9.19: Computational Psycholinguistics, Pset 7
due 24 November 2021

10 November 2021

As with Pset 6, we provide code that does most of the heavy lifting for this pset, but you will need to keep track of everything as you progress through the assignment. You can access this code and complete this pset through an interactive Colab notebook we've prepared, at https://colab.research.google.com/drive/1KPDfk_ItQauV8RgyPyQ-3QL7-OPERBQz?usp=sharing. Please read and follow all instructions in the notebook. You will need to write code or textual responses in every place marked as TODO.

Testing a neural language model like a psycholinguistics subject

Some English verbs, such as *accept*, *anticipate*, *regret*, and *believe*, have two possible SUBCATEGORIZATION FRAMES: they can take either a direct object (DO) NP or a complement clause (CC). Complement clauses in turn can be optionally introduced by the word *that*. Thus all the following sentences are OK in English:

(1) a. The congresswoman accepted [$_{\text{DO}}$ the changes to the bill].
b. The congresswoman accepted [$_{\text{CC}}$ the changes would be met with resistance].
c. The congresswoman accepted [$_{\text{CC}}$ that the changes would be met with resistance].

(2) a. The loan officer anticipated [$_{\text{DO}}$ a deluge of applications].
b. The loan officer anticipated [$_{\text{CC}}$ a huge number of people would submit applications].
c. The loan officer anticipated [$_{\text{CC}}$ that a huge number of people would submit applications].

(3) a. The administrators regretted [$_{\text{DO}}$ their remarks].
b. The administrators regretted [$_{\text{CC}}$ their remarks were susceptible to misinterpretation out of context].

c. The administrators regretted [$_{\text{CC}}$ that their remarks were susceptible to misinterpretation out of context].

(4) a. The attorney believed [$_{\text{DO}}$ the senator.]
b. The attorney believed [$_{\text{CC}}$ the senator was lying].
c. The attorney believed [$_{\text{CC}}$ that the senator was lying].

As some of the above examples make clear, for this type of verb, incremental sentences of the form

(5) Subject Verb NP ...

have a LOCAL SYNTACTIC AMBIGUITY: the NP could be the direct object of the verb, or it could be the beginning of a complement clause. On the other hand,

(6) Subject Verb that NP ...

are not ambiguous: the word *that* makes clear that the NP is the beginning of a complement clause. One of the earliest results in the use of eye movements during reading to monitor incremental language understanding is that this local syntactic ambiguity can lead to garden-path disambiguation effects (Frazier & Rayner, 1982). Sometimes this garden-path effect rises to the level of conscious awareness (e.g., you may have found the end of (3-b) a bit jarring), but even more subtle cases of this garden-pathing effect that do not rise to conscious awareness are measurable in the eye movement record.

Among these verbs, the SUBCATEGORIZATION BIAS varies: some verbs take direct objects more often than complement clauses (e.g., *accept*), and some take complement clauses more often than direct objects (e.g., *believe*). Garnsey et al. (1997) showed that the garden-path effect is modulated by verb subcategorization bias, with resolution to a low-probability subcategorization frame creating a bigger garden-path effect than resolution to a high-probability subcategorization frame. In the first part of this pset, you will test a neural language model as you would a human psycholinguistic subject, to see whether it shows this human-like incremental processing effect.

Like Garnsey et al. (1997), we will compare versions of complement-clause sentences with versus without *that*. **The following experiment design has a problem**, and it is your job to fix that problem.

Design: consider the following sentence prefix:

(7) The senator accepted the changes...

Now, *the changes* is ambiguous between being a direct object versus the start of a complement clause for *accepted*. If it is the start of a complement clause, there is a good chance that the next word will be the verb of the complement clause. If it is the direct object, then something other than a verb will come next. The probability of whatever material X comes

next can be written as the following marginalization:

$$
\begin{aligned}
&P(X|\text{Context}) = \\
&\quad P(X|\text{Context}, \textit{accepted} \text{ is DO}) \times P(\textit{accepted} \text{ is DO}|\text{Context}) + \\
&\quad P(X|\text{Context}, \textit{accepted} \text{ starts CC}) \times P(\textit{accepted} \text{ starts CC}|\text{Context})
\end{aligned}
$$

The second part of each term on the right-hand side, $P(\textit{accepted}$ {is DO, starts CC}$|$Context), makes manifest the role of the verb bias.

