Against Bracketing Erasure in English triconstituent compounds:

an investigation of acoustic constituent durations

Annika Schebesta – Heinrich-Heine Universität Düsseldorf
FOR 2373 workshop – project: EMB
PI’s: Gero Kunter, Ingo Plag
Today’s talk

[health\textsubscript{N1} care\textsubscript{N2}] law\textsubscript{N3} corner\textsubscript{N1} [drug\textsubscript{N2} store\textsubscript{N3}]

Does the morphological structure of compounds have an effect on the acoustic durations of N1, N2 and N3?
Lexical Phonology (Kiparsky 1982)

Assumption:

strict division of

- the application of morphological and phonological rules to a lexical item
Lexical Phonology (Kiparsky 1982)

Assumption:

strict division of

- the application of morphological and phonological rules to a lexical item \(ightarrow\) lexical level
Lexical Phonology (Kiparsky 1982)

Assumption:

strict division of

- the application of morphological and phonological rules to a
  lexical item \(\rightarrow\) lexical level

- the embedding into a syntactic structure and the phonetic
  implementation of a lexical item
Lexical Phonology (Kiparsky 1982)

Assumption:

strict division of

- the application of morphological and phonological rules to a lexical item $\rightarrow$ lexical level

- the embedding into a syntactic structure and the phonetic implementation of a lexical item $\rightarrow$ post-lexical level
Lexical Phonology (Kiparsky 1982)

Assumption:

strict division of

- the application of morphological and phonological rules to a lexical item

- the embedding into a syntactic structure and the phonetic implementation of a lexical item
Lexical Phonology (Kiparsky 1982)

Assumption:

strict division of

- the application of morphological and phonological rules to a lexical item

- the embedding into a syntactic structure and the phonetic implementation of a lexical item
Bracketing Erasure

After each application of a morphological rule, the internal morphological brackets in the complex word are erased.

→ morphological structure is not visible
→ phonetic signal has no access to the morphological structure of the complex word
→ factors related to the morphological structure do not affect the phonetic signal
Bracketing Erasure

$\text{health}_{N_1} \text{ care}_{N_2} \text{ law}_{N_3}$  $\text{corner}_{N_1} \text{ drug}_{N_2} \text{ store}_{N_3}$

$\rightarrow$ phonetic signal cannot reflect the morphological structure of the compound
$\rightarrow$ relations (e.g. embeddedness) between constituents should be undetectable
Contrary assumption

The morphological structure is encoded in the phonetic signal.

Sproat & Fujimura (1993): gradient variation of /l/ realizations according to the morphological boundary they attach to

Hay (2007): un- shorter in words with weaker boundaries (less decomposable), and longer in words with stronger boundaries (more decomposable)

Phonetic realization of segments at a morphological boundary is sensitive to the degree of boundary strength
Contrary assumption

The morphological structure is encoded in the phonetic signal.

Hay & Plag (2004): in suffixed words, inner boundaries are weaker than outer boundaries; suffixes with weaker boundaries are closer to the base

[aim-less]-ness  [king-dom]-ful

morphological embeddedness of affixes correlated with boundary strength
Contrary assumption

The morphological structure is encoded in the phonetic signal.

\[
\text{health}_{N1} \; \text{care}_{N2} \; \text{law}_{N3} \quad \quad \quad \text{corner}_{N1} \; \text{drug}_{N2} \; \text{store}_{N3}
\]

→ the phonetic implementation of the three constituents should be different due to the different boundary strengths
Contrary assumption

Kunter & Plag (2016) present the **Embedded Reduction Hypothesis**

In a complex word with more than two constituents, the embedded constituents are acoustically shorter than constituents at higher derivational levels.
Contrary assumption

**Embedded Reduction Hypothesis** tested with

a) experimental data  
b) corpus data
Kunter & Plag (2016)
Kunter & Plag (2016)

\[ \text{health}_{N1} \text{ care}_{N2} \text{ law}_{N3} \quad \text{corner}_{N1} [\text{drug}_{N2} \text{ store}_{N3}] \]

Predictions:
a. The embedded constituents are relatively short.
Kunter & Plag (2016)

\[
[\text{health}_{N1} \text{ care}_{N2}] \text{ law}_{N3} \quad \text{corner}_{N1} [\text{drug}_{N2} \text{ store}_{N3}]
\]

Predictions:

a. The embedded constituents are relatively short.
b. The free constituent is relatively long.
Kunter & Plag (2016)

\[ \text{health}_{N1} \text{ care}_{N2} \text{ law}_{N3} \quad \text{corner}_{N1} [\text{drug}_{N2} \text{ store}_{N3}] \]

Predictions:
a. The embedded constituents are relatively short.
b. The free constituent is relatively long.
c. This effect is independent from the branching direction.

