What do RNN Language Models Learn about Filler-Gap Dependencies?

Ethan Wilcox, Roger Levy, Takashi Morita and Richard Futrell

25.1.19

Filler-gap dependency

- **filler** is a wh-complementizer such as 'what' or 'who'.
- **gap** is an empty syntactic position licensed by the **filler**.

Examples

- I know what the lion devoured __ at sunrise.
- *I know that the lion devoured __ at sunrise.

Method

Models

- **Google model** - BIG LSTM+CNN Inputs - *Jozefowicz et al. (2016)* - uses the output of a character-level CNN as input to the LSTM, trained on One Billion Word Benchmark, two hidden layers with 8196 units each.
- **Gulordava model** - LSTM *Gulordava et al. (2018)* - trained on 90 million tokens of English Wikipedia, it has two hidden layers of 650 units each.
- **Baseline** - 5-gram model trained on One Billion Word Benchmark.

Method

Suprisal

- **suprisal values**: $S(x_i) = -log_2 p(x_i|h_i - 1)$, where x_i is the current word or character and $h_i - 1$ is the LSTM's hidden state before consuming x_i, with the probability calculated from the RNN's soft-max activation.
- The surprisal of a word or a sentence tells us the extend to which the given word or sentence is unexpected under the language model's probability distribution.
- Wilcox et al. measure surprisal in two places, at the word following the gap and summed over the whole embedded clause following the gap.

Method

Expectations

- **wh-licensor, no gap** - expect higher summed surprisal.
- **no wh-licensor, gap** - expect higher summed surprisal.
- The wh-licensing interaction is calculated by:
  (S(wh-licensor, no gap) - S(no wh-licensor, no gap)) -
  (S(wh-licensor, gap) - S(no wh-licensor, gap))

Example

- I know that the lion devoured a gazelle at sunrise.
- *I know what the lion devoured a gazelle at sunrise.
- *I know that the lion devoured __ at sunrise.
- I know what the lion devoured __ at sunrise.

