When to drop “that”: Revisiting the UID hypothesis on complementizer-dropping in English

Penny Pan & Alex Warstadt
UC San Diego

CAMP 8, November 15, 2025

My boss confirmed that we were absolutely crazy.

My boss confirmed that we were absolutely crazy.

My boss confirmed we were absolutely crazy.

complementizer

My boss confirmed [$_{CC}$ that we were absolutely crazy].

My boss confirmed [$_{CC}$ we were absolutely crazy].

complement clause

→ My boss confirmed [$_{CC}$ that we were absolutely crazy].

My boss confirmed [$_{CC}$ we were absolutely crazy].

→ My boss confirmed [$_{cc}$ that we were absolutely crazy].

My boss confirmed [$_{cc}$ we were absolutely crazy].

I think [$_{cc}$ that we were absolutely crazy].

I think [$_{cc}$ we were absolutely crazy].

→ My boss confirmed [$_{cc}$ that we were absolutely crazy].

My boss confirmed [$_{cc}$ we were absolutely crazy].

I think [$_{cc}$ that we were absolutely crazy].

→ I think [$_{cc}$ we were absolutely crazy].

Uniform Information Density (UID)

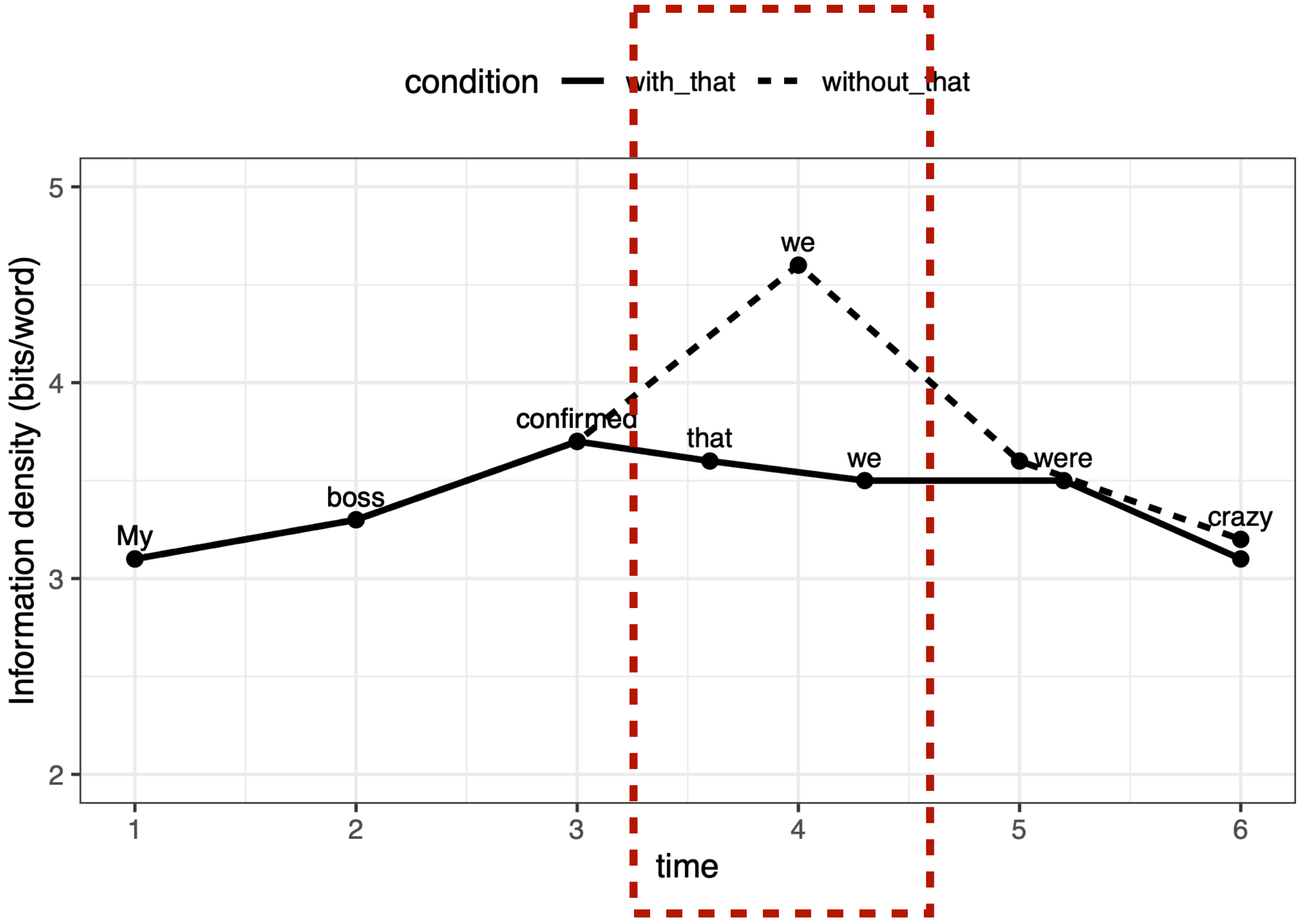
Levy & Jaeger (2007), Jaeger (2011)

Uniform Information Density (UID)

speakers prefer an even rate of information when uttering a sentence.

Levy & Jaeger (2007), Jaeger (2011)

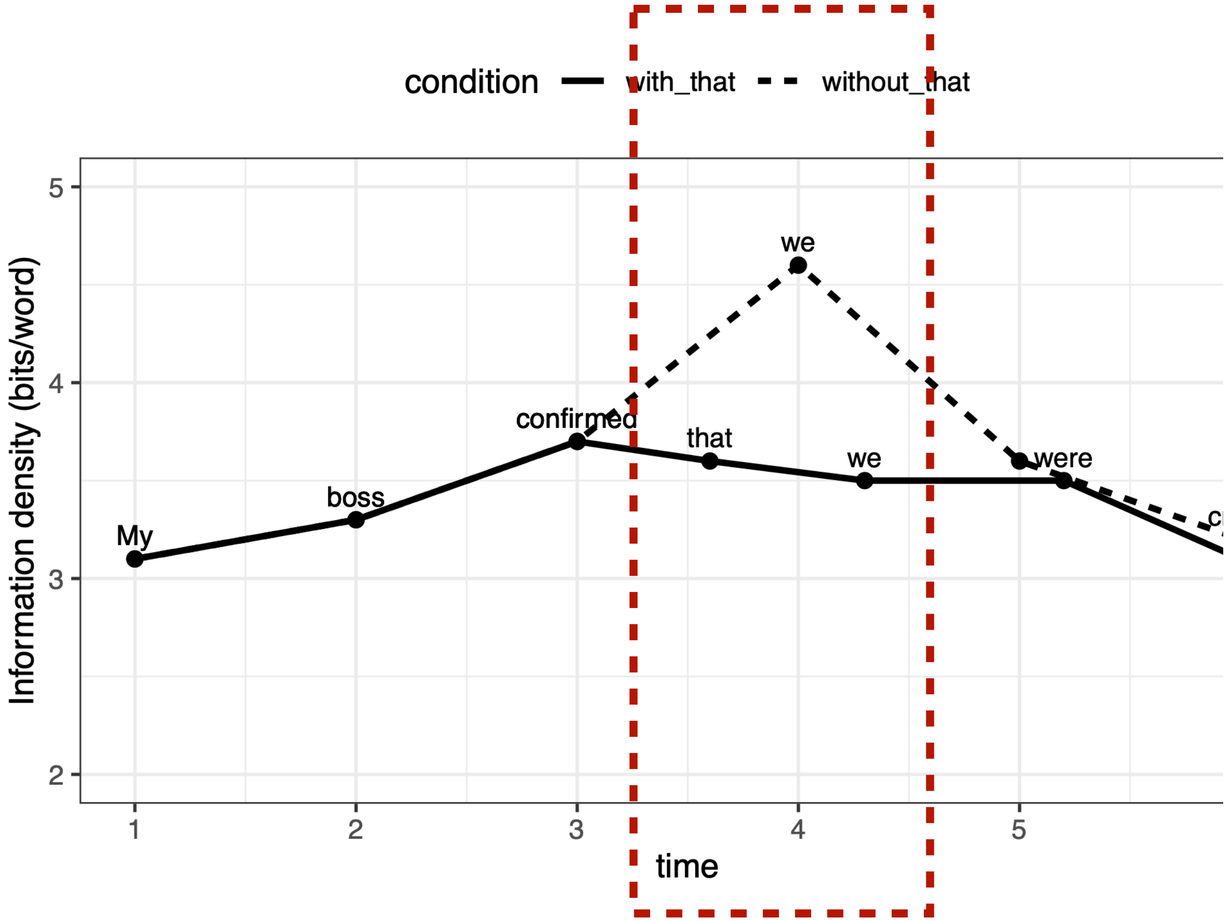
Uniform Information Density (UID)