→ interaction between constituents and branching direction of the compound needed
Kunter & Plag (2016)

\([\text{health} \text{care}] \text{law} \quad \text{corner} \quad \text{drug} \text{store}\]

-data set: experimental data (Kösling 2013, Kösling et al. 2013)
-477 English triconstituent NNN compounds

left = 239  right = 238
Kunter & Plag (2016)

- statistical analysis: lmer modelling

dependent variable: constituent duration

predictors:
- constituent number
- branching
- frequencies of each constituent
- bigram frequency N1N2 and bigram frequency N2N3
- trigram frequency N1N2N3
- accent
- pitch range
- phonological length

random effect: speaker

central interactions:
- constituent number * branching * bigramFreqN1N2
- constituent number * branching * bigramFreqN2N3
Shortcomings

- across-boundary frequencies kept low

- across-boundary frequencies may not be informative

→ therefore: focus on embedded constituent frequencies

   N1N2 for left-branching compounds
   N2N3 for right-branching compounds
left-branching: [N1 N2] N3

N1N2 bigram frequency (=embedded constituent):

N1 is relatively short regardless of N1N2 freq.
N2 is relatively short regardless of N1N2 freq.
N3 is relatively long regardless of N1N2 freq.
left-branching: \([N1 \ N2] \ N3\)

\(N1N2\) bigram frequency (\(=\)embedded constituent):

N1 is relatively short regardless of \(N1N2\) freq.  
N2 is relatively short regardless of \(N1N2\) freq.  
N3 is relatively long regardless of \(N1N2\) freq.  

EXPECTED
right-branching: N1 [N2 N3]

N2N3 bigram frequency (=embedded constituent):

N1 is relatively long with higher N2N3 freq.
N2 is relatively short with higher N2N3 freq.
N3 is relatively short regardless of N2N3 freq.
right-branching: N1 [N2 N3]

N2N3 bigram frequency (=embedded constituent):
N1 is relatively long with higher N2N3 freq.  
N2 is relatively short with higher N2N3 freq.  
N3 is relatively short regardless of N2N3 freq.  

EXPECTED
## Results

<table>
<thead>
<tr>
<th>embedded constituent frequency</th>
<th>Kunter &amp; Plag (2016)</th>
<th>implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>left-branching (N1N2)</td>
<td>N1 short N2 short N3 long</td>
<td>challenges Bracketing Erasure supports ERH predictions</td>
</tr>
<tr>
<td>right-branching (N2N3)</td>
<td>N1 long N2 short N3 short</td>
<td>challenges Bracketing Erasure supports ERH predictions</td>
</tr>
</tbody>
</table>
Corpus Study
Corpus Study

\[ \text{[health}_{N1} \text{ care}_{N2}] \text{ law}_{N3} \]

\[ \text{corner}_{N1} \text{ [drug}_{N2} \text{ store}_{N3}] \]

- Data from BURSC (data set by Köslin & Plag 2009)
- 451 English triconstituent NNN compounds

left = 331  right = 120
Corpus study

[health$_{N1}$ care$_{N2}$] law$_{N3}$

corner$_{N1}$ [drug$_{N2}$ store$_{N3}$]

Predictions:
a. The embedded constituents are relatively short.
b. The free constituent is relatively long.
c. This effect is independent from the branching direction.
Corpus study

- statistical analysis: lmer modelling

dependent variable: constituent duration

predictors:
- constituent number
- branching frequencies of each constituent
- bigram frequency N1N2 and bigram frequency N2N3
- trigram frequency N1N2N3
- pitch range
- phonological length

random effect: speaker

central interactions:
- constituent number * branching * bigramFreqN1N2
- constituent number * branching * bigramFreqN2N3
left-branching: [N1 N2] N3

N1N2 bigram frequency (=embedded constituent):

N1 is relatively short
N2 is relatively short
N3 is relatively long
left-branching: [N1 N2] N3

N1N2 bigram frequency (=embedded constituent):

N1 is relatively short  EXPECTED
N2 is relatively short  EXPECTED
N3 is relatively long   EXPECTED
right-branching: N1 [N2 N3]

N2N3 bigram frequency (=embedded constituent):

N1 is longer than N2, but shorter than N3
N2 is relatively short
N3 is relatively long
right-branching: N1 [N2 N3]

N2N3 bigram frequency (=embedded constituent):

N1 is longer than N2, but shorter than N3
N2 is relatively short
N3 is relatively long

difference to Kunter & Plag (2016) analysis:

3-way interaction constituent number * branching * bigramFreq N2N3 not significant