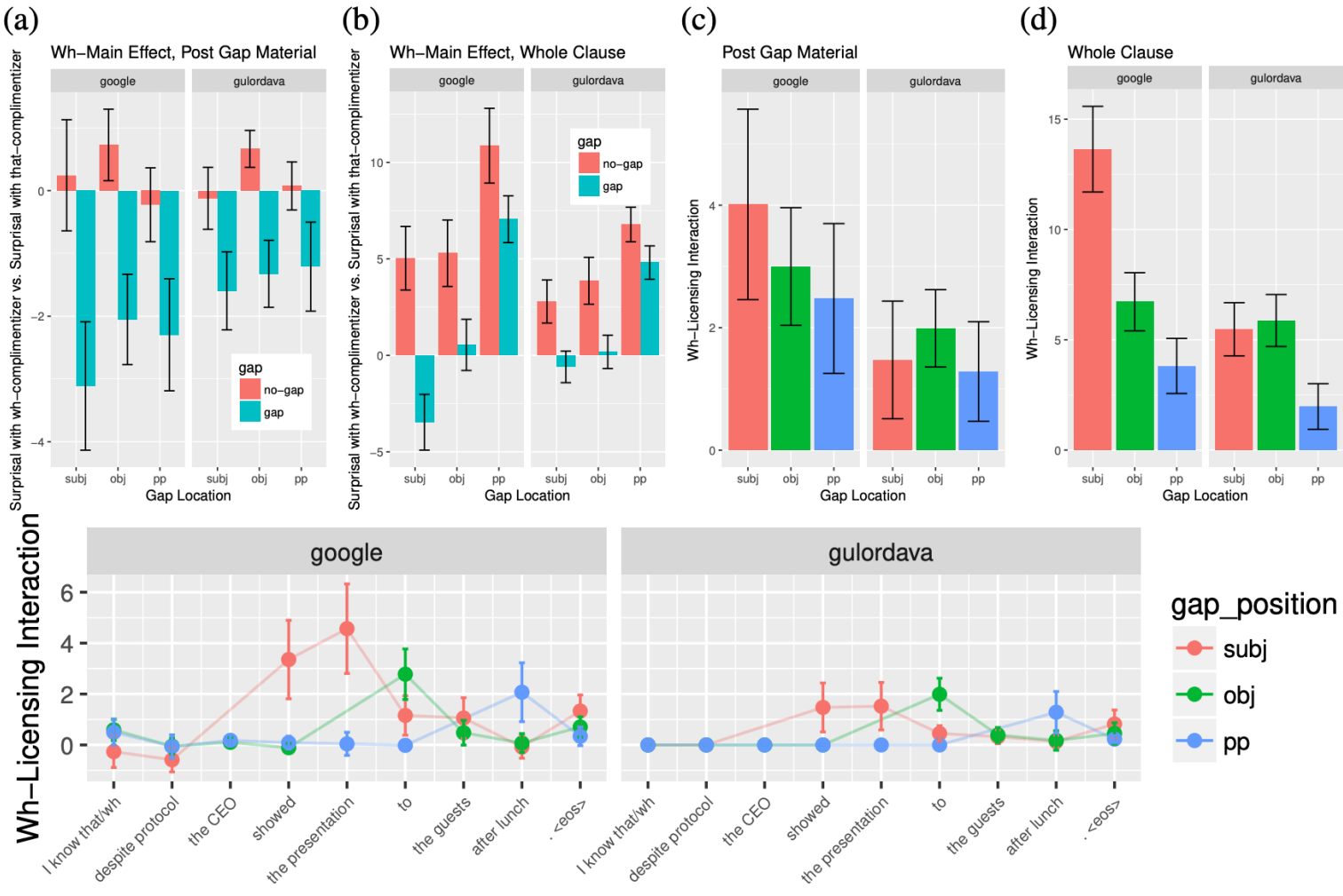
Filler-gap dependencies

Flexibility of Wh-licensing:

Do LSTM's learn filler gap dependencies when the gap appears in subject, object and indirect object?

- I know who __ showed the presentation to the visitors yesterday. (subject)
- I know what the businessman showed __ to the visitors yesterday. (object)
- I know who the businessman showed the presentation to __ yesterday. (indirect object)

Filler-gap dependencies



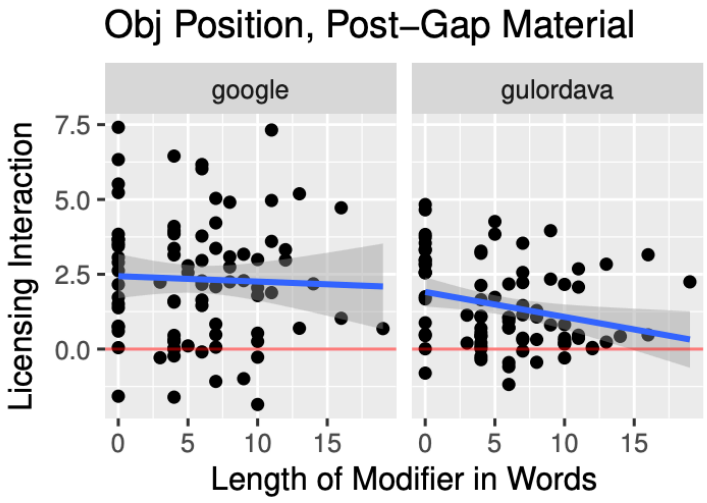
Filler-gap dependencies

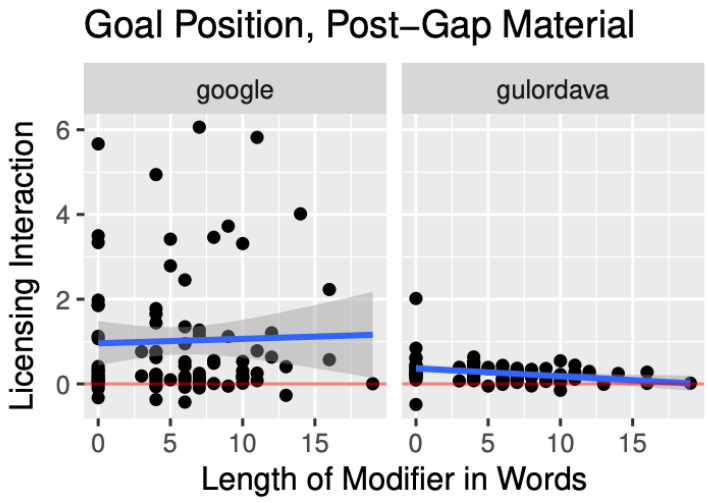
Robustness of Wh-licensing:

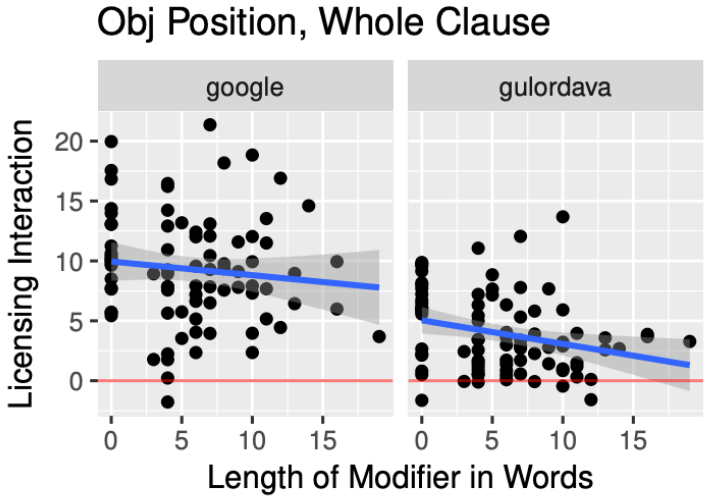
Does the number of words separating filler from gap affect LSTM learning filler-gap dependencies? Small intervention: 3-5 words, medium interventions: 6-8 words, long intervention: 8-12 words.

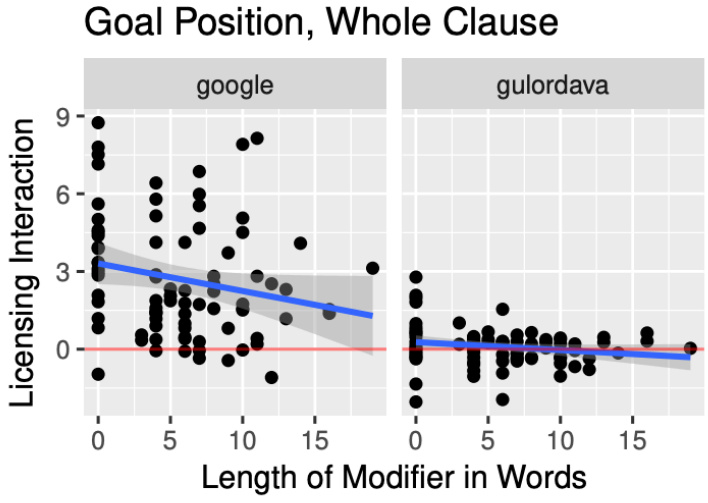
- I know what your friend gave __ to Sam during the picnic yesterday.
- I know what your **new** friend **from the south of France who only just arrived last week** gave __ to Sam during the picnic yesterday. (12 words - long intervention)