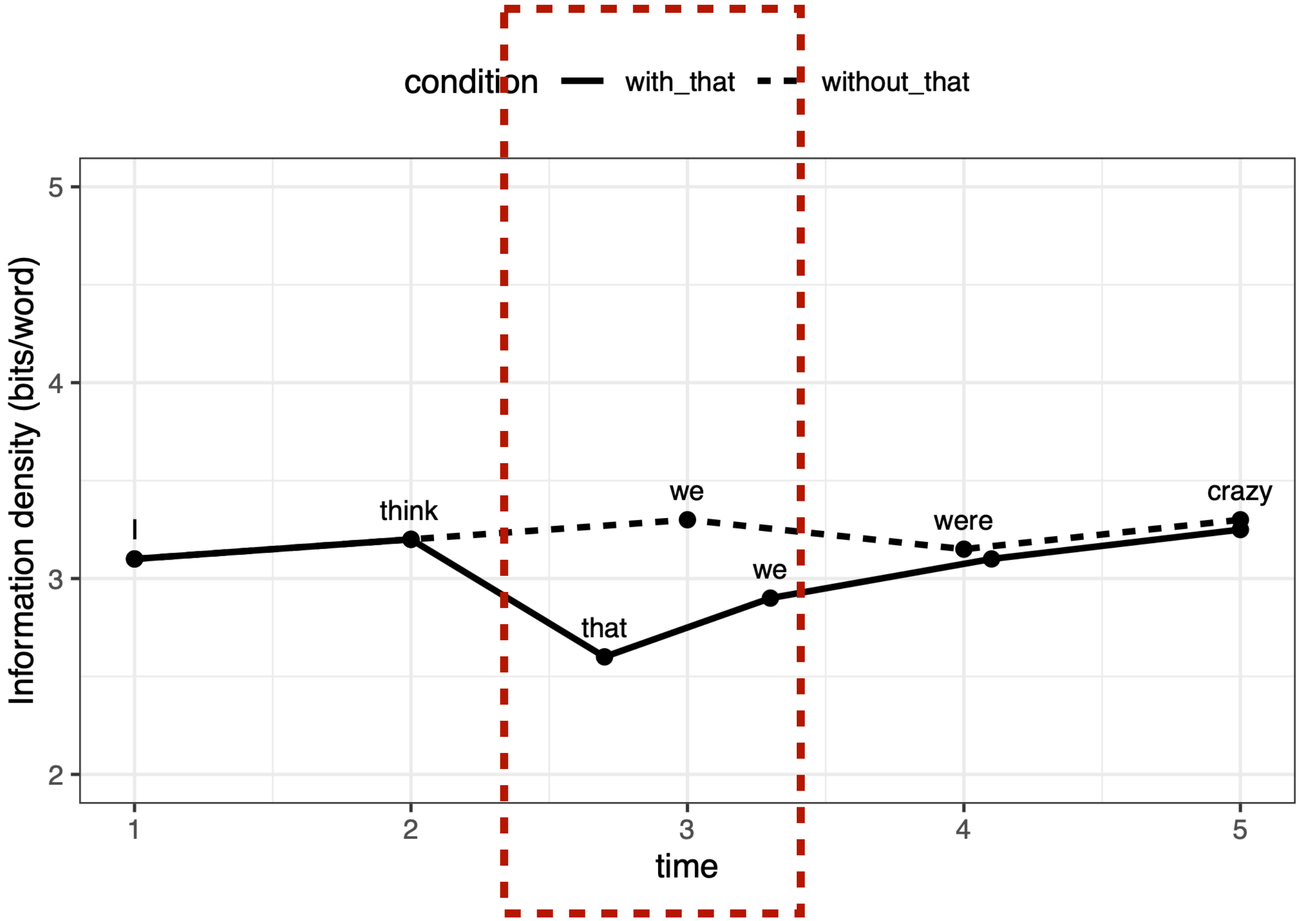


speakers prefer an even rate of information when uttering a sentence.

Levy & Jaeger (2007), Jaeger (2011)

Uniform Information Density (UID)





speakers prefer an even rate of information when uttering a sentence.

Levy & Jaeger (2007), Jaeger (2011)

Non-contextual informativity of the CC

Jaeger (2011)

Non-contextual informativity of the CC

- Subcategorization probability of the verb

$$I(CC \mid V_k)$$

Jaeger (2011)

Non-contextual informativity of the CC

- Subcategorization probability of the verb: how likely the verb takes a complement clause.

$$I(CC \mid V_k) \approx -\log(\frac{\#CC, V_k}{\sum_i \#V_k})$$

Jaeger (2011)

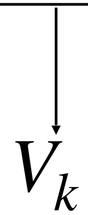
Non-contextual informativity of the CC

- Subcategorization probability of the verb: how likely the verb takes a complement clause.

$$I(CC \mid V_k) \approx -\log(\frac{\#CC, V_k}{\sum_i \#V_k})$$

My boss confirmed [**cc** we were absolutely crazy].

I think [**cc** we were absolutely crazy].



Jaeger (2011)

Non-contextual informativity of the CC

- Subcategorization probability of the verb: how likely the verb takes a complement clause.

$$I(CC \mid V_k) \approx -\log(\frac{\#CC, V_k}{\sum_i \#V_k})$$

My boss confirmed [**cc** we were absolutely crazy].

I think [**cc** we were absolutely crazy].

V_k

$$I(CC \mid think) < I(CC \mid confirm)$$

Jaeger (2011)

Non-contextual informativity of the CC

- Subcategorization probability of the verb: how likely the verb takes a complement clause.

$$I(CC \mid V_k) \approx -\log(\frac{\#CC, V_k}{\sum_i \#V_k})$$

My boss confirmed [**cc** that we were absolutely crazy].

I think [**cc** we were absolutely crazy].

V_k

$$I(CC \mid think) < I(CC \mid confirm)$$

Jaeger (2011)

Contextual informativity of w_1

Levy & Jaeger (2007), Rabinovich (2024)

Contextual informativity of w_1

- Surprisal of w_1

$$I(w_1 \mid context_k, V_k)$$

Levy & Jaeger (2007), Rabinovich (2024)

Contextual informativity of w_1

- Surprisal of w_1: the predictability of the first word w_1 in the complement clause given the context and the verb. $I(w_1 \mid context_k, V_k) = -\log P(w_1 \mid context_k, V_k)$

Levy & Jaeger (2007), Rabinovich (2024)

Contextual informativity of w_1

- Surprisal of w_1: the predictability of the first word w_1 in the complement clause given the context and the verb. $I(w_1 \mid context_k, V_k) = -\log P(w_1 \mid context_k, V_k)$

My boss confirmed [$_{CC}$ **we** were absolutely crazy].

I think [$_{CC}$ **we** were absolutely crazy].

w_1

Levy & Jaeger (2007), Rabinovich (2024)

Contextual informativity of w_1

- Surprisal of w_1: the predictability of the first word w_1 in the complement clause given the context and the verb. $I(w_1 \mid context_k, V_k) = -\log P(w_1 \mid context_k, V_k)$

My boss confirmed [cc **we** were absolutely crazy].

I think [cc **we** were absolutely crazy].

$context_k$ w_1

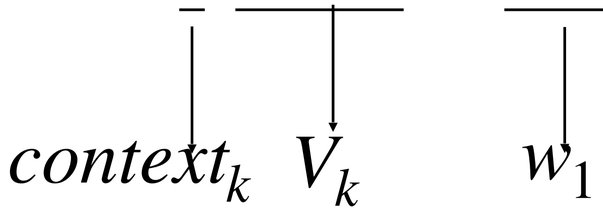
Levy & Jaeger (2007), Rabinovich (2024)

Contextual informativity of w_1

- Surprisal of w_1: the predictability of the first word w_1 in the complement clause given the context and the verb. $I(w_1 \mid context_k, V_k) = -\log P(w_1 \mid context_k, V_k)$

My boss confirmed [cc **we** were absolutely crazy].

I think [cc **we** were absolutely crazy].