Considering all this, we will look at the effect of including or omitting the word *that* before a complement clause on the surprisal of what comes after a post-verbal *the Noun* sequence. The template we will use is given below, with the part of the sentence for which we will compute surprisals underlined:

(8) a. The senator accepted the changes to the bill...
b. The senator accepted that the changes to the bill...
c. The senator believed the changes to the bill...
d. The senator believed that the changes to the bill...

Since adding *that* rules out the direct-object interpretation and makes it possible for a verb to appear after *the changes*, it should lower the probability of *to the bill* and thus make this continuation more surprising. But this effect should be smaller for *believed* than for *accepted*, because the probability of a direct-object interpretation for *believed* was low in the first place (due to its subcategorization bias). So, if we denote by S(example) the surprisal given the context of the underlined region of the example referred to, we predict that human-like processing behavior would entail the following inequality:

$$S((8\text{-b})) - S((8\text{-a})) > S((8\text{-d})) - S((8\text{-c})) \tag{1}$$

Thus, we will feed sentences like (8) into a left-to-right neural language model, assess the surprisals of the underlined regions, and determine whether the inequality of Equation (1) holds.

Task 1: most of the above reasoning is sound, but there is a subtle problem with the logic that has to do with the syntax of sentences like (8). Identify the problem with the logic, fix the design and crucial tests to eliminate this problem, develop any new experimental materials you need, and test your neural language model on these experimental materials. You should create at least 8 items following the template (8), so you will need to use at least 8 DO-biased verbs and 8 CC-biased verbs. (You can get additional verbs of varying bias from the appendices of Garnsey et al. (1997), which you can download from `Stellar | Materials | Readings`.)

Task 2: choose a human syntactic or sentence processing phenomenon from the list below. Develop a set of materials (with at least 8 items; preferably 16 or more) and an evaluation criterion. You can use the SyntaxGym website (http://syntaxgym.org), which is under active development in our lab, and consult our lab's recent papers, including Futrell et al. (2018), Wilcox et al. (2018), and others, for numerous examples. **Considerations:** a

good test suite should be one that an n-gram model with reasonably low n (e.g., a 5-gram model) can possibly succeed on. It is also crucial not to use out-of-vocabulary words in your test suite, as the model could fail for uninteresting reasons (it has never encountered the word). The best test suite items are ones where the model you're testing has at least several exposures to every word appearing in the test suite.

Optional Task 3 (no credit, but we will read and provide feedback on your work if you do this): use a different language model—train the Pset 6 model architecture on a different (maybe bigger) dataset, and/or change the language model architecture. How does your new language model's performance on the MV/RR, DO/CC, and your test suite compare with the model from Pset 6? How do the models' perplexities compare? What conclusions do you draw?

Syntax and/or sentence processing phenomena

We will continue to add phenomena to this list between the pset's initial release and due date. If you have an idea for a phenomenon not on this list that you would like to develop a test for, please contact the instructors and describe your idea; we will give you feedback and help you determine whether it is a suitable test.

Noun phrases versus adjective phrases. Adjectives can be used either inside a noun phrase, as in (9-a), or in a predicative position as an adjective phrase on their own, as in (9-d). If they're used inside a noun phrase that starts with an indefinite determiner, there should be a noun after the adjective phrase.

(9) a. That is a [$_{\text{AdjP}}$ very tall and rather beautiful] tree.
b. *That is a [$_{\text{AdjP}}$ very tall and rather beautiful].
c. *That is [$_{\text{AdjP}}$ very tall and rather beautiful] tree.
d. That is [$_{\text{AdjP}}$ very tall and rather beautiful].

V2 in English. V2 is a construction English has inherited from German. When an adjunct get topicalized because the speaker wishes to emphasize it, the following verb can invert its normal position and comes second in the sentence right after the topic (hence V(erb)2; "verb second"). When there is no special topic, the order should not change, even when there is preceding material due (for example) to an initial subordinate clause.