→ effect of N2N3 similar for left-branching
left-branching & right-branching
left-branching & right-branching

low bigramFreq N1N2

average bigramFreq N1N2
left-branching & right-branching

low bigramFreq N1N2
average bigramFreq N1N2
high bigramFreq N1N2
left-branching & right-branching

low bigramFreq N1N2

average bigramFreq N1N2

high bigramFreq N1N2
left-branching & right-branching

low bigramFreq N1N2

average bigramFreq N1N2

high bigramFreq N1N2
left-branching & right-branching

low bigramFreq N1N2

average bigramFreq N1N2

high bigramFreq N1N2
## Results

<table>
<thead>
<tr>
<th>Embedded constituent frequency</th>
<th>Corpus study</th>
<th>Implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>left-branching (N1N2)</td>
<td>N1 short N2 short N3 long</td>
<td>challenges Bracketing Erasure supports ERH predictions</td>
</tr>
<tr>
<td>right-branching (N2N3)</td>
<td>N1 short N2 short N3 long</td>
<td>challenges Bracketing Erasure no support ERH predictions</td>
</tr>
</tbody>
</table>
Kunter & Plag (2016) & corpus study: comparisons
## Kunter & Plag (2016) & corpus study

<table>
<thead>
<tr>
<th>Embedded Constituent Frequency</th>
<th>Kunter &amp; Plag (2016)</th>
<th>Corpus Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>left-branching (N1N2)</strong></td>
<td>N1 short N2 short N3 long</td>
<td>N1 short N2 short N3 long</td>
</tr>
<tr>
<td><strong>right-branching (N2N3)</strong></td>
<td>N1 long N2 short N3 short</td>
<td>N1 short N2 short N3 long</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Across-Boundary Frequency</th>
<th>Kunter &amp; Plag (2016)</th>
<th>Corpus Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>left-branching (N2N3)</strong></td>
<td>N1 short N2 long N3 short</td>
<td>N1 short N2 short N3 long</td>
</tr>
<tr>
<td><strong>right-branching (N1N2)</strong></td>
<td>N1 short N2 short N3 long</td>
<td>N1 short N2 short N3 long</td>
</tr>
</tbody>
</table>
Contrasting ERH and Bracketing Erasure

Bracketing Erasure claims

- no difference in acoustic durations among all constituents of a complex word
- no effect of branching direction on the acoustic duration of constituents

Embedded Reduction Hypothesis claims

- differences in acoustic durations among constituents of a complex word
- shorter durations with embedded constituents, longer durations with free constituents
## Contrasting ERH and Bracketing Erasure

<table>
<thead>
<tr>
<th></th>
<th>Kunter &amp; Plag (2016)</th>
<th>corpus study</th>
<th>implications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>embedded constituent frequency</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>left-branching (N1N2)</td>
<td>N1 short N2 short N3 long</td>
<td>N1 short N2 short N3 long</td>
<td>challenges Bracketing Erasure supports ERH predictions</td>
</tr>
<tr>
<td>right-branching (N2N3)</td>
<td>N1 long N2 short N3 short</td>
<td>N1 short N2 short N3 long</td>
<td>challenges Bracketing Erasure partly supports ERH predictions</td>
</tr>
<tr>
<td><strong>across-boundary frequency</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>left-branching (N2N3)</td>
<td>N1 short N2 long N3 short</td>
<td>N1 short N2 short N3 long</td>
<td>challenges Bracketing Erasure partly support ERH predictions</td>
</tr>
<tr>
<td>right-branching (N1N2)</td>
<td>N1 short N2 short N3 long</td>
<td>N1 short N2 short N3 long</td>
<td>challenges Bracketing Erasure no support for ERH predictions</td>
</tr>
</tbody>
</table>
Contrasting ERH and Bracketing Erasure

Bracketing Erasure

cannot explain the effects found in both studies.

Embedded Reduction Hypothesis

cannot explain all the effects, either.

More research needed:

experimental data, controlled for n-gram frequencies, not only duration but also phonetic reduction taken into account, ...
References


Contrasting ERH and Bracketing Erasure

<table>
<thead>
<tr>
<th>embedded constituent frequency</th>
<th>Kunter &amp; Plag (2016)</th>
<th>corpus study</th>
<th>implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>left-branching (N1N2)</td>
<td>N1 short N2 short N3 long</td>
<td>N1 short N2 short N3 long</td>
<td>challenges Bracketing Erasure supports ERH predictions</td>
</tr>
<tr>
<td>right-branching (N2N3)</td>
<td>N1 long N2 short N3 short</td>
<td>N1 short N2 short N3 long</td>
<td>challenges Bracketing Erasure partly supports ERH predictions</td>
</tr>
<tr>
<td>across-boundary frequency</td>
<td>left-branching (N2N3)</td>
<td>N1 short N2 long N3 short</td>
<td>N1 short N2 short N3 long</td>
</tr>
<tr>
<td>right-branching (N1N2)</td>
<td>N1 short N2 short N3 long</td>
<td>N1 short N2 short N3 long</td>
<td>challenges Bracketing Erasure no support for ERH predictions</td>
</tr>
</tbody>
</table>