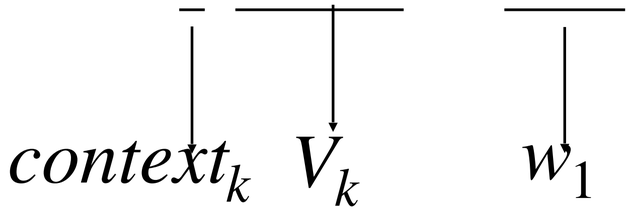
Levy & Jaeger (2007), Rabinovich (2024)

Contextual informativity of w_1

- Surprisal of w_1: the predictability of the first word w_1 in the complement clause given the context and the verb. $I(w_1 \mid context_k, V_k) = -\log P(w_1 \mid context_k, V_k)$

My boss confirmed [cc **we** were absolutely crazy].

I think [cc **we** were absolutely crazy].



$$I(we \mid Ithink) < I(we \mid Mybossconfirmed)$$

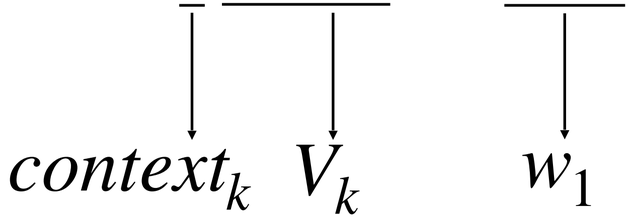
Levy & Jaeger (2007), Rabinovich (2024)

Contextual informativity of w_1

- Surprisal of w_1: the predictability of the first word w_1 in the complement clause given the context and the verb. $I(w_1 \mid context_k, V_k) = -\log P(w_1 \mid context_k, V_k)$

My boss confirmed [cc that **we** were absolutely crazy].

I think [cc **we** were absolutely crazy].



$$I(we \mid Ithink) < I(we \mid Mybossconfirmed)$$

Levy & Jaeger (2007), Rabinovich (2024)

Expected contextual informativity

Cohen Priva (2015)

Expected contextual informativity

- The amount of information that a linguistic element σ usually has in a language, across all contexts that it occurs in.

$$\mathbb{E}_j[I(\sigma \mid context_j)]$$

Cohen Priva (2015)

Expected contextual informativity of the verb

- Global contextual information of V_k: The amount of information that a verb usually has in a language, across all contexts that it occurs in.

$$\mathbb{E}_j[I(V_k \mid context_j)]$$

My boss confirmed [$_{CC}$ that we were absolutely crazy].

I think [we were absolutely crazy].

Thompson & Mulac (1991)

Expected contextual informativity of the verb

- Global contextual information of V_k: The amount of information that a verb usually has in a language, across all contexts that it occurs in.

$$\mathbb{E}_j[I(V_k \mid context_j)]$$

My boss confirmed [$_{CC}$ that we were absolutely crazy].

I think [we were absolutely crazy].

I think it's time for bed.
I thought it's a good idea.
…

"think" is usually predictable across its contexts

Thompson & Mulac (1991)

Expected contextual informativity of the verb

- Global contextual information of V_k: The amount of information that a verb usually has in a language, across all contexts that it occurs in.

$$\mathbb{E}_j[I(V_k \mid context_j)]$$

My boss confirmed [cc that we were absolutely crazy].

I think [we were absolutely crazy].

I think it's time for bed.
I thought it's a good idea.
…

"think" is usually predictable across its contexts

$$\mathbb{E}_j[I(think \mid context_j)] < \mathbb{E}_j[I(confirm \mid context_j)]$$

Thompson & Mulac (1991)

Expected contextual informativity of the verb

- Global contextual information of V_k: The amount of information that a verb usually has in a language, across all contexts that it occurs in.

$$\mathbb{E}_j[I(V_k \mid context_j)]$$

My boss confirmed [$_{cc}$ that we were absolutely crazy].

I think [we were absolutely crazy].

I think it's time for bed.
I thought it's a good idea.
…

"think" is usually predictable across its contexts

"grammaticalization"

$$\mathbb{E}_j[I(think \mid context_j)] < \mathbb{E}_j[I(confirm \mid context_j)]$$

Thompson & Mulac (1991)

Non-contextual informativity of the CC

subcategorization probability of the verb

Contextual informativity of w_1

surprisal of w_1

Expected contextual informativity of the verb

global information of V_k

Research questions

- How is the information density computed?
  - Is the information density computed online during processing or based on accumulated contextual experience or some non-contextual heuristics (e.g., frequency)?

Research questions

- How is the information density computed?
  - Is the information density computed online during processing or based on accumulated contextual experience or some non-contextual heuristics (e.g., frequency)?
- What is the context window for the uniform information density?
  - Is it before, at, or after the possible reduction site?

Non-contextual informativity of the CC

subcategorization probability of the verb

Contextual informativity of w_1

surprisal of w_1

Expected contextual informativity of the verb

global information of V_k

My boss confirmed [$_{CC}$ we were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity		subcategorization probability of the verb	
local contextual informativity			surprisal of w_1
expected contextual informativity	global information of V_k		

My boss confirmed [$_{CC}$ we were absolutely crazy].

How ↓

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity		subcategorization probability of the verb	
local contextual informativity			surprisal of w_1
expected contextual informativity	global information of V_k		

My boss confirmed [CC we were absolutely crazy].

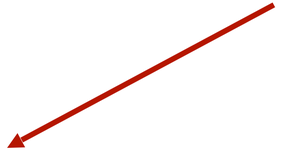
Where →	**verb** (V_k)	**complement clause** (CC)	**first word in CC** (w_1)
non-contextual informativity		subcategorization probability of the verb	
local contextual informativity			surprisal of w_1
expected contextual informativity	global information of V_k		

My boss confirmed [cc we were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity	← non-contextual heuristics	subcategorization probability of the verb	
local contextual informativity			surprisal of w_1
expected contextual informativity	global information of V_k		

$xt_j, V_j)$

My boss confirmed [**cc** we were absolutely crazy].



	verb (V_k)	**complement clause (CC)**	first word in CC (w_1)
non-contextual informativity	← non-contextual heuristics	subcategorization probability of the verb	
local contextual informativity			surprisal of w_1
expected contextual informativity	global information of V_k		

$xt_j, V_j)$

My boss confirmed [CC **we** were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity		subcategorization probability of the verb	
local contextual informativity	← online processing based on the context		surprisal of w_1
expected contextual informativity	global information of V_k		

$i, V_j)$

My boss confirmed [CC **we** were absolutely crazy].

	verb (V_k)	complement clause (CC)	**first word in CC (w_1)**
non-contextual informativity		subcategorization probability of the verb	
local contextual informativity	← online processing based on the context		surprisal of w_1
expected contextual informativity	global information of V_k		

$i, V_j)$

My boss **confirmed** [CC we were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity		subcategorization probability of the verb	
local contextual informativity			surprisal of w_1
expected contextual informativity	global information of V_k		

← precomputed contextual information based on past experiences

My boss **confirmed** [CC we were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity		subcategorization probability of the verb	
local contextual informativity			surprisal of w_1
expected contextual informativity	global information of V_k		

← precomputed contextual information based on past experiences

$xt_j, V_j)$

Measures

My boss confirmed [$_{CC}$ we were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity		subcategorization probability of the verb	
local contextual informativity			surprisal of w_1
expected contextual informativity	global information of V_k		

Measures

My boss confirmed [$_{CC}$ we were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity	relative frequency of V_k compared to other verbs	subcategorization probability of the verb	relative frequency of w_1 compared to other words
local contextual informativity	surprisal of V_k	surprisal of CC	surprisal of w_1
expected contextual informativity	global information of V_k	global information of CC	global information of w_1

Measures

My boss confirmed [CC we were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity	$I(V_k)$	$I(CC \mid V_k)$	$I(w_1)$
local contextual informativity	$I(V_k \mid context_k)$	$I(CC \mid context_k, V_k)$	$I(w_1 \mid context_k, V_k)$
expected contextual informativity	$\mathbb{E}_j[I(V_k \mid context_j)]$	$\mathbb{E}_j[I(CC \mid context_j, V_k)]$	$\mathbb{E}_j[I(w_1 \mid context_j, V_k)]$

