2002; Corley & Stewart 2008; Watanabe et al. 2008)
Research question

To what extent do disfluencies in interpreting depend on informational complexity?

- Lexical content
- Numbers
- Syntactic embedding
Research question

To what extent do disfluencies in interpreting depend on informational complexity?

• Lexical content
• Numbers
• Syntactic embedding
Data

European Parliament Interpreting Corpus – Ghent
Plenary sessions of the European Parliament
2006-2008
French, Spanish, Dutch, and English
190 000 tokens… and rising
Data

European Parliament Interpreting Corpus – Ghent

Transcribed according to VALIBEL-corpus (Bachy et al. 2007)  
POS-tagged and chunked by means of LeTs (Van de Kauter et al. 2013)
Data

Reference corpus: Spoken Dutch Corpus (Oostdijk 2000)
Component: Parliamentary debates
POS-tagged
10 million tokens
  • Flanders: 1/3
  • The Netherlands: 2/3
Data

Reference corpus: Spoken Dutch Corpus (Oostdijk 2000)
Component g: Parliamentary debates

360 000 tokens
• Flanders: 140 000
• The Netherlands: 220 000
## Data

<table>
<thead>
<tr>
<th></th>
<th>Nr. of files</th>
<th>Nr. of sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPICg – FRA (source)</td>
<td>108</td>
<td>1458</td>
</tr>
<tr>
<td>EPICg – DUT (target)</td>
<td>108</td>
<td>1437</td>
</tr>
<tr>
<td>SDCfl</td>
<td>155</td>
<td>8293</td>
</tr>
<tr>
<td>SPCnl</td>
<td>85</td>
<td>10753</td>
</tr>
</tbody>
</table>
Method

Per sentence:
- Nr. of *uh(m)*
- Nr. of content words
- Nr. of numerals
- Nr. of function words (= remainder)
- …
Method

Predict Nr. of $uh(m)$ on the basis of content words, numerals, function words AND ‘language’

i.e. non-interpreted Dutch, interpreted Dutch, and French source

Poisson regression
(Verified with Robust regression)
Analysis

1. At the level of the sentences
2. To measure the effect of the French source load on the Dutch interpretations:
   At the level of the files
Analysis

1. At the level of the sentences
2. To measure the effect of the French source load on the Dutch interpretations:
   At the level of the files
The graph shows the predicted number of UH(m) values across different numerals. The x-axis represents the number of numerals, ranging from 0 to 120. The y-axis represents the predicted number of UH(m), ranging from 0 to 35. Two sets of data are presented: one for lang: NED or and another for lang: NED in. The data is represented by solid and dashed lines.
Conclusion

Confirmation:

• More $uh(m)$ with interpreters than non-interpreters
• Lexical content has enhancing effect
• Numbers lead to high score of $uh(m)$ throughout

Negative effect of lexical content for non-interpreters: scripted nature of parliamentary speeches
Conclusion

Results demonstrate informational load:
Positive effect of grammatical material on the frequency of $uh(m)$ for non-interpreters
Absent for interpreters
HENCE: Non-interpreters produce more $uh(m)$ when they speak longer, interpreters when processing more content
Conclusion

Future prospects:
• Syntactic embedding
• Position of $uh(m)$ in utterance
Thank you!

koen.plevoets@ugent.be