Filler-gap dependencies









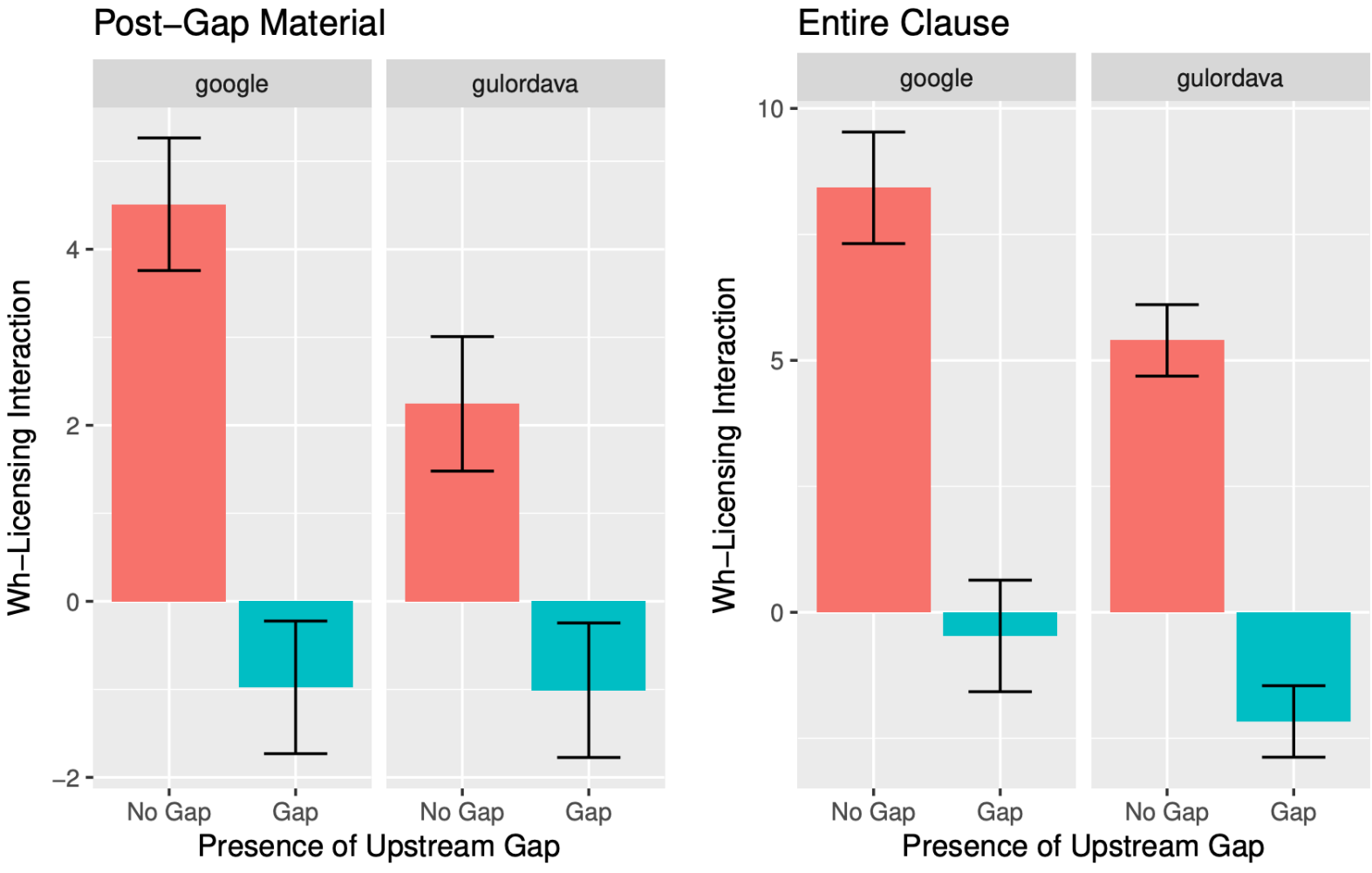
Syntactic Islands

Wh-Island Constraint:

Do LSTM's learn that a gap cannot appear inside double nested clauses headed by wh complementizers?

- I know what Alex said your friend devoured __ at the party. *null-comp*
- I know what Alex said that your friend devoured __ devoured at the party. *that-comp*
- *I know what Alex said whether your friend devoured __ at the party. *wh-comp*