(10) a. Only a single time before today have I seen such beautiful orchids.
b. ?Only a single time before today I have seen such beautiful orchids.
c. *While visiting the park before today have I seen such beautiful orchids.
d. While visiting the park before today I have seen such beautiful orchids.

Auxiliary inversion and free relative clauses. When an English sentence begins with a *wh-* word like *what* or *who*, there are two possibilities: (i) it is the beginning of a question, or

(ii) it is the beginning of a FREE RELATIVE CLAUSE. If it is the beginning of a question, then the sentence should have AUXILIARY INVERSION: the subject and verb should be swapped in position, as in (11-a). (If the finite verb of the sentence is not an auxiliary verb or the copula, then there is also *do*-support, which e.g. Wikipedia has a good article on.) If it is the beginning of a free relative clause, then there is no auxiliary inversion, as in (11-d).

(11) a. What would you like to do this afternoon before we go home?
b. *What you would like to do this afternoon before we go home?
c. *What would you like to do this afternoon is very important to me.
d. What you would like to do this afternoon is very important to me.

Progressive Sentential Subjects. NP + V-ing (present progressives) can serve as the subjects of sentences, in which case a matrix clause must follow as in (12-a) but not (12-b). Progressives cannot serve as the main clause of the sentence, which must be tensed, as in (12-d)

(12) a. You joking with the director will only get us in trouble.
b. *You joked with the director will only get us in trouble.
c. *You joking with the director.
d. You joked with the director.

Tense Adverbs. In some languages, such as Chinese, tense marking is done entirely by tense adverbs like "tomorrow" or "yesterday." English has both tense marking on verbs as well as adverbs of time, but there must be agreement between these two. Future tense adverbs can only modify VPs that have future tense as in (13-a), while past tense adverbs can only modify VPs that have past tense as in (13-d).

(13) a. He will go to the concert in the park tomorrow.
b. *He went to the concert in the park tomorrow.
c. *He will go to the concert in the park yesterday.
d. He went to the concert in the park yesterday.

Optional: Compare the results obtained with different positioning of the adverb—not just sentence-final but also sentence-initial and before post-verbal prepositional phrases:

(14) a. He will go to the concert in the park tomorrow.
b. Tomorrow he will go to the concert in the park.
c. He will go tomorrow to the concert in the park.

Does the model's success/failure to learn about tense adverbs generalize across various adverb positionings?

Solution for Task 0

The results you get for Task 0 will depend a bit on the exact parameters of the language model you wound up training in Pset 6. You are likely to see that the main verb/reduced relative clause comparison are reasonably successful for:

1. the effect of relative clause reduction for part-of-speech-ambiguous participial verbs: surprisal at the main verb is generally greater when the relative clause is reduced than when it is unreduced (though the effect size may be small, and not for hugely above 50% of items);

2. the effect of part of speech ambiguity: surprisal at the main verb is generally greater when the participial verb has the same form as in the simple past (as with *brought*) than when the forms are different (as with *given* versus *gave*).

Typically we have found that LSTM models trained on a small amount of text, as you did in Pset 6, do not perform very well on the most stringent MV/RR prediction, that the effect of relative clause reduction on main-verb surprisal is greater for part-of-speech-ambiguous verbs than for part-of-speech-unambiguous verbs (see Futrell et al. (2019), Figure 7, lower right panel). This is likely because there are a lot of facts to put together (the syntactic patterns, the overlapping but distinct distributional possibilities for the different wordforms) in order to get this pattern, and without explicit supervision for grammatical structure more data may be required for these generalizations to emerge.

References

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*, 178–210.

Futrell, R., Wilcox, E., Morita, T., & Levy, R. (2018). RNNs as psycholinguistic subjects: Syntactic state and grammatical dependency.

Futrell, R., Wilcox, E., Morita, T., Qian, P., Ballesteros, M., & Levy, R. (2019). Neural language models as psycholinguistic subjects: Representations of syntactic state, In *Proceedings of the 18th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*.

Garnsey, S. M., Perlmutter, N. J., Meyers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, *37*, 58–93.

Wilcox, E., Levy, R. P., Morita, T., & Futrell, R. (2018). What do RNN language models learn about filler–gap dependencies? In *Proceedings of the workshop on analyzing and interpreting neural networks for NLP*.