Do discourse connectives influence likelihood and form of reference in the absence of verb semantics?

Jennifer Spenader
Artificial Intelligence
University of Groningen
j.spenader@ai.rug.nl

Abstract

We tested the influence of connectives on the grammatical role of continuations (subject or object) and the referential form of continuations (name or pronoun). We factor out verb semantics by using nonsense transitive verbs followed one of three connectives, because, so and but. We found an unexpected general tendency for object continuations. Object continuations were significantly more likely with but and so compared to because. Because was also significantly more likely to be followed by a pronoun than a name compared to but. We also found evidence that grammatical role expectations do not interact with referential form choice.

Likelihood of reference: connectives without verb semantics

Verb semantics influences the likelihood of subsequent reference to verbal arguments. Consider the verb like (1).

(1) a. John<experiencer> likes Bill<stimulus> because he ...
(preferred continuation: he = Bill)

b. John<experiencer> likes Bill<stimulus> so he (preferred continuation: he = John)

The subject of like has the semantic role of Experiencer, while the object is the Stimulus. Like and similar verbs belong to the class of Experience-Stimulus verbs (ES). In continuation experiments, where participants are asked to write continuations of sentences/clauses after certain verbs followed by a connective and (often) an ambiguous pronoun, as in (1), participants show a strong and robust tendency to continue with a reference to the Stimulus argument: in (1-a), he = Bill. But connectives also influence continuation likelihood. Stimulus continuations are preferred after because but Experiencer continuations are preferred after so, see (1-b).

ES verbs and Stimulus-Experiencer verbs (SE) like amaze, that have Experiencer objects and Stimulus subjects are all State verbs. Stevenson et al. (1994) and Fukumura and Von Gompel (2010) found semantic role continuation preferences hold for State verbs with consistent effects for connectives, regardless of the grammatical role of the Experiencer or Stimulus. However, not all verb groups that show consistent continuation biases find an interaction with connectives. Stevenson et al. (1994) found connective choice did not influence Transfer verbs(e.g. give) or Action verbs (e.g. hit). Transfer verbs are consistently biased to Goal/object continuations, and Action verbs to Patient/object continuations. An explanation may be that connective-verb semantic interactions only occur when connective meaning relates to verb meaning. For State verbs the Stimulus is the cause, so it is not surprising that after because most continuations refer to this semantic role. The opposite is the case for so (used with Cause-so-Effect sequences). The connection between Agent/Patient roles or Source/Goal roles and causal connectives may be more indirect with Action and Transfer verbs.

Nonsense verbs can be used to study the contribution of connectives separate from preferences stemming from verb biases. Given this background, our first research question is:

- Do discourse connectives influence the likelihood of continuing a discourse with a subject or an object in the absence of verb semantics?

At issue is also a broader research question: Do expectations interact with referential form choice? For verb semantic biases previous work disagrees as to whether or not continuation likelihood influences referent accessibility. Recall that transfer verbs are biased to Goal continuations. In a continuation experiment with Transfer verbs without connectives, Arnold (2001) found that speakers used pronouns more than proper nouns to refer to Goal continuations, suggesting that a strong likelihood of reference correlates with referential forms associated with higher levels of activation (e.g. pronouns, rather than names). However, Fukumura and Von Gompel (2010) found the opposite effect with State verbs: arguments consistent with continuation preferences were not more likely to be referred to with pronouns. Fukumura and Von Gompel (2010) conclude that verb-based expectations are separate from referent accessibility. They did however find that grammatical role affected form choice: continuations with subject antecedents tended to be pronouns. Based on these results our second research question is:

- Do arguments consistent with continuation biases have a greater chance to be referred to with a pronoun?

To answer these questions, we did a written continuation experiment with nonsense verbs. We had three Connective Types: because, so and but. We used because and so because these have been extensively used in previous continuation tasks and this allows useful comparisons. We also included but because a previous experiment Ehrlich (1980) argued that it has the ability to reverse continuation preferences found with because. However, it is not clear if this characterization is accurate: but tends to mark cause-effect sequences, like so, but with but the effect is unexpected (e.g. It’s raining but I didn’t get wet. Perhaps but patterns with so because of similar causal structure.

Method

We created 24 simple transitive sentence fragments with nonsense verbs. For each verb we created three versions each with because, so and but, e.g.

(2) Marcie knoffed Jared because/solbut...
Our results are evidence that connectives do contribute inde-
but (and marginally significantly less likely after so). Names were significantly more likely after but (and marginally significantly less likely after so) (See Table 2)

General Discussion
Our results are evidence that connectives do contribute inde-
but (and marginally significantly less likely after so). Names were significantly more likely after but (and marginally significantly less likely after so) (See Table 2)

Half the stimuli used an unambiguous female name as the subject, with an unambiguously male name as the object. The other half reversed this pattern. From three balanced lists of items-connectives we created eight random orders.

