

Thanks for virtually visiting our poster! Humans can use pragmatic reasoning in discourse processing, making reference to world knowledge and experiences beyond language use. Our work asks if models trained with no knowledge beyond what is available in language data can acquire linguistic behavior akin to those making reference to pragmatic knowledge. In doing so, we show that world knowledge and experience are not necessary for exhibiting at least some aspects of pragmatic reasoning.

Zoom in to background

In particular, we are interested in stimuli like this. In reading, "the knight killed" there is an ambiguity in whether "killed" is a main verb or a reduced relative. If "killed" is a main verb then we expect continuations like (a) "the knight killed the dragon and fell to the ground". If instead, "killed" is a reduced relative, then we expect continuations like, "the knight killed by the dragon fell to the ground". In isolation, continuations like (b) lead to a reading time slow down compared to (a), known as a garden path effect.

scroll down to context alleviation

However, we know that prior context can alleviate these effects for humans. In particular, we can use discourse context to accommodate conversational implicatures of the relative clause in (b). I've schematized this in the following diagram. If the prior context introduces two knights, one of whom was killed by a dragon, then reading "the knight killed by the dragon fell to the ground" becomes more preferred. That is, we expect "knight" to be modified when there are multiple knights, which can alleviate the garden path effect. We refer to this as referent context alleviation. Additionally, temporal contexts can alleviate garden path effects. For example, if we are talking about the future actions of a knight and then read "the knight killed" we prefer the reduced relative because "killed" as a main verb would mean we were talking about the past instead of the future. We refer to this as temporal context alleviation. Finally, information structure and definiteness can alleviate garden path effects for a related type of stimulus. We give more details on this in our paper. All together these alleviation effects have been shown for humans in various studies.

scroll over to research question

This leads us to our overarching research question, which is: how much pragmatic reasoning, like these garden path alleviation effects, is contained within the linguistic signal? More specifically, can human-like garden path alleviation effects be learned from text data alone? To address this question, we used recurrent neural network language models trained on Wikipedia text. The models were trained on one of two conditions: the first is trained on ordered Wikipedia text, which preserves the contexts that each sentence occurs in. The second shuffles the same data by sentence. Shuffled data may seem odd, but it is not uncommon in computational linguistics being used in a number of papers that use psycholinguistic experiments and methods to probe the linguistic representations learned by recurrent neural networks. In our study,

training on shuffled data serves as our baseline condition, as it is trained with no meaningful contextual information.

scroll down to results

Turning to our results, the first plot in this figure shows the canonical garden path effect for both model training conditions. That is, ambiguous stimuli that cause garden path effects for humans are more surprising than those that don't. The next four plots show the amount of contextual alleviation for each training condition for each type of context. We see across the board that models trained on shuffled data (in other words trained without knowledge of larger contexts) do not have any alleviation effects. For models trained on ordered data, we see that future contexts and indefinites reduce the models' dispreference for ambiguous stimuli. This is in-line with the human results. However, contexts with 2 referents and information status do not alleviate garden path effects for the models, but see our paper for more details on these results.

scroll to conclusion

To conclude, we see that both training conditions exhibit the garden path effect. Temporal contexts and definiteness can alleviate these effects, but only for models trained on ordered data. This suggests that multi-sentence discourse effects such as tense and use of indefinites on sentence processing can be acquired from language data alone. Zooming out a little, we take this work to point to two possible conclusions: 1) some kinds of garden path alleviation may not be evidence of the use of pragmatic reasoning in humans, or 2) some aspects of pragmatic reasoning are learnable from just language data.

Thanks and feel free to contact me if you have any questions!