Puzzles of Agreement: Syntactic, Semantic, and Psycholinguistic perspectives
October 24–25, 2024

Modelling agreement attraction effects in vector space

Eva Neu[a], Maayan Keshev[b] and Brian Dillon[a]

[a]University of Massachusetts Amherst, [b]The Hebrew University of Jerusalem

Agreement attraction effects

Agreement dependencies are subject to interference from distractors (e.g., Bock and Miller, 1991; Wagers et al., 2009):

(1) The **key** to the **cabinets** is/*are on the table.

Not all nouns are equally distracting (Franck et al., 2002; Bock and Cutting, 1992):

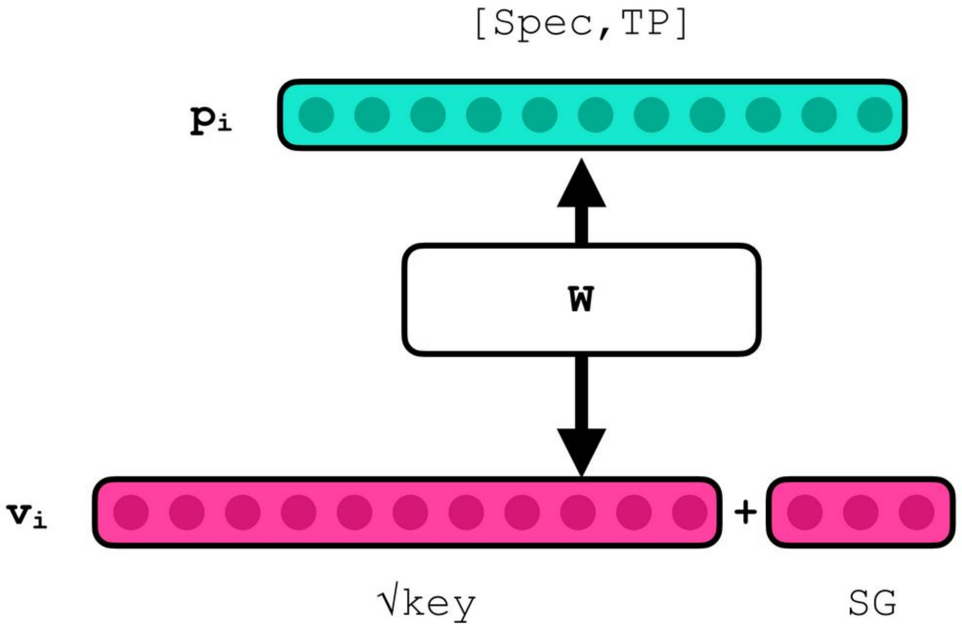
(2) The **helicopter** for the **flight** over the **canyon** is rusty.

Both syntactic and semantic similarity play a role – we focus here on syntax.

Agreement attraction effects

Keshev et al. (accepted): agreement attraction effects are due to the way sentences are encoded in working memory

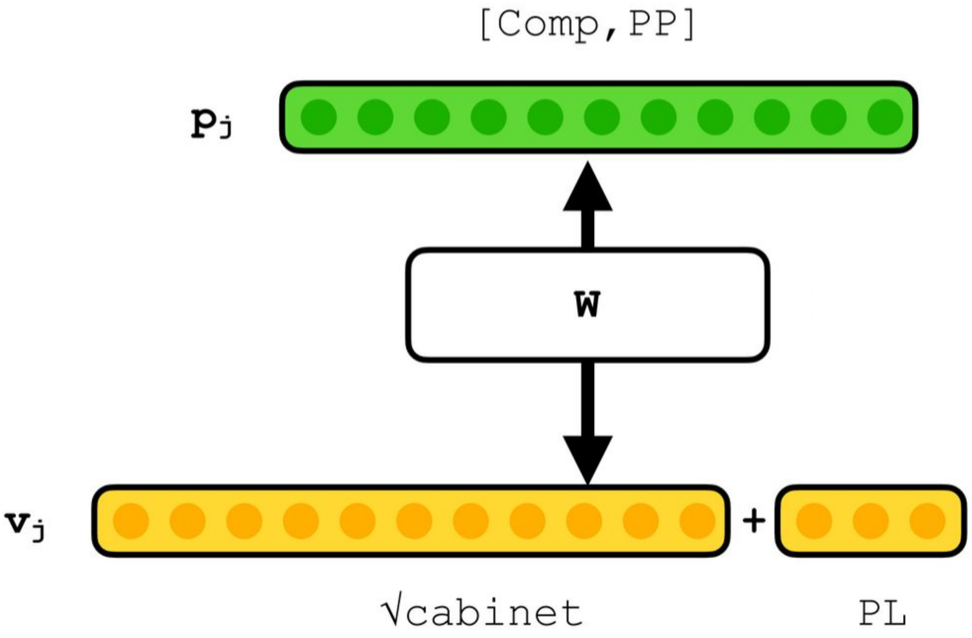
- Lexical items are bound to syntactic positions
- Items and positions are each represented as vectors