Measures

My boss confirmed [CC we were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity	$I(V_k) = -\log(\frac{\#V_k}{\sum_i \#V_i})$	$I(CC \mid V_k) \approx -\log(\frac{\#CC, V_k}{\sum_i \#V_k})$	$I(w_1) = -\log(\frac{\#w_1}{\sum_i \#w_i})$
local contextual informativity	$I(V_k \mid context_k) = -\log P(V_k \mid context_k)$	$I(CC \mid context_k, V_k) = -\log P(CC \mid context_k, V_k)$	$I(w_1 \mid context_k, V_k) = -\log P(w_1 \mid context_k, V_k)$
expected contextual informativity	$\mathbb{E}_j[I(V_k \mid context_j)] \approx \frac{1}{\#V_k}\sum_{j:V_j=V_k} I(V_j \mid context_j)$	$\mathbb{E}_j[I(CC \mid context_j, V_k)] \approx \frac{1}{\#V_k}\sum_{j:V_j=V_k} I(CC \mid context_j, V_j)$	$\mathbb{E}_j[I(w_1 \mid context_j, V_k)] \approx \frac{1}{\#V_k}\sum_{j:V_j=V_k} I(w_1 \mid context_j, V_j)$

Measures

My boss confirmed [CC we were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity	$I(V_k) = -\log(\frac{\#V_k}{\sum_i \#V_i})$	$I(CC \mid V_k) \approx -\log(\frac{\#CC, V_k}{\sum_i \#V_k})$	$I(w_1) = -\log(\frac{\#w_1}{\sum_i \#w_i})$
local contextual informativity	$I(V_k \mid context_k) = -\log P(V_k \mid context_k)$	$I(CC \mid context_k, V_k) = -\log P(CC \mid context_k, V_k)$	$I(w_1 \mid context_k, V_k) = -\log P(w_1 \mid context_k, V_k)$
expected contextual informativity	$\mathbb{E}_j[I(V_k \mid context_j)] \approx \frac{1}{\#V_k} \sum_{j:V_j=V_k} I(V_j \mid context_j)$	$\mathbb{E}_j[I(CC \mid context_j, V_k)] \approx \frac{1}{\#V_k} \sum_{j:V_j=V_k} I(CC \mid context_j, V_j)$	$\mathbb{E}_j[I(w_1 \mid context_j, V_k)] \approx \frac{1}{\#V_k} \sum_{j:V_j=V_k} I(w_1 \mid context_j, V_j)$

Measures

My boss confirmed [CC we were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity	$I(V_k) = -\log(\frac{\#V_k}{\sum_i \#V_i})$	$I(CC \mid V_k) \approx -\log(\frac{\#CC, V_k}{\sum_i \#V_k})$	$I(w_1) = -\log(\frac{\#w_1}{\sum_i \#w_i})$
local contextual informativity	$I(V_k \mid context_k) = -\log P(V_k \mid context_k)$	$I(CC \mid context_k, V_k) = -\log P(CC \mid context_k, V_k)$	$I(w_1 \mid context_k, V_k) = -\log P(w_1 \mid context_k, V_k)$
expected contextual informativity	$\mathbb{E}_j[I(V_k \mid context_j)] \approx \frac{1}{\#V_k} \sum_{j:V_j=V_k} I(V_j \mid context_j)$	$\mathbb{E}_j[I(CC \mid context_j, V_k)] \approx \frac{1}{\#V_k} \sum_{j:V_j=V_k} I(CC \mid context_j, V_j)$	$\mathbb{E}_j[I(w_1 \mid context_j, V_k)] \approx \frac{1}{\#V_k} \sum_{j:V_j=V_k} I(w_1 \mid context_j, V_j)$

Measures

My boss confirmed [CC we were absolutely crazy].

	verb (V_k)	complement clause (CC)	first word in CC (w_1)
non-contextual informativity	$I(V_k) = -\log(\frac{\#V_k}{\sum_i \#V_i})$	$I(CC \mid V_k) \approx -\log(\frac{\#CC, V_k}{\sum_i \#V_k})$	$I(w_1) = -\log(\frac{\#w_1}{\sum_i \#w_i})$
local contextual informativity	$I(V_k \mid context_k) = -\log P(V_k \mid context_k)$	$I(CC \mid context_k, V_k) = -\log P(CC \mid context_k, V_k)$	$I(w_1 \mid context_k, V_k) = -\log P(w_1 \mid context_k, V_k)$
expected contextual informativity	$\mathbb{E}_j[I(V_k \mid context_j)] \approx \frac{1}{\#V_k}\sum_{j:V_j=V_k} I(V_j \mid context_j)$	$\mathbb{E}_j[I(CC \mid context_j, V_k)] \approx \frac{1}{\#V_k}\sum_{j:V_j=V_k} I(CC \mid context_j, V_j)$	$\mathbb{E}_j[I(w_1 \mid context_j, V_k)] \approx \frac{1}{\#V_k}\sum_{j:V_j=V_k} I(w_1 \mid context_j, V_j)$

Monte Carlo estimation

Corpus study

- 1,000,000 sentences from the dolma corpus (written)

Corpus study

- 1,000,000 sentences from the dolma corpus (written)

My boss confirmed that we were absolutely crazy.

I know him.

I don’t know when he will come.

He said “it is beautiful outside.”

I didn’t expect that it would be so good.

He told her it would take him hours to get there.

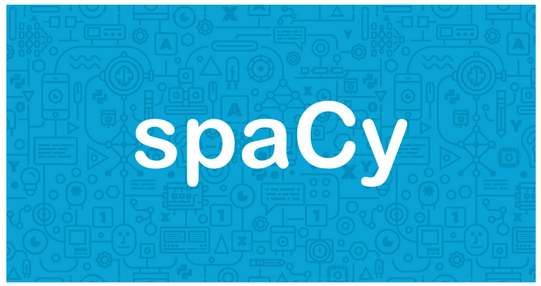
…

I think you do have to be in LA.

…

Corpus study

- 1,000,000 sentences from the dolma corpus (written)



My boss confirmed that we were absolutely crazy.

I know him.

I don’t know when he will come.

He said “it is beautiful outside.”

I didn’t expect that it would be so good.

He told her it would take him hours to get there.

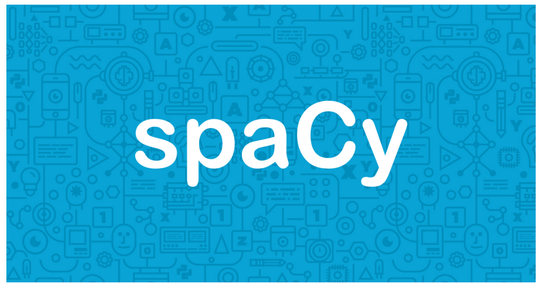
…

I think you do have to be in LA.

…

Corpus study

- 1,000,000 sentences from the dolma corpus (written)



verb + CC + (that)

I think you do have to be in LA.

My boss confirmed that we were absolutely crazy.

I didn’t expect that it would be so good.

He told her it would take him hours to get there.

…

others

I know him.

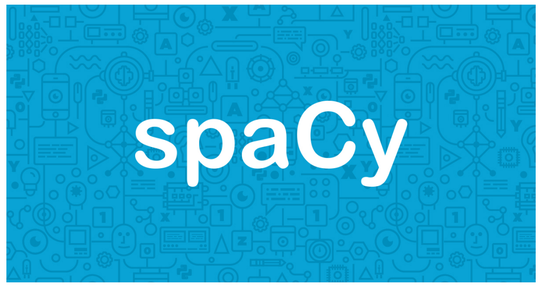
I don’t know when he will come.

He said “it is beautiful outside.”

…

Corpus study

- 1,000,000 sentences from the dolma corpus (written)



verb + CC + (that)

I think you do have to be in LA.

My boss confirmed that we were absolutely crazy.

I didn’t expect that it would be so good.

He told her it would take him hours to get there.

…

others

NP

I know him.