Syntactic Islands



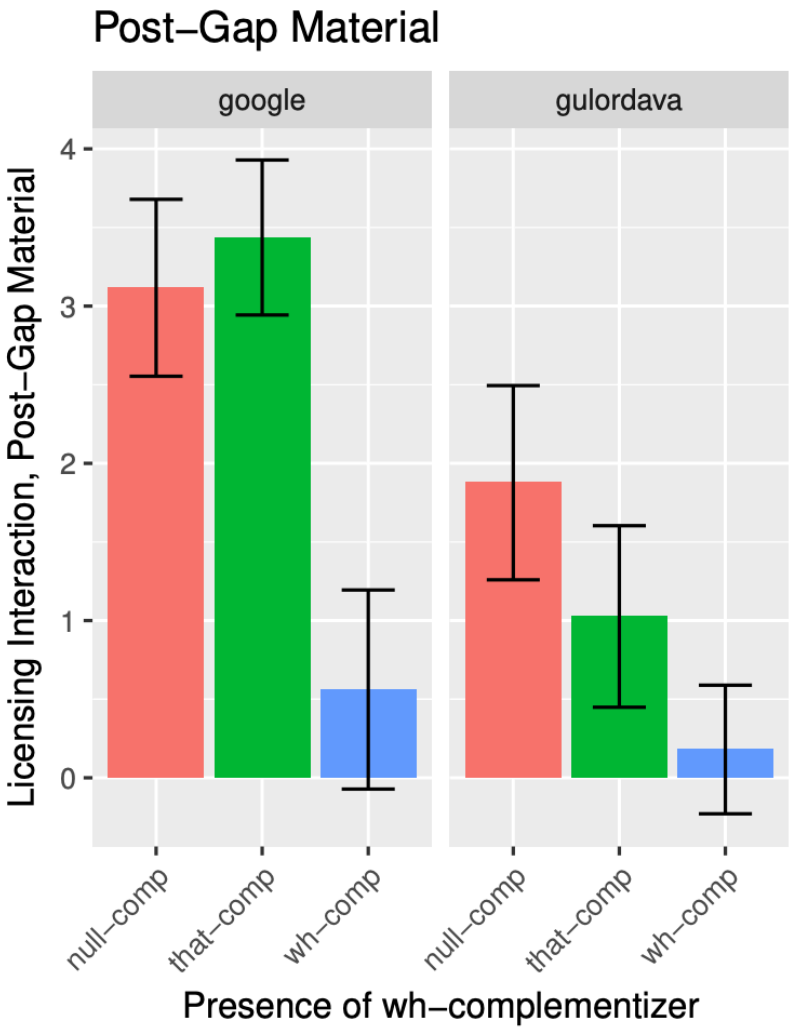
Syntactic Islands

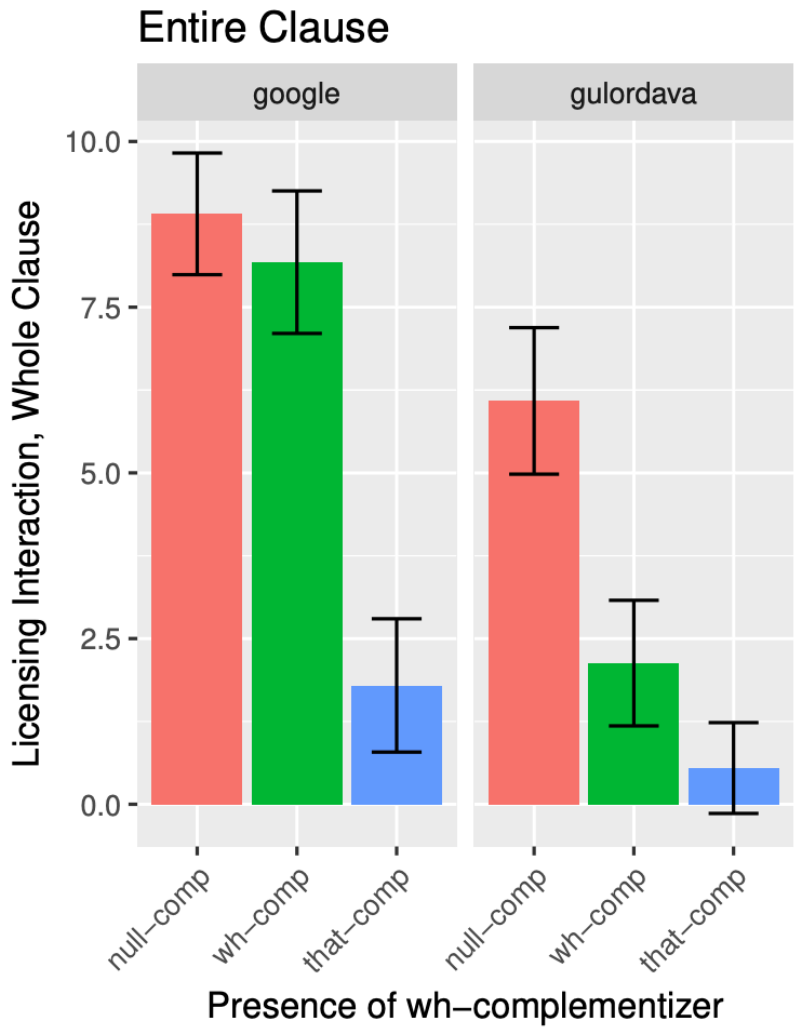
Adjunct Island Constraint:

Do LSTM's learn that a gap cannot be licensed in an adjunct clause?

