1 | Introduction
Previous studies have related the differences between corpus frequency data and judgment data to different processing modules [3]:
• Judgments represent the cognitive workload for structure processing
  - The data structure reveals finer-grained distribution of the relevant factors
• Frequency reflects output selection for linguistic production
  - The data structure reveals strong preferences of few particular over other competitive structures

In our study:
1 We compare different data types resulting from different methods...
2 Along with the phenomenon of word order freedom in German accusative experiencer structures (ACCEXP)

Here:
• We obtained relative instead of absolute judgments.
• We tested intonation-based frequency arising from choices out of manipulated sets of alternatives instead of corpus frequencies [2], [3], [6].

2 | Hypotheses

Method-related predictions:
• Judgment vs. frequency:
  • The relative distributions of the two measures match per condition, but they differ in strength
  • Choice frequency vs. corpus frequency:
  • The relative frequencies reflect gradient well-formedness rather than production (cf. [3]):
    - Phenomenon under investigation: German object-experiencer verbs are known to license object-before-subject order (O+S).
    - The factorial design contains three binary variables:
      • CONTEXT: licensing context (triggering of object fronting) vs. non-licensing context (no triggering of object fronting)
      • VERBTYPE: object-experiencer vs. canonical verbs
      • WordOrder: O+S vs. O+S

Condition-related predictions:
• The factors CONTEXT and VERBTYPE have significant impact on WordOrder.

3 | Data

Judgment data:
• Obtained via split-100 rating [2]; subjects rated the S+O and O+S version of the same utterance corresponding to one of the four permutations of CONTEXT*VERBTYPE. Out of a score of 100 subjects award points to both alternatives (e.g. 50/50, 0/100, 80/20...).
• Thus, all structural alternatives were explicitly evaluated by the subjects → gives a relative judgment of the well-formedness of the alternatives

Frequency data:
• Obtained via two-alternatives forced choice study [1], [4], [7]; subjects chose between the S+O and O+S version of a contextually embedded sentence.
• Thus, not all structural alternatives are explicitly evaluated → it is possible that subjects reject a well-formed structure or that they decide for a non-well-formed

4 | Results

<table>
<thead>
<tr>
<th>7-Point Rating (check study)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Non-relative judgment of context-target pair</td>
</tr>
<tr>
<td>• slight rise of acceptability</td>
</tr>
<tr>
<td>• no significant effects of verb type or context</td>
</tr>
<tr>
<td>• scalar endpoints match (no random distribution)</td>
</tr>
<tr>
<td>• no general alignment with the distribution of the relative data</td>
</tr>
</tbody>
</table>

**Forced choice task (frequency study)**
- Choice out of two minimal different context-target pairs
- Rise of choice probability
- Significant effects of verb type and context
- The most acceptable structure is the most frequent, the less acceptable structure is the least frequent

<table>
<thead>
<tr>
<th>Fitted effects</th>
<th>Estimate</th>
<th>Standard error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.21</td>
<td>0.285</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>VerbType</td>
<td>0.097</td>
<td>0.229</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Random Subject</td>
<td>0.191</td>
<td>0.219</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Error</td>
<td>0.207</td>
<td>0.713</td>
<td></td>
</tr>
</tbody>
</table>

**Split-100 task (rating study)**
- Relative judgment of two minimal different context-target pairs
  - Rise of acceptability
  - Significant but lower effects
  - Low frequent structures are not minimum acceptable.
  - High frequent structures are not maximum acceptable.

**Stability across data types?** → No stability of effects with 7-point judgments

5 | Conclusion & Discussion

Method-related conclusion:
• The relative distribution of the data points is the same for both data types but for some conditions the effect size is larger for frequency than for judgments.
• There is an alignment of effects on well-formedness and choice probabilities.
• This reinforces the ranking of the effects of the underlying processes.
• The differences in strength cannot be associated with non-controlled factors, but rather with the underlying process.

Choice frequency vs. corpus frequency:
• Corpus frequencies typically reveal very strong preference for one and zero for other structures. We find a high occurrence probability for more than one condition and no condition with highest or zero frequency.
• Choice vs. corpus: conscious output selection process out of a limited number of overt competitors vs. unconscious output selection out of an unlimited number of non-overt competitors.
• Choice & corpus: not all structural alternatives are explicitly evaluated (vs. scalar judgment & categorical yes/no-frequencies).

**Forced choice frequencies share properties of both levels:** They reflect gradient well-formedness as well as strong preferences of an output selection process. Contrary to the above hypotheses, they uncover both cognitive workload and production-related choices.

Content-related conclusion:
• The manipulated context licenses object fronting & ACCEXP verbs license O+S-order without the contextual licensing.

References | Footnotes