I don’t know when he will come.

embedded wh

He said “it is beautiful outside.”

quotation

…

Corpus study

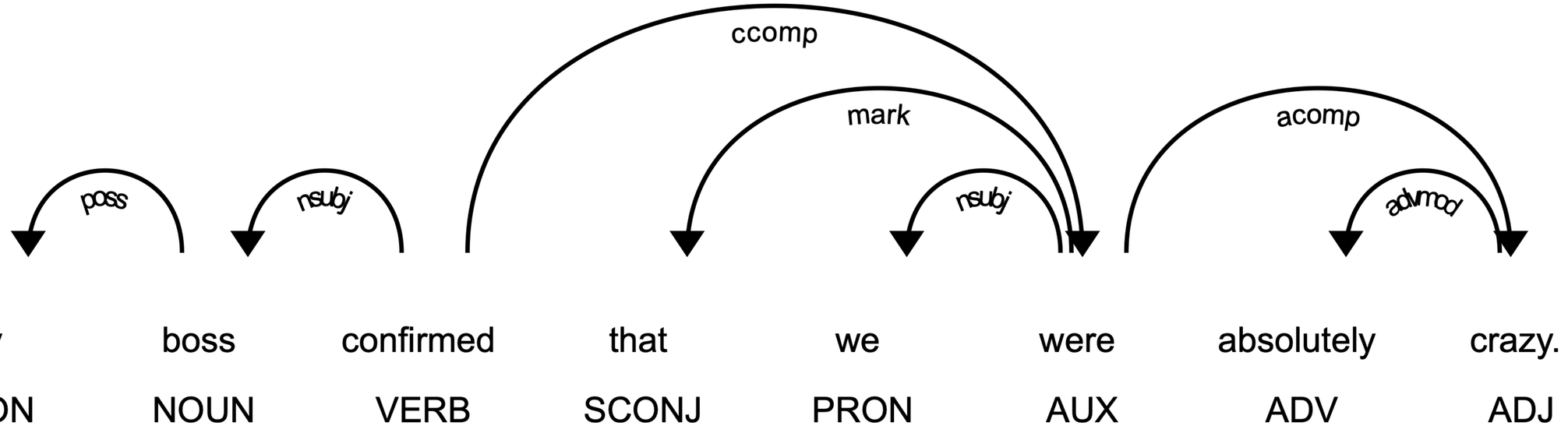
- 1,000,000 sentences from the dolma corpus (written)

My boss confirmed that we were absolutely crazy.

Corpus study

- 1,000,000 sentences from the dolma corpus (written)

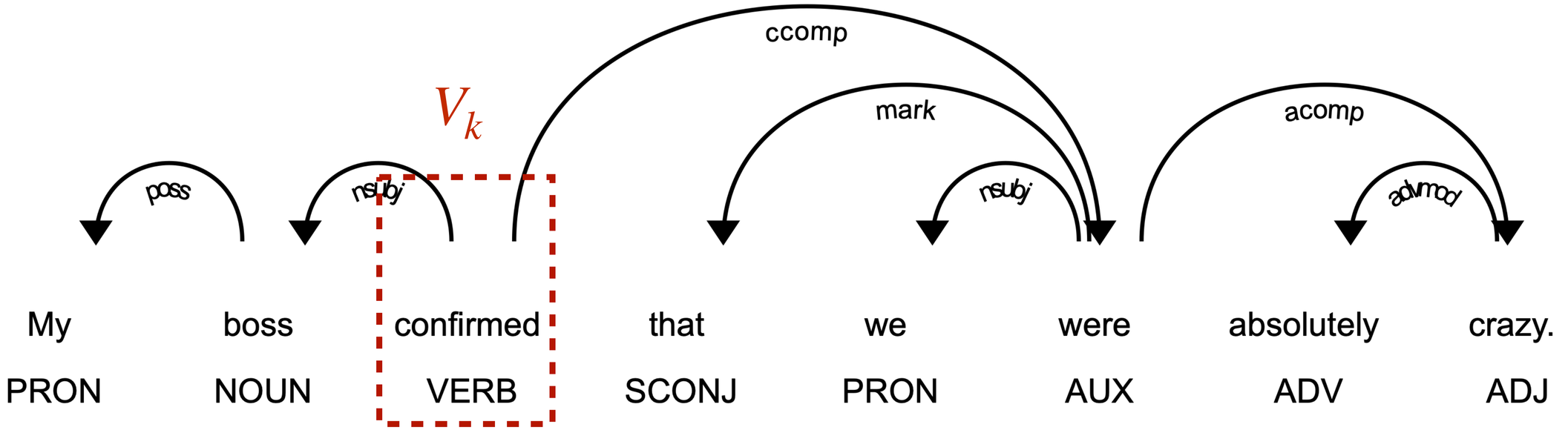
My boss confirmed that we were absolutely crazy.



Corpus study

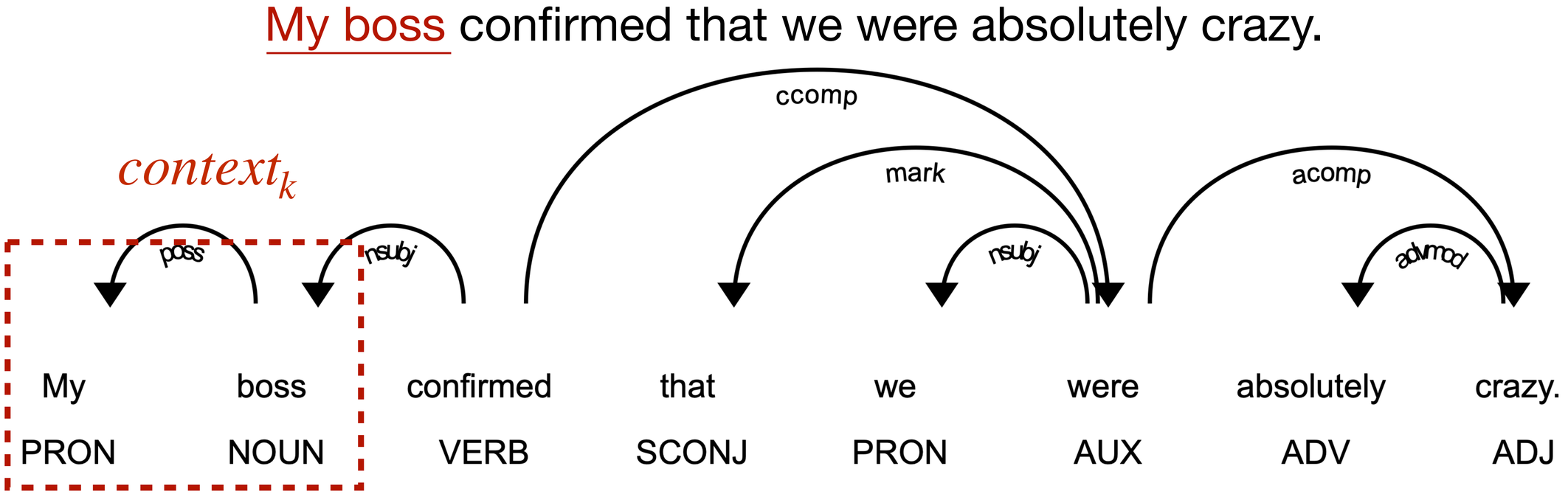
- 1,000,000 sentences from the dolma corpus (written)

My boss confirmed that we were absolutely crazy.



Corpus study

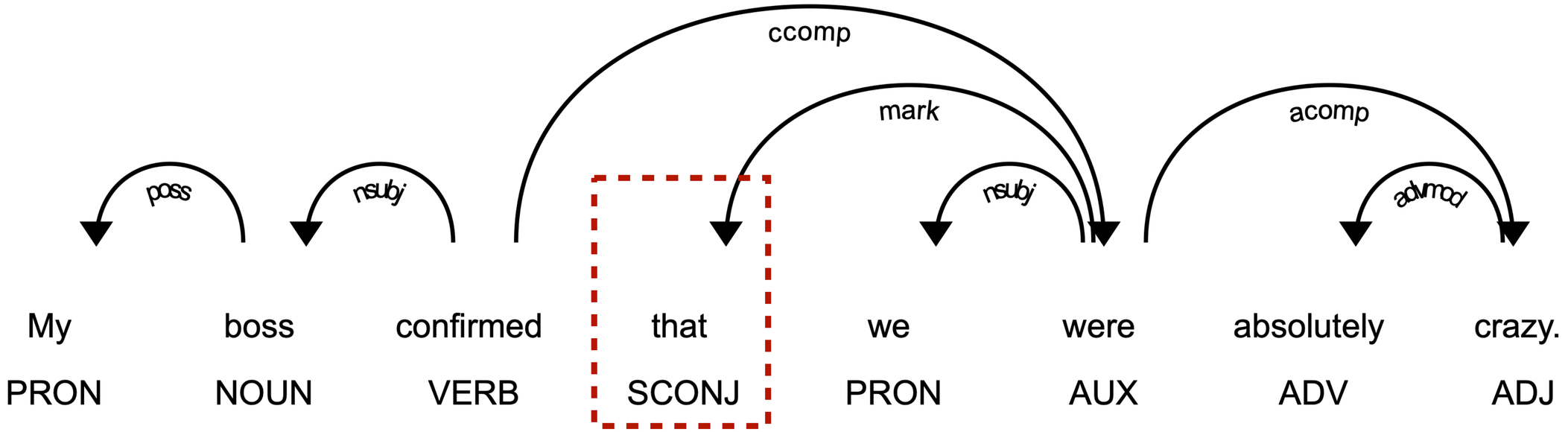
- 1,000,000 sentences from the dolma corpus (written)