- I know what the librarian in the dark blue glasses placed __ on the wrong shelf. *object*
- *I know what the patron got mad after the librarian placed __ on the wrong shelf. *adjunct-back*
- *I know what, after the librarian placed __ on the wrong shelf, the patron got mad. *adjunct-front*

Syntactic Islands





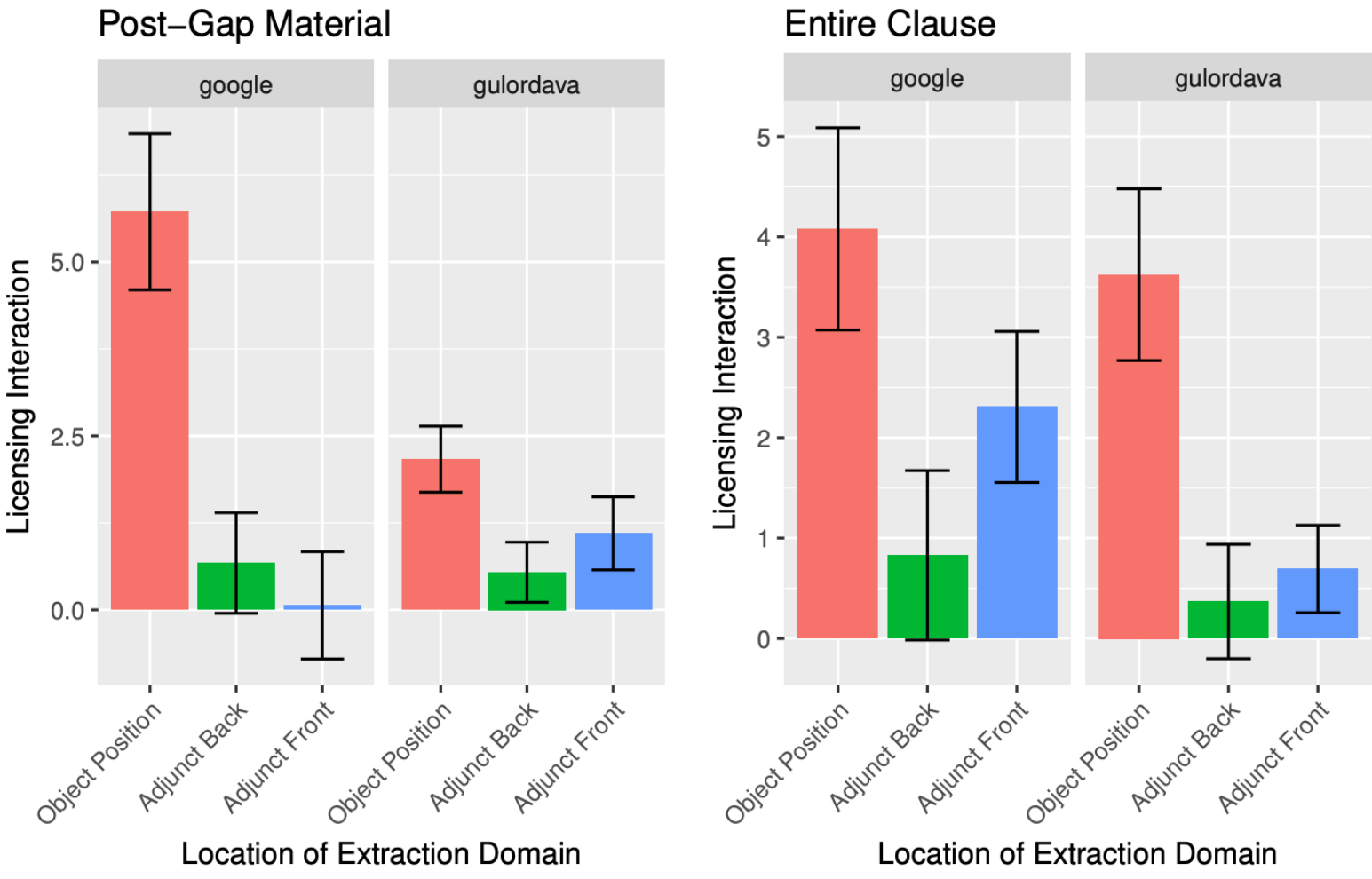
Complex NP and Subject Islands

Complex NP Constraint:

Do LSTM's learn that a gap cannot be hosted in a sentential clause dominated by a noun phrase with a lexical head noun?

- I know what the family bought __ last year. *(object)*
- *I know who the family bought the painting that depicted __ last year. *(that-rc/obj)*

Syntactic Islands



Complex NP and Subject Islands

Subject Constraint:

While CNPC generally doesn't apply to other NP modifiers, such as PPs, NP occurring in subject position isn't acceptable none the less.

- I know who the family bought the painting by __ last year. *(prep/obj)*
- *I know who the painting by __ fetched a high price at auction. *(prep/subj)*

Syntactic Islands