Morphemes in written word production

Please note that the results reported in this talk are outdated. Updated results have been reported at the annual meeting of the German Society for Linguistics (DGfS) in Munich, February 26, 2003 and are available at:


Guido Nottbusch, Angela Grimm & Ruediger Weingarten
University of Osnabrueck, Germany
Overview

• Introduction
  • Morphological composition
  • Lexical access of complex words: holistic or compositional?

• Method

• Results
  • Exp.1: Prefix, suffix and compound
  • Exp.2: whole word frequency or base frequency?

• Discussion
  • Lexical access: holistic or compositional?
Introduction

Subsumption in the framework of text production models

Writing single words involves the following sub-processes:

• Lexical access
• Graphemic encoding
• Orthographic encoding
• Graphomotoric execution
Introduction

Measuring the time course of writing can give insights into processes of word production after the initiation of writing.
Introduction

Morphological composition is evident in written language production as suggested by studies on written word production using healthy people (Will, et al. 2002\textsuperscript{a,b}) and on spelling errors of acquired dysgraphics (Badecker, et al., 1990, 1996).

- Lengthened interkey intervals at stem-morpheme boundaries depending on word frequency were found by Will, et al.
- Transposition errors over a stem-morpheme boundary almost never occur (e.g. Kindergarten as Kindergraten but not Kindegrarten).
Introduction

• Badecker, et al. found the rate at which inflectional word-final forms were preserved by their dysgraphic patient (e.g. surfed as sourphed but not as sourpht) to be clearly distinguishable from that of non-inflectional word-final forms (e.g. crypt as cript but not crippled).

• They therefore assume a lexical orthographic system in which morphologically complex forms have to be composed in production.
Introduction

Lexical access to complex words

- holistic access: independent lexical entry of the complex word
- composition: complex words are composed from their parts

In current models both routes are available and compete (e.g. Caramazza et al., 1988) or can converge on a single representation (Baayen & Schreuder, 1999).
Introduction

These results pose two questions:

1. Are stems and prefixes/suffixes represented separately in the lexicon? (Experiment 1)
2. Do within word frequency effects depend on whole word frequency or on the frequency of the base? (Experiment 2)
Method

Measured stimuli are matched for bigrams, i.e. only the intervals within a certain bigram occurring in different words are analysed.

| hindurch | hi n+d urch | [throughout] |
| Linde    | Li n+d e   | [lime tree]  |
| Kind     | Ki n+d    | [child]      |

This is done in order to control the following effects:
Method

- typing skill
- peripheral keys are struck slower
- alternating hands keystrokes are faster than keys struck with different fingers from the same hand
- According to Gentner (1983) a strong influence is exerted by the immediately preceding character
- graphotactic probability
- grapheme and bigram frequency
- ...
The bigram <nd> is present in all stimuli.
Method

The graph illustrates the time in milliseconds (ms) it took to write words corresponding to different types of stimuli: pictorial, printed, and acoustic. The x-axis represents the characters written, and the y-axis shows the time in ms. The graph shows the average time for each character for each type of stimulus.
Method (both experiments)

Participants (both experiments):

- 78 students of the University of Osnabrueck.
- All were native speakers of German.
- All were able to type fluently, although no strict criteria were applied (average writing speed: 44.2 words/min, std.dev.: 10.1).
- 63 female, 15 male.
- Mean age: 24.7 years, std.dev.: 3.8
- 72 students were right-handed, 6 left-handed.
Method *(both experiments)*

*Procedure (both experiments):*

- Stimuli appeared in a randomised fashion in the upper half of a 19” computer screen.
- Participants were instructed to read the stimulus and to type the word on the keyboard as fast as possible without errors.
- Simultaneously, with the typing of the first key of the target word, the stimulus disappeared from the screen, i.e. viewing times were self paced.
Method Experiment 1

• Are stems and prefixes/suffixes represented separately in the lexicon?

Stimuli:

Prefix+Prefix+Stem vs. Stem+Stem
vor+ent+halten [to deprive] vs. Meer+enge [strait]
Prefix+Stem vs. Stem+Stem (Control)
ver+edeln [to ennoble] vs. Rohr+ende [tube end]

Stem+Stem vs. Stem+Suffix
Hotel+koch [hotel cook] vs. Eitel+keit [vanity]

Note: All compared stimuli pairs contain the same bigrams.
Method Experiment 1

Stimuli:

- All items were derived exclusively from productive paradigms.
- All items were controlled for relative frequencies, i.e. the stem being more frequent than the whole term.
- All items were semantically transparent and unambiguous.
- Within bigram sets the number of syllables right from the relevant keys were matched.
## Results  
**Experiment 1**

### TABLE 1

<table>
<thead>
<tr>
<th>Morpheme Construction</th>
<th>Prefix + Prefix + Stem</th>
<th>Prefix + Stem</th>
<th>Stem + Stem</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>408 (174)</td>
<td>387 (188)</td>
<td>401 (180)</td>
</tr>
<tr>
<td></td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Morpheme Construction</th>
<th>Stem + Stem</th>
<th>Stem + Suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>421 (189)</td>
<td>366 (161)</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td></td>
</tr>
</tbody>
</table>

Note: standard deviations and number of cases are reported in parenthesis.
Results summary *Experiment 1*

- The slight differences between the initiation of prefixes and stems were nonsignificant.