Corpus study

- 1,000,000 sentences from the dolma corpus (written)

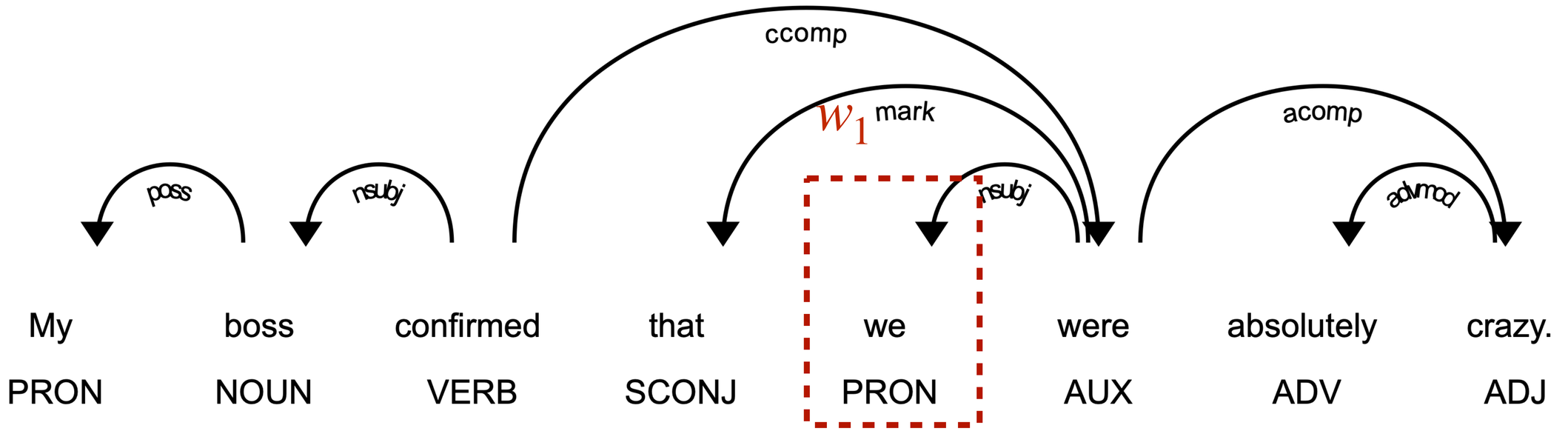
My boss confirmed that we were absolutely crazy.



Corpus study

- 1,000,000 sentences from the dolma corpus (written)

My boss confirmed that we were absolutely crazy.



Corpus study

- 1,000,000 sentences from the dolma corpus (written)
- 69,888 sentences with a CC, headed by 126 clause-embedding predicates of interest, and 26,888 of these sentences contain an overt complementizer (37.7%).

My boss confirmed that we were absolutely crazy.

Measuring local contextual informativity

My boss confirmed that we were absolutely crazy.

Measuring local contextual informativity

My boss confirmed we were absolutely crazy.

Measuring local contextual informativity

My boss confirmed we were absolutely crazy.

GPT-2

P(confirmed | My boss)

P(we | My boss confirmed)

Measuring local contextual informativity

My boss confirmed we were absolutely crazy.

GPT-2

Local informativity of the verb, $I(V_k \mid context_k)$

P(confirmed | My boss)

Local informativity of $w_1, I(w_1 \mid context_k, V_k)$

P(we | My boss confirmed)

Measuring expected contextual informativity

My boss confirmed we were absolutely crazy.

Local informativity of
the verb, $I(V_k \mid context_k)$

Local informativity of
w_1, $I(w_1 \mid context_k, V_k)$

Measuring expected contextual informativity

My boss confirmed we were absolutely crazy.

She confirms we need more chairs for the guests.
The waiter finally confirmed the reservation was made.
….

Local informativity of the verb, $I(V_k \mid context_k)$

Local informativity of w_1, $I(w_1 \mid context_k, V_k)$

Monte Carlo estimation

Expected informativity of the verb, $\mathbb{E}_j[I(V_k \mid context_j)]$

$$\approx \frac{1}{\#V_k} \sum_{j:V_j=V_k} I(w_1 \mid context_j)$$

Expected informativity of w_1, $\mathbb{E}_j[I(w_1 \mid context_j, V_k)]$

$$\approx \frac{1}{\#V_k} \sum_{j:V_j=V_k} I(V_j \mid context_j)$$

Analysis

Bayesian logistic regression

dropping of the complementizer ~ control variables + **main effects**

Analysis

Bayesian logistic regression

position of the verb, distance between verb and CC,
length of CC onset, length of the remaining CC

dropping of the complementizer ~ control variables + **main effects**

Analysis

Bayesian logistic regression

position of the verb, distance between verb and CC,
length of CC onset, length of the remaining CC

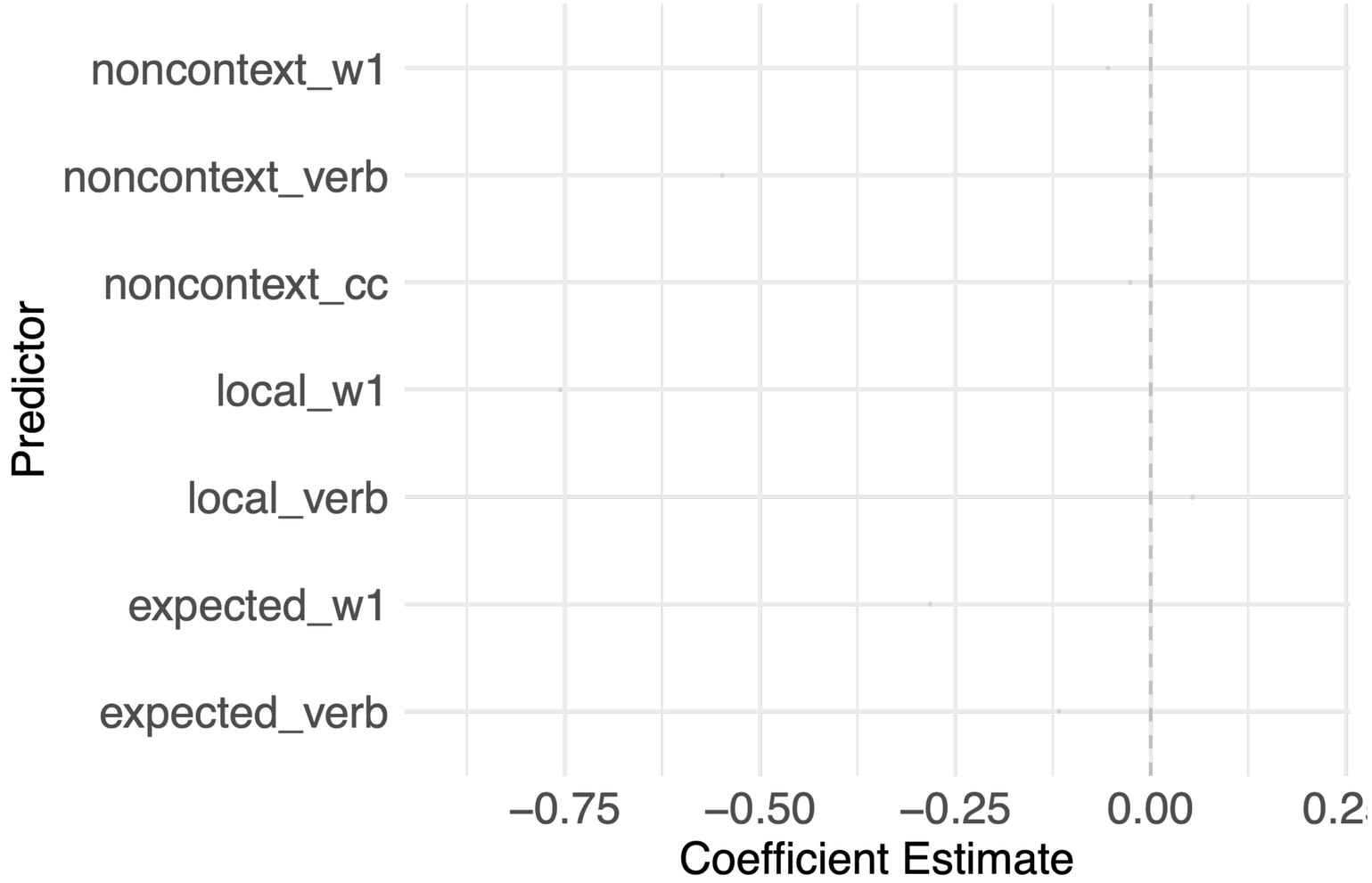
dropping of the complementizer ~ control variables + **main effects**

	V_k	CC	w_1
non-contextual	**noncontext_verb**	**noncontext_cc**	**noncontext_w1**
local contextual	**local_verb**		**local_w1**
expected	**expected_verb**		**expected_w1**

Results

My boss confirmed [CC we were absolutely crazy].