Agreement attraction effects

Keshev et al. (accepted): agreement attraction effects are due to the way sentences are encoded in working memory

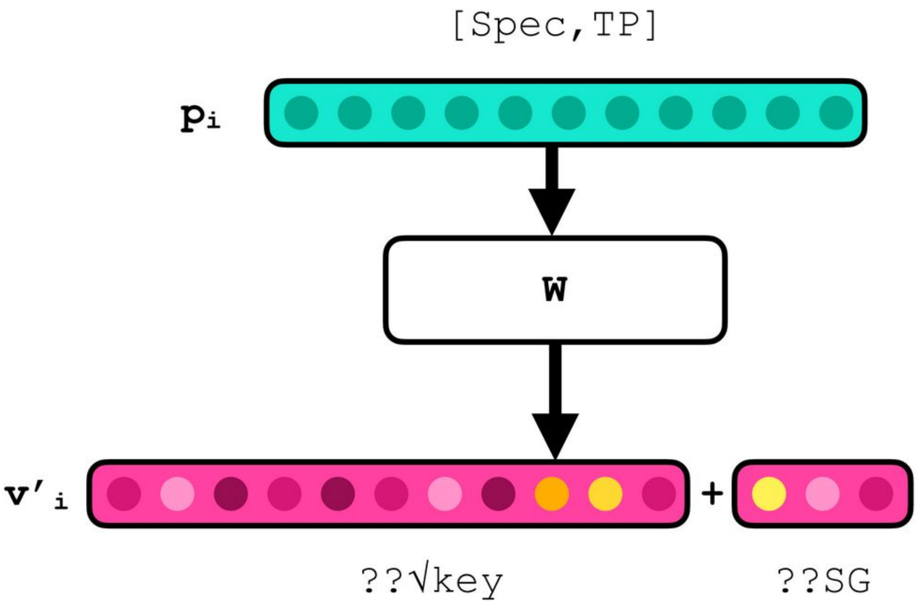
- ➢ Lexical items are bound to syntactic positions
- ➢ Items and positions are each represented as vectors



Agreement attraction effects

Agreement: speakers need to retrieve item bound to subject position

- Retrieval is prone to errors
- Gradient similarity between syntactic positions
- Higher similarity leads to higher rates of erroneous retrieval



Agreement attraction effects

Our goal for today: build a vector-symbolic representation of syntax such that…

- ➢ Each syntactic position is assigned a position vector
- ➢ Higher (cosine) similarity between target and distractor corresponds to higher rates of interference

A vector-symbolic metric of syntactic similarity

Step 1: Assign a constituency parse

- ➢ Berkeley Neural Parser
- ➢ Binary branching is enforced

Step 2: Assign each node a base vector depending on category

- ➢ E.g., same base vector for all N/NP nodes
- ➢ Different base vectors are orthogonal to each other

A vector-symbolic metric of syntactic similarity

Step 3: Assign each node a position vector

- ➢ Weighted sum of 1. the node's base vector and 2. the position vector of the node's mother
- ➢ Free parameter α
- ➢ Update rule from the Temporal Context Model of memory (Howard and Kahana, 2002)

position vector (X) =

α × base vector (X) + (1 − α) × position vector (X's mother)

A vector-symbolic metric of syntactic similarity

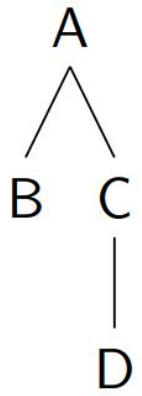
Example:

$$\text{position vec}(A) = \text{base vec}(A)$$

$$\text{position vec}(B) = \alpha \times \text{base vec}(B) + (1 - \alpha) \times \text{position vec}(A)$$

$$\text{position vec}(C) = \alpha \times \text{base vec}(C) + (1 - \alpha) \times \text{position vec}(A)$$

$$\text{position vec}(D) = \alpha \times \text{base vec}(D) + (1 - \alpha) \times \text{position vec}(C)$$



… but the position vector of C contains the position vector of A!