24 native speakers of English recruited via Amazon’s Me-
chancial Turk service completed the task. Subjects were in-
structed to write a natural continuation to each sentence.

Only continuations beginning with a reference to one of the characters in the fragment were analyzed. Thus we excluded cases where neither character was referred to (N=6), where both were referred to with e.g. they (N=6), and numerous other errors or non-target responses. Unexpectedly, many so responses were PURPOSE continuations (N=60) rather than RESULT and had to be excluded. For pronoun responses gender identified whether the antecedent was the previous subject or object. Two participants were excluded because they misunderstood the task.

Results
Subject or object continuations?
We used mixed-effect models (lmer in R) for analysis. The optimal model used connective as a predictor, and had participant as a random effect. There was a significant difference between because compared to but and so. (See Table 1) After but and so there is a significantly higher chance of an object continuation compared with because. There was no difference between but and so.

Name or pronouns continuations?
Using a mixed-effect model analysis as above we found the best model had connective as a predictor and participant and a random effect. Pronouns were significantly more likely after because than after but (and marginally significantly more likely after so). Names were significantly more likely after but (and marginally significantly less likely after so) (See Table 2)

Table 1: Percentage continuation with an object antecedent

<table>
<thead>
<tr>
<th>connective</th>
<th>because</th>
<th>so</th>
<th>but</th>
</tr>
</thead>
<tbody>
<tr>
<td>%age object-ant</td>
<td>67%</td>
<td>83%</td>
<td>87%</td>
</tr>
<tr>
<td>N=</td>
<td>105</td>
<td>116</td>
<td>104</td>
</tr>
</tbody>
</table>

Table 2: Percentage continuations that were names

<table>
<thead>
<tr>
<th>connective</th>
<th>because</th>
<th>so</th>
<th>but</th>
</tr>
</thead>
<tbody>
<tr>
<td>%age names</td>
<td>11%</td>
<td>20%</td>
<td>27%</td>
</tr>
<tr>
<td>N=</td>
<td>18</td>
<td>28</td>
<td>33</td>
</tr>
</tbody>
</table>

Table 3: Continuations by role, form and connective

<table>
<thead>
<tr>
<th>continuation</th>
<th>form</th>
<th>because</th>
<th>but</th>
<th>so</th>
</tr>
</thead>
<tbody>
<tr>
<td>object names</td>
<td>91</td>
<td>72</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>subjects pron</td>
<td>48</td>
<td>15</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>names</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

their continuations and for the referential form of such continu-
ations. The difference found between the connectives is con-
sistent with results for Transfer and Action verbs in Stevenson et al. (1994) where there was a general preference for object continuations, but like or results because showed a lower rate of object continuations than so (57% (Transfer) and 64% (Ac-
tion) for because compared to 77% (Transfer) and 80% (Ac-
tion) for so). They explain this as a general tendency to focus on the endpoint of events. If nonsense verbs are interpreted more like Transfer and Action verbs with event structures, contrary to State verbs, this could account for some of our re-
sults. However this account does not give an explanation for why in both studies these preferences would be reduced with because.

For referential choice, the difference between because and but is most interesting. Given the gender manipulation, all continuations were unambiguous, so for referent identifica-
tion names were unnecessary. Still, in approx 30% of continuations with but, names were used. A possible explanation may be the ability of but to associate with focus in marking contrasts between arguments. In our stimuli, using a name focusses attention, perhaps signalling contrast. Alternatively, perhaps but lead to more names because it signals an un-
expected continuation. Further analysis of the meanings of the actual continuations might help distinguish between these two explanations and is planned for future work.

Theoretically, the overall preference for continuations to refer to a previous object argument was unexpected. Mul-
tiple theories suggest a strong overall tendency for subjects to tend to have previous subjects as their antecedents (e.g. the Parallel Preference Theory, Centering Theory). However our results are consistent with previous continuation studies finding object continuation preferences (e.g. (Arnold, 2001), with transfer verbs and Fukumura and Von Gompel (2010) with State verbs). Arnold attributes the discrepancy to a task effect. Fukumura and Von Gompel (2010) instead argue that it is additional evidence that referential choice is separate from grammatical role continuation preferences: by frequency alone, objects were slightly more likely as con-
tinuations in their study, but participants used proportionally more pronouns with subject continuations than object continuations.

In our results, object continuations were much more likely than subject continuations. Also, like Fukumura and Von Gompel (2010), 93% of subject continuations were pronouns compared to 77% of object continuations, consistent with the
conclusion that referential choice is separate from structural expectations. (See Table 3 for an overview).

Our next step is to follow the methodology of Fukumura and Von Gompel (2010) and explicitly mark which character (the subject or the object) that experiment participants should begin their continuations with, allowing us to collect a more balanced set of subject and object. With such results we should be able to examine the interaction between the grammatical role and referential choice in more depth.

References