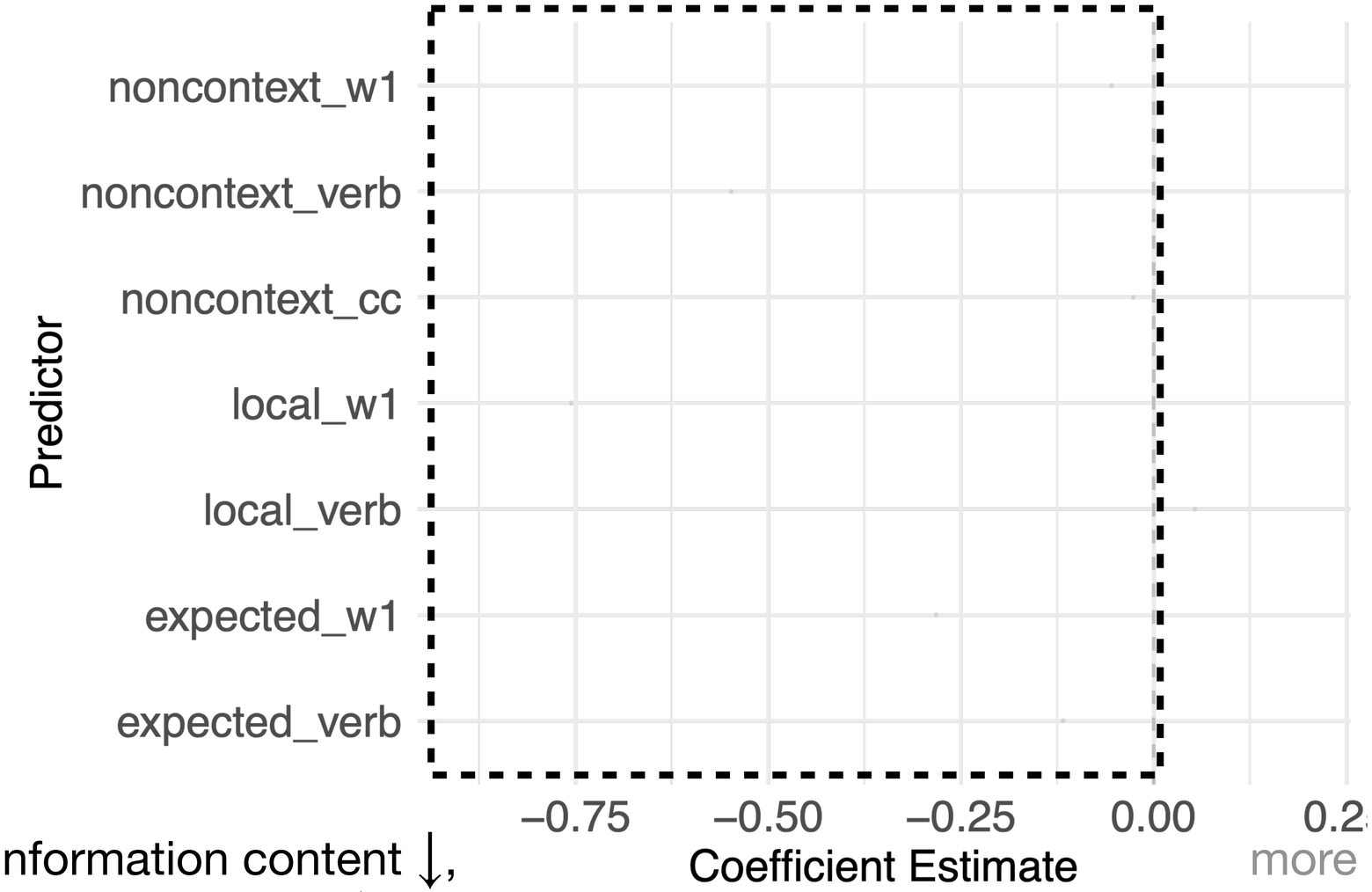
predicting the dropping of “that”



Results

My boss confirmed [CC we were absolutely crazy].

predicting the dropping of “that”



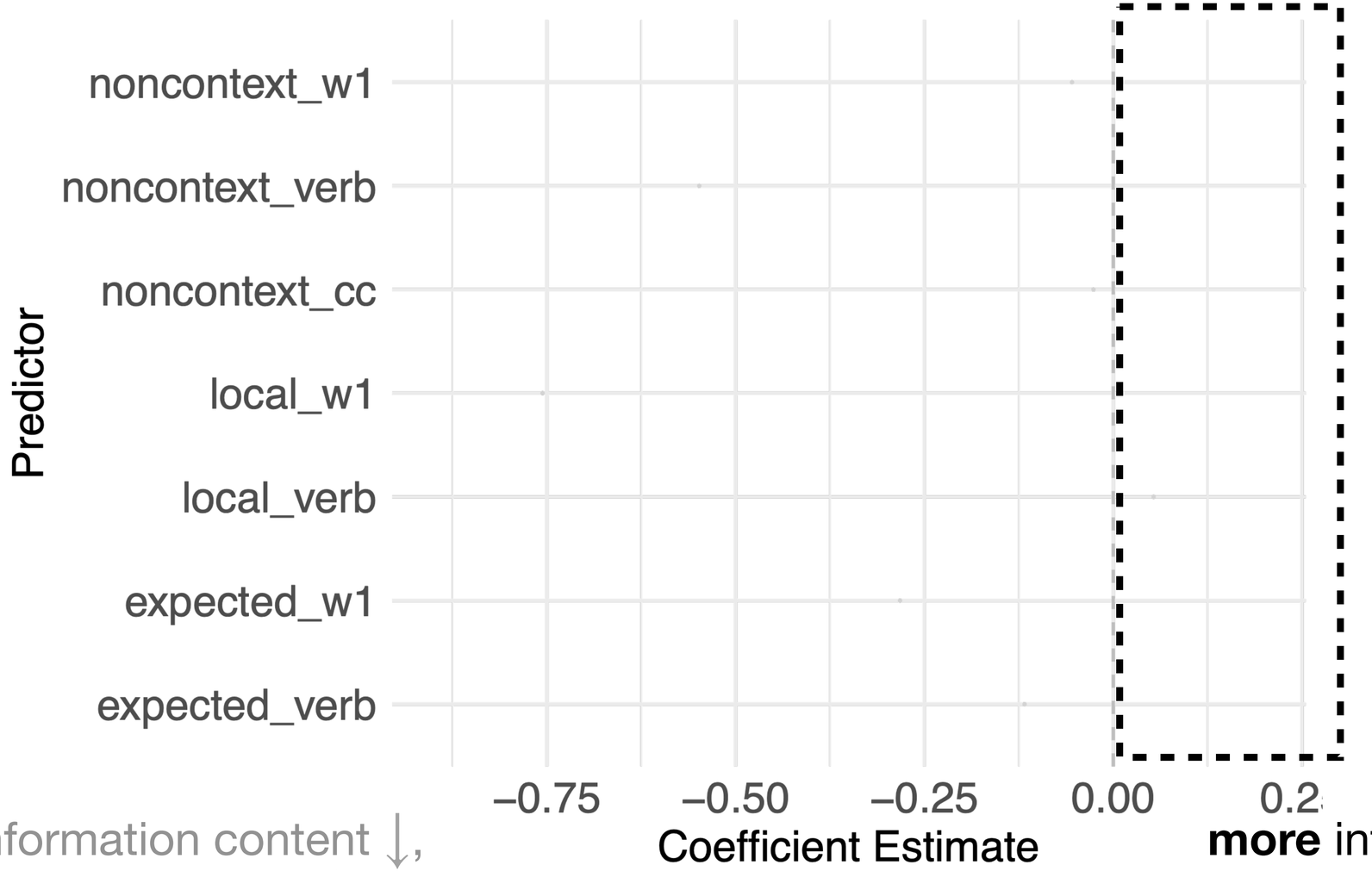
less information content ↓,
more dropping “that” ↑

more information content ↑,
more dropping “that” ↑

Results

My boss confirmed [CC we were absolutely crazy].

predicting the dropping of “that”



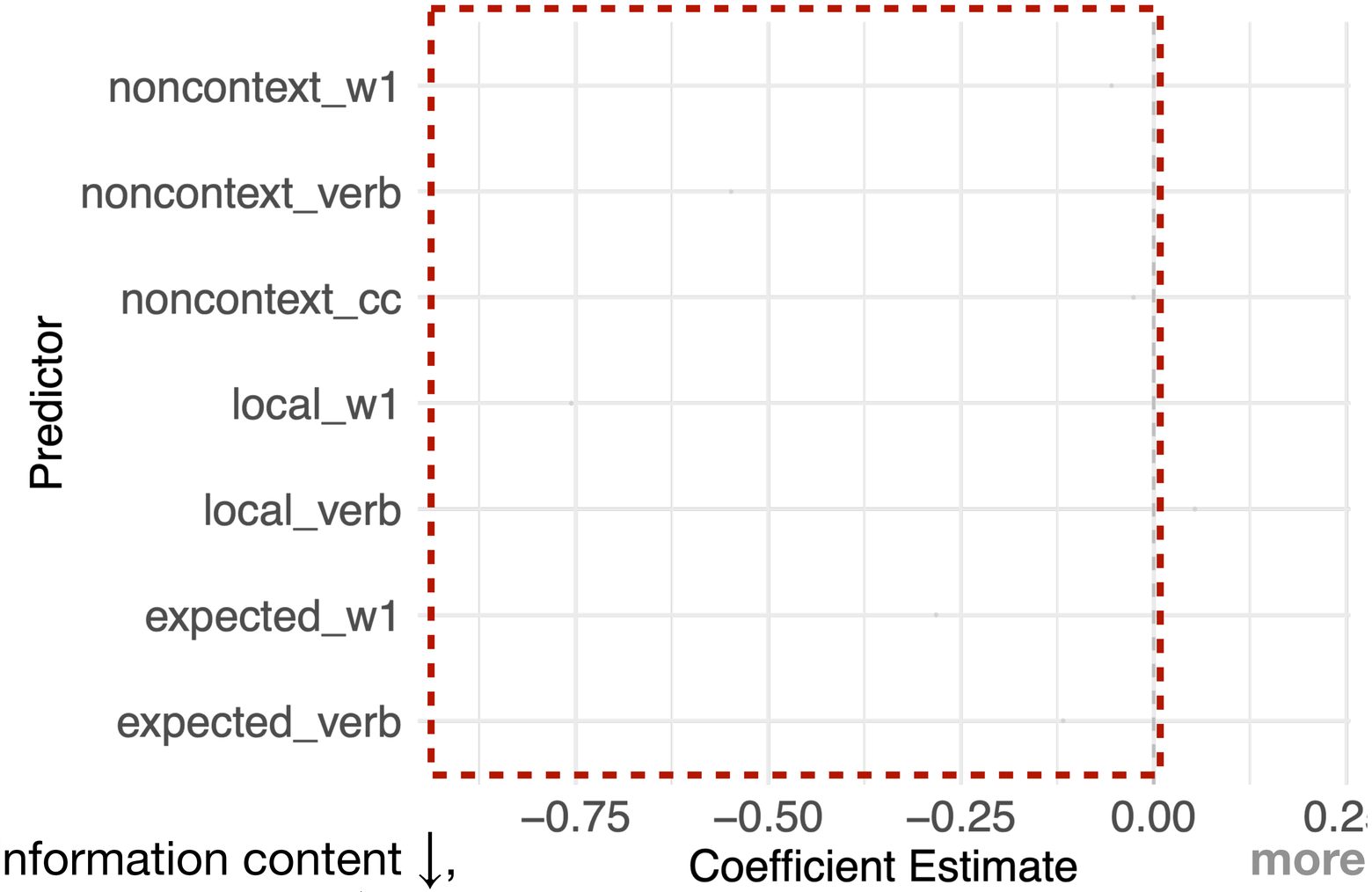
less information content ↓,
more dropping “that” ↑

more information content ↑,
more dropping “that” ↑

Results

My boss confirmed [CC we were absolutely crazy].

predicting the dropping of “that”



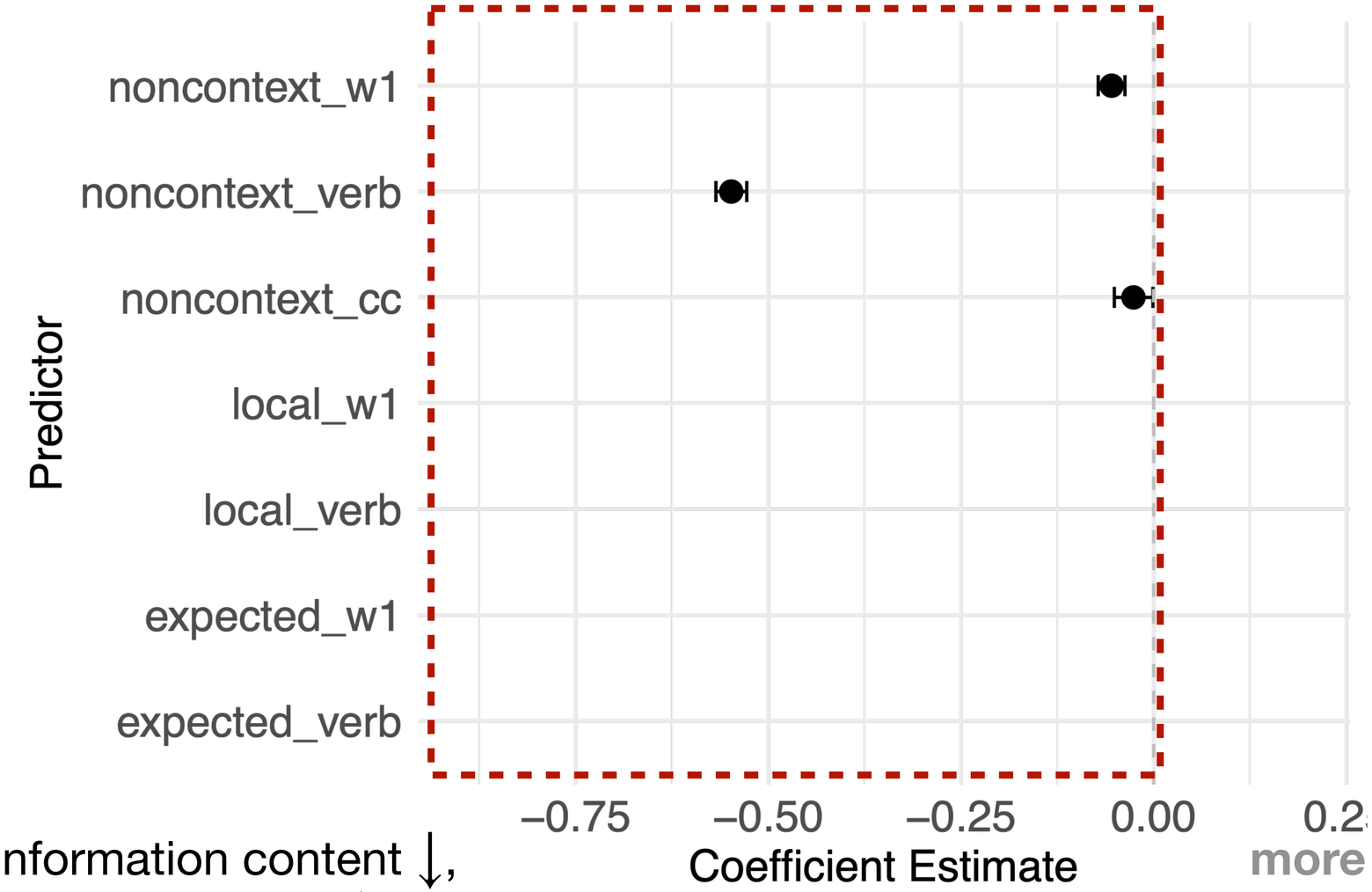
less information content ↓,
more dropping “that” ↑

more information content ↑,
more dropping “that” ↑

Results

My boss confirmed [CC we were absolutely crazy].

predicting the dropping of “that”



we

confirm

confirm + CC