A vector-symbolic metric of syntactic similarity

Upshot:

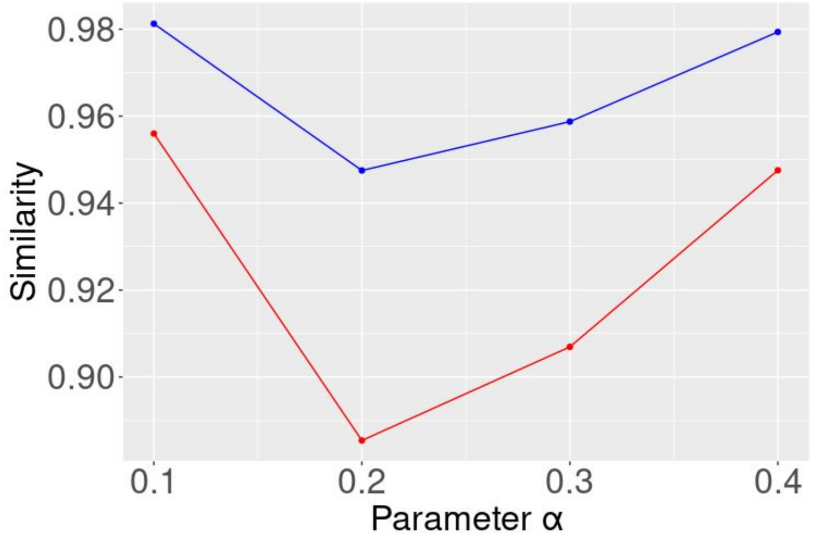
- ➢ Position vector encodes category information of current node and all dominating nodes.
- ➢ More distant nodes make up smaller part of representation.

Let’s apply this metric to model agreement attraction! Remember:

higher cosine similarity → higher likelihood of interference

Results (I)

(3) The **helicopter** for the **flight** over the **canyon** is rusty.

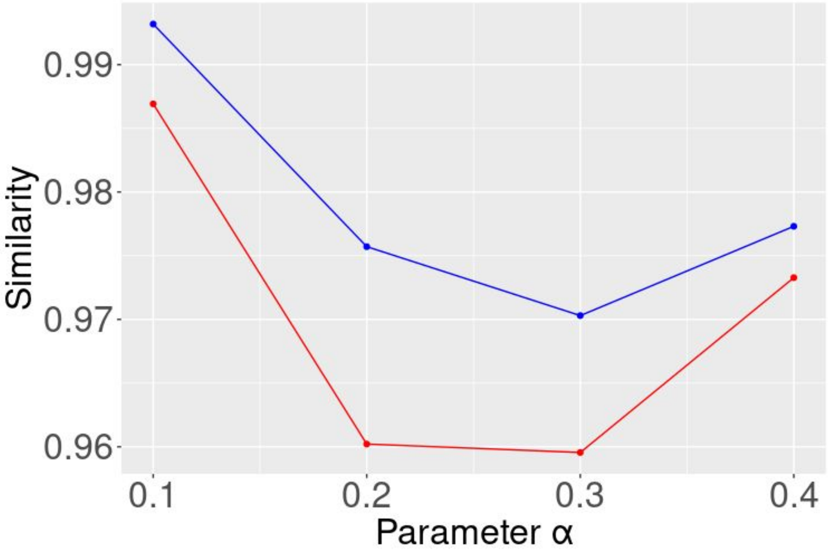


- ➢ Empirical observation: interference from ‘flight’ > interference from ‘canyon’
- ➢ Model prediction: similarity of ‘flight’ > similarity of ‘canyon’

Results (II)

(4a) The **editor** of the **book** was tired.

(4b) The **editor** who rejected the **book** was tired.



