LSTMs Can Learn Basic Wh- and Relative Clause Dependencies in Norwegian

Anastasia Kobzeva *Norwegian University of Science and Technology (NTNU)*
anastasia.kobzeva@ntnu.no
Suhas Arehalli *John Hopkins University*
Tal Linzen *New York University*
Dave Kush *NTNU and University of Toronto*

BACKGROUND

Filler-gap dependencies: contingencies between **fillers** (like *what*) and **gaps** – positions where fillers are interpreted in a sentence.

(1) I know **what** the priest revealed __ at the party.

(2) I heard about **the secret** that the priest revealed __ at the party.

Wilcox et al. (2018, 2019) found that neural network without specific language bias can learn complex generalizations about *wh*-filler-gap dependencies like (1) from raw English text data.

We test the generality of this result by:

- Training and testing a similar model on Norwegian data
- Including relative clause (RC) dependencies like (2) into the test set

MODELS

1. LSTM RNN model trained on next word prediction task (ppl 30.4)
   - 113 million tokens from Norwegian Bokmål Wikipedia
   - Trained following procedure by Gulordava et al. (2018)
2. Baseline model: 5-gram model with Knesser-Ney smoothing (ppl 133.5)

DEPENDENT VARIABLE

Surprisal – inverse log probability that the model assigns to a word given the previous context: $S(w_n) = -log_2 p(w_n|h_{n-1})$, where h_{n-1} is LSTM's hidden state before consuming w_n.

MEASURING FILLER-GAP DEPENDENCIES

Manipulating the presence of a filler and the presence of a gap:

-GAP CONDITIONS:

a. She knows that the priest revealed *the secret* at the party. -FILLER
b. *She knows what the priest revealed *the secret* at the party. +FILLER

Filled-gap effect (FGE): surprisal difference between **b** and **a** at *the secret*. FGEs should be **positive** (surprisal: high - low)

+GAP CONDITIONS:

c. She knows what the priest revealed *__ at the party.* +FILLER
d. *She knows that the priest revealed *__ at the party.* -FILLER

Unlicensed gap effect (UGE): surprisal difference between **c** and **d** at *at the party*. UGE should be **negative** (surprisal: low - high)

EXPERIMENTS

Across two dependency types, we explore whether the model:

1. Can learn the **flexibility** of filler-gap licensing: fillers can license gaps in multiple syntactic positions (Subject, DO, and OBL):

(3) She knows that the priest revealed the secret in front of the guests at the party.

2. Can establish dependencies across increased linear **distance:**

(4) I heard about the secret that the priest [in a black robe and white collar] revealed __ at the party.

Tested no, short (2-4), medium (5-8) and long (9-12) subject modifiers.

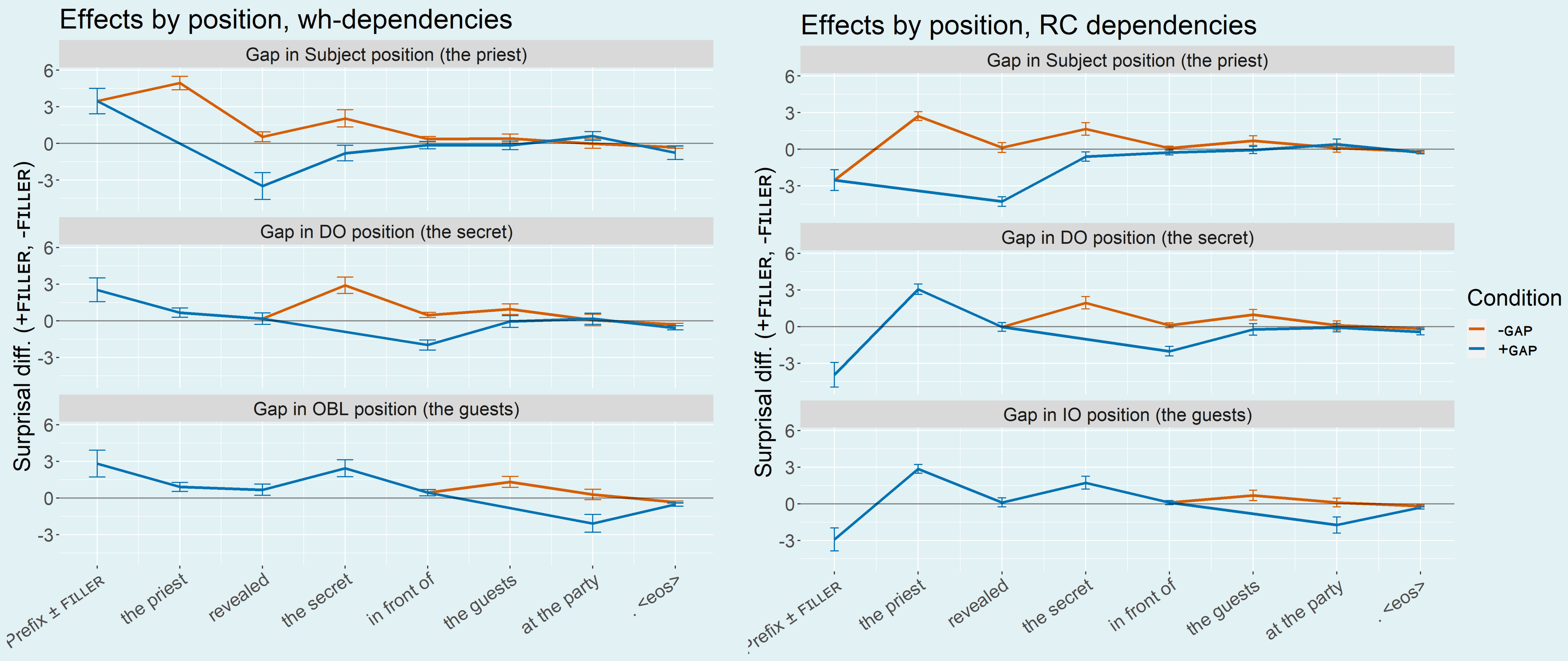
RESULTS

The LSTM model represents filler-gap dependencies across two dependency types in Norwegian by:

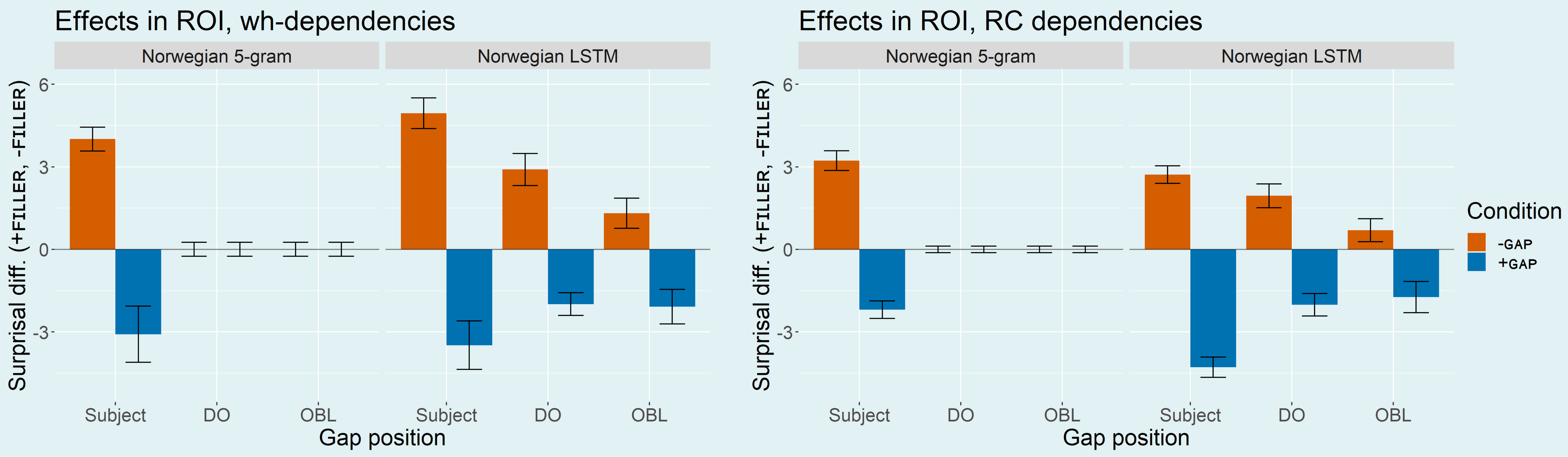
- Showing filled-gap and unlicensed gap effects
- Exhibiting them across all syntactic positions despite increased linear distance between the filler and the gap

General-purpose neural networks can learn basic syntactic generalizations about filler-gap dependencies in Norwegian

1. Flexibility of filler-gap licensing

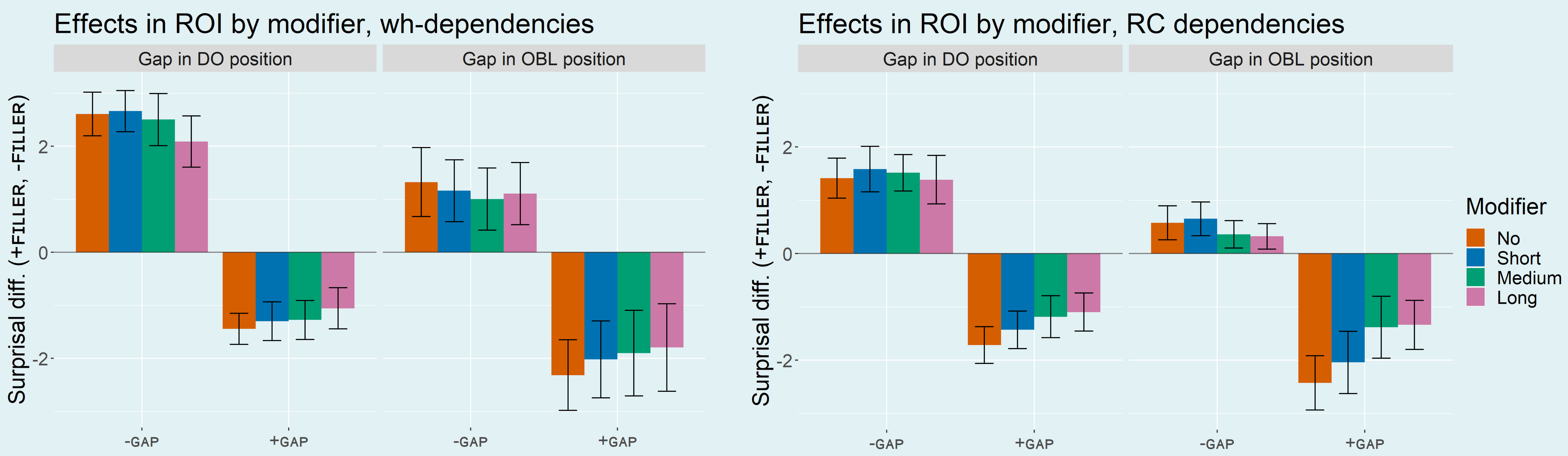


- FGEs and UGEs across all positions for both dependency types
- FGEs are found in multiple filled NP positions



- 5-gram model showed the effects only in Subj position
- Size of the effects varies by position: Subj > DO > OBL

2. Linear distance between the filler and the gap



- The effects are present across all modifier lengths
- Small effect of increased linear distance between the filler and the gap

CONCLUSIONS & FUTURE WORK

The LSTM model could learn basic properties of filler-gap dependencies in Norwegian, in line with Wilcox et al.'s findings for English. It had strongest expectation for Subject gaps, followed by DO and OBL gaps for both dependency types

- Humans do not exhibit subject FGEs with *wh*-dependencies but do so with relative clauses (Stowe, 1986; Lee, 2004)
- The model is more likely to be affected by corpus statistics

Future work will test if the model can learn more properties of filler-gap dependencies, as well as constraints on them known as *islands* (Ross, 1967)