- Experiment 1 showed a significant difference in timing at the start of stem morphemes and suffixes, the latter being initiated faster.
Discussion *Experiment 1*

Prefix + [Prefix+Stem]

- For prefixes a following unit is obligatory.
- The access to prefixes seems to be influenced by frame information of the following unit containing a stem.

[Stem] + [Stem]

[Stem + Suffix]

- A following unit after a suffix is only optional. (This can only be another suffix or a new frame containing a stem.)
- In the case of Stem+Suffix constructions the access to the suffix is faster - no further information is needed.
Method *Experiment 2*

In order to investigate the origin of the within word frequency effect (Will, et al.), two types of stimulus words were used:

**Stimuli:** whole word Freq. > base Freq.

vs. whole word Freq. < base Freq.

Prefix+Stem: ver+bessert [advanced] $\rightarrow$ wwF-item

vs. ver+blüht [withered] $\rightarrow$ bF-item

Stem+Stem: Welt+all [universe] $\rightarrow$ wwF-item

vs. Wort+art [part of speech] $\rightarrow$ bF-item

Note: All compared stimuli pairs contain the same bigrams.
### Results Experiment 2

#### TABLE 2

<table>
<thead>
<tr>
<th>frequency type</th>
<th>whole word frequency</th>
<th>base frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prefix+Stem</td>
<td>354 (166)</td>
<td>377 (168)</td>
</tr>
<tr>
<td>Stem+Stem</td>
<td>413 (177)</td>
<td>423 (188)</td>
</tr>
</tbody>
</table>

Note: standard deviations and number of cases are reported in parenthesis.
### Results Experiment 2

All frequency types consisted of high and low frequency items (split by 100), e.g.:

<table>
<thead>
<tr>
<th>type</th>
<th>level</th>
<th>word</th>
<th>wwf</th>
<th>(base) stem</th>
<th>baseF</th>
</tr>
</thead>
<tbody>
<tr>
<td>wwf</td>
<td>high</td>
<td>Jahr+hundert</td>
<td>484</td>
<td>Hundert</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[century]</td>
<td></td>
<td>[hundred]</td>
<td></td>
</tr>
<tr>
<td>wwf</td>
<td>low</td>
<td>Ver+hinderung</td>
<td>43</td>
<td>Hinderung</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[prevention]</td>
<td></td>
<td>[prevention]</td>
<td></td>
</tr>
<tr>
<td>bF</td>
<td>high</td>
<td>Haupt+person</td>
<td>0</td>
<td>Person</td>
<td>278</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[main person]</td>
<td></td>
<td>[person]</td>
<td></td>
</tr>
<tr>
<td>bF</td>
<td>low</td>
<td>Trost+pflaster</td>
<td>1</td>
<td>Pflaster</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[consolation]</td>
<td></td>
<td>[plaster]</td>
<td></td>
</tr>
</tbody>
</table>
## Results Experiment 2

### TABLE 3

Experiment 2: Mean IKIs in ms (complete dataset)

<table>
<thead>
<tr>
<th>freq. type</th>
<th>high frequency</th>
<th>low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>P+S wwF</td>
<td>329 (158)</td>
<td>343 (161)</td>
</tr>
<tr>
<td>P+S baseF</td>
<td>335 (162)</td>
<td>406 (198)</td>
</tr>
<tr>
<td>S+S wwF</td>
<td>413 (181)</td>
<td>413 (174)</td>
</tr>
<tr>
<td>S+S baseF</td>
<td>411 (182)</td>
<td>446 (196)</td>
</tr>
</tbody>
</table>

- **ns** indicates no significant difference
- **s** indicates significant difference
Discussion *Experiment 2*

- In Experiment 2 the overall comparison between higher whole word and base frequency did not lead to significant differences.

- The difference becomes clearer if the frequency levels of the items are taken into account.
Discussion

- Significantly increased latencies at morpheme onsets were found only in low frequency items with relatively higher base frequency.
- This is interpreted as an effect of compositional word production.
- There may also be compositional processes in high frequency items with higher base frequency.
- Possibly in these cases the compositional processes are too fast to be detected within the time course of word writing.
References


