A Self-Paced Reading Study on Processing Constructions with Different Degrees of Compositionality

Giulia Rambelli, Alessandro Lenci, Philippe Blache

To cite this version:

Giulia Rambelli, Alessandro Lenci, Philippe Blache. A Self-Paced Reading Study on Processing Constructions with Different Degrees of Compositionality. The 35th Annual Conference on Human Sentence Processing, Mar 2022, UC Santa Cruz, United States. hal-03620795

HAL Id: hal-03620795
https://hal.science/hal-03620795
Submitted on 26 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A Self-Paced Reading Study on Processing Constructions with Different Degrees of Compositionality

Giulia Rambelli, Alessandro Lenci, Philippe Blache
giulia.rambelli@psy.unip.it

Introduction

Many theoretical questions are still debated in the study of language comprehension: How do speakers process compositional sequences? Which factors can facilitate language comprehension? While the principle of compositionality is traditionally considered as the primary mean of explaining language processing, behavioral data support the idea that it is just a default option within a more complex scenario, where a series of noncompositional mechanisms can be used in processing. analogies with stored exemplars, shallow processing, activation of a network of mutual expectations, etc.

Experimental Hypothesis

The experiment is designed to examine reading times (RTs) of verb-noun constructions with a different degree of compositionality. We compare 3 conditions:

1. Idiomatic expressions (ID) such as: 'bury the hatchet'
2. Compositional and high-frequency expressions (HF) such as: 'bury the treasure'
3. Compositional and low-frequency expressions (LF) such as: 'bury the machete'

The novelty of this work is twofold:

1. we do not deal with fixed multi-word sequences as lexical bundles but specifically with processes on a group of expressions.
2. we compare in the same experiment both idiomatic and high frequent expressions, longer for infrequent expressions than for frequent ones.

Hypothesis RTs are longer for compositional expressions than for idiomatic ones, and RTs are facilitated when reading idiomatic expressions instead of compositional ones.

Method

Participants 90 L1 English speakers from the US and Canada, aged 18 to 50 (M±29.6 ± 7.55). The experiment was delivered remotely, and participants were recruited using Prolific.

Material 48 VERB-determinant=NOUN idioms and corresponding high and low-frequency bigrams, objects matched by frequency and character length. Constraint: V-N association score for HF 2(hex)-N association score for ID.

Method Each stimulus consisted of a context sentence presented for the participant to read in one instance and a sentence with the target phrase embedded, displayed word-by-word using the moving-window SPR paradigm.

Results

Analysis reveals no difference between processing the figurative meaning of idioms and the compositional one of HF; there are facilitation effects in the comprehension of both expressions. Two plausible explanations:

1. both idioms and HF expressions are stored as unanalyzed wholes and directly retrieved once recognized, following usage-based perspective;
2. processing HF relies on a co-activated network of representations operating with analogy-based mechanisms leading to sentence meaning construction; facilitation effects are similar for ID and HF but depend on different mechanisms.

The present experiment offers many exciting avenues for conducting additional exploratory and targeted analyses. Although the results of this study require further interrogation, we wish these findings contributed to the existing research in compositionality.

Acknowledgements

Research supported by grants ANR-16-CVen-0002-1390 and the Excellence Initiative of Aix-Marseille University (A*MIDEX).

References


Data Analysis

We analyzed the cleaned RTs of the phrase-final words using a linear mixed-effects model: $RT_{ij} \sim Condition \times Age \times WordLength \times VerbFreq_{ij} \times PosInList_{i}$

$\text{Table 2. Fixed effects for final model.}$

<table>
<thead>
<tr>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (ID)</td>
<td>3.5699</td>
<td>0.0074</td>
<td>47</td>
<td>486.93</td>
</tr>
<tr>
<td>HF</td>
<td>0.0022</td>
<td>0.0126</td>
<td>38.38</td>
<td>0.0173</td>
</tr>
<tr>
<td>LF</td>
<td>0.0312</td>
<td>0.0037</td>
<td>94.89</td>
<td>0.0535</td>
</tr>
<tr>
<td>PosInList</td>
<td>0.0004</td>
<td>0.0001</td>
<td>3891</td>
<td>0.0000</td>
</tr>
<tr>
<td>WordLength</td>
<td>0.0177</td>
<td>0.0004</td>
<td>134.81</td>
<td>0.0000</td>
</tr>
<tr>
<td>VerbFreq</td>
<td>0.0099</td>
<td>0.0037</td>
<td>137.99</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 1. Example of stimuli.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Critical region</th>
<th>Non-critical region</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Finn changed his life.</td>
<td>All of a sudden he kicked the ball into the net and won the match.</td>
</tr>
<tr>
<td>HF</td>
<td>All of a sudden he kicked the ball into the net and won the match.</td>
<td>Finn changed his life.</td>
</tr>
<tr>
<td>LF</td>
<td>His brother kicked the sister of his best friend in the head.</td>
<td>All of a sudden he kicked the ball into the net and won the match.</td>
</tr>
</tbody>
</table>

Discussion

How do speakers process compositional sequences? Which factors can facilitate language comprehension? While the principle of compositionality is traditionally considered as the primary mean of explaining language processing, behavioral data support the idea that it is just a default option within a more complex scenario, where a series of noncompositional mechanisms can be used in processing. analogies with stored exemplars, shallow processing, activation of a network of mutual expectations, etc.

While the speakers process compositional sequences as lexical bundles but specifically with processes on a group of expressions.

Hypothesis RTs are longer for compositional expressions than for idiomatic ones, and RTs are facilitated when reading idiomatic expressions instead of compositional ones.

The novelty of this work is twofold:

1. we do not deal with fixed multi-word sequences as lexical bundles but specifically with processes on a group of expressions.
2. we compare in the same experiment both idiomatic and high frequent expressions, longer for infrequent expressions than for frequent ones.

Hypothesis RTs are longer for compositional expressions than for idiomatic ones, and RTs are facilitated when reading idiomatic expressions instead of compositional ones.

The advantage for infrequent phrases was relatively small: maybe context sentences reduce the advantage for infrequent phrases was relatively small: maybe context sentences reduce the advantage for infrequent phrases.