CONFIDENTIAL - FOR PEER-REVIEW ONLY

rnn word order - Heavy NP Shift (#11400)

Created: 05/30/2018 11:54 AM (PT)
Shared: 05/30/2018 11:55 AM (PT)

1) Have any data been collected for this study already?
It's complicated. We have already collected some data but explain in Question 8 why readers may consider this a valid pre-registration nevertheless.

2) What's the main question being asked or hypothesis being tested in this study?
RNNs show heavy NP shift: a preference for the order V PP NP only when the NP is long.

3) Describe the key dependent variable(s) specifying how they will be measured.
Total sentence surprisal in the Jozefowicz et al. (2016) LSTM and the Gulordava et al. (2018) LSTM.

4) How many and which conditions will participants be assigned to?
2x2x2

* Order is V NP PP (unshifted) or V PP NP (shifted)
* NP is short or long
* the V is NP-biased or NP/S-biased

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We predict surprisal from the experimental conditions with linear regression. We expect an interaction whereby the (shifted, long) condition is easier than would be expected from shiftedness and NP length.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
We will remove all items that turn out to contain UNKs.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
We have adapted 40 items from Stallings et al. (1998) and we believe that will be enough.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
Surprisals have been calculated but not looked at or analyzed.

We will also look for an effect of verb bias modulating the badness of shiftedness and modulating the heavy NP shift interaction, as reported in Stallings et al. (1998).

We will also look at where the surprisal penalties appear. We expect that in the [unshifted, long] condition there will be large penalty at the PP which is reduced in the [shifted, long] condition, while the penalty at the NP remains roughly the same.