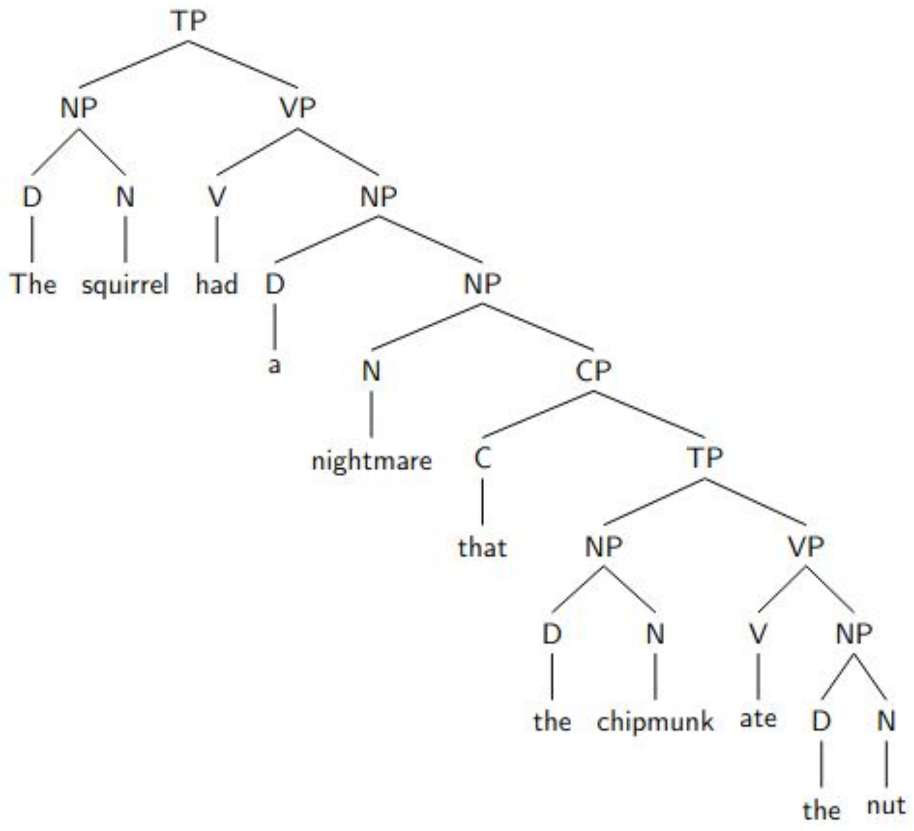
- ➢ Empirical observation: interference from P complement (4a) > interference from RC object (4b)
- ➢ Model prediction: similarity of P complement (4a) > similarity of RC object (4b)

Results (III)

(5) The **squirrel** had a **nightmare** that the **chipmunk** ate the **nut**.

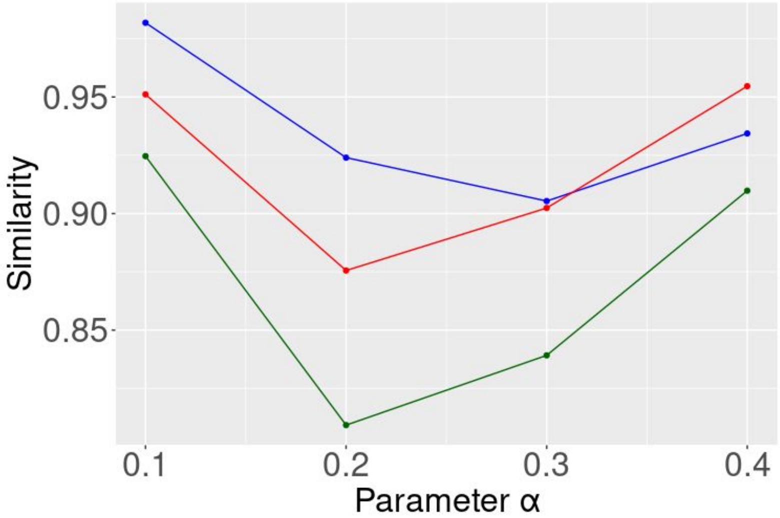
➢ 'Nightmare' is structurally closer but 'chipmunk' is also inside a subject

➢ Which one is more similar to target?



Results (III)

(5) The **squirrel** had a **nightmare** that the **chipmunk** ate the **nut**.



➢ Model prediction: for low α, similarity of 'nightmare' > similarity of 'chipmunk'; for high α, the opposite

position vector (X) = α × base vector (X) + (1 − α) × position vector (X's mother)

Discussion

Experimentally…

- ➢ Effects observed so far are captured
- ➢ Many concrete and testable predictions

Computationally…

- ➢ Different potential modifications of the model (e.g., dependency parse, non-orthogonal base vectors, probabilistic base vector assignment…)

Theoretically…

- ➢ Support for the view that agreement attraction is due to the way that sentences are encoded in working memory
- ➢ Grammatical positions are encoded relationally (no ‘subject feature’)

Thank you!

Many thanks to Samuel Amouyal, Mandy Cartner, Niki Koesterich, Aya Meltzer-Asscher and Stephanie Rich

This research was supported by the National Science Foundation (NSF BCS-1941485 to BD) and the Binational Science Foundation (NSF-BSF 2146798 to BD)