Introduction

The theory

• Bybee (2002): articulatory reduction is observed in high-frequency words because high frequency of use makes the production of a word more automatic.
• Automatic behavior is less subject to conscious inhibition.
• Therefore, if high frequency of use automates word production, the production of high-frequency words should be harder to interrupt than the production of low-frequency words.

The dependent variable: Does the speaker interrupt the production of a word s/he intends to replace or repeat?

<table>
<thead>
<tr>
<th>Repetition repair</th>
<th>Interrupted:</th>
<th>It is a fi-, fixed female, by the way.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not interrupted</td>
<td>There was another, another amusement park down there.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Replacement repair</th>
<th>Interrupted</th>
<th>It was pathē-, I mean, it was horrible.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not interrupted</td>
<td>That's why we were surprised to see 'Toyota' written, I mean, imprinted on the engine.</td>
</tr>
</tbody>
</table>

The hypothesis: Words that are interrupted have a lower token frequency than words that are not interrupted.

e.g., frequency of fixed should be lower than frequency of another pathetic written

Data

Repair tokens from the Switchboard Corpus (Godfrey et al. 1992)
Checked for transcriber error

Repetition repairs:
Included only:
Single-word repetitions
more than 1 syllable, 4-8 segments
more than 1 segment pre-interruption
Excluded:
Personal names
Nonce formations
Words with a higher-frequency homonym
Very, 'really': uninterrupted repetition used for intensification

1018 total

Replacement repairs:
Included only:
Single-word replacements
Semantically related replacements
replaced word 3-8 segments in length

1485 total

Conclusion

• The production of a high-frequency word is less likely to be interrupted than the production of a low-frequency word even if the word is to be replaced
• This is consistent with Logan (1982) who found that typists, asked to stop typing a sentence when presented with a signal, stop less quickly if the signal is presented during the highest-frequency word the
• These results support the hypothesis that the production of high-frequency word is more automatic and therefore harder to interrupt than the production of low-frequency words (Bybee 2002)
• The frequencies of the replaced word and the replacement word show a positive correlation: semantically similar confusable words tend to have similar frequencies (Hotopf 1980)
• High-frequency words come to mind before low-frequency words: the replaced word is usually more frequent than the replacement word (not found previously, see Garrett 2001 for a review)

Results

Interrupted words tend to have lower frequency than uninterrupted words

<table>
<thead>
<tr>
<th>Repetition repairs</th>
<th>Interrupted word frequency</th>
<th>Uninterrupted word frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>N= 290</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>% interrupted: 26%</td>
<td>50%</td>
<td>60%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Replacement repairs</th>
<th>Interrupted word frequency</th>
<th>Uninterrupted word frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>N= 569</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>% interrupted: 28%</td>
<td>50%</td>
<td>70%</td>
</tr>
</tbody>
</table>

High frequency replacements could become available earlier, leading to interruption

But interrupted words actually tend to be replaced by words that are lower in token frequency than the words that replace uninterrupted words

If frequency of the replacement word mattered, we would expect significant differences in the other direction

Hence, it is the frequency of the pre-interruption word that matters for whether or not the words are interrupted

The relation between frequency of the replaced word and frequency of the replacement word

• Replacement more frequent than replaced words (p<.001)
• The frequency of replaced and the frequency of the replacement are positively correlated (p<.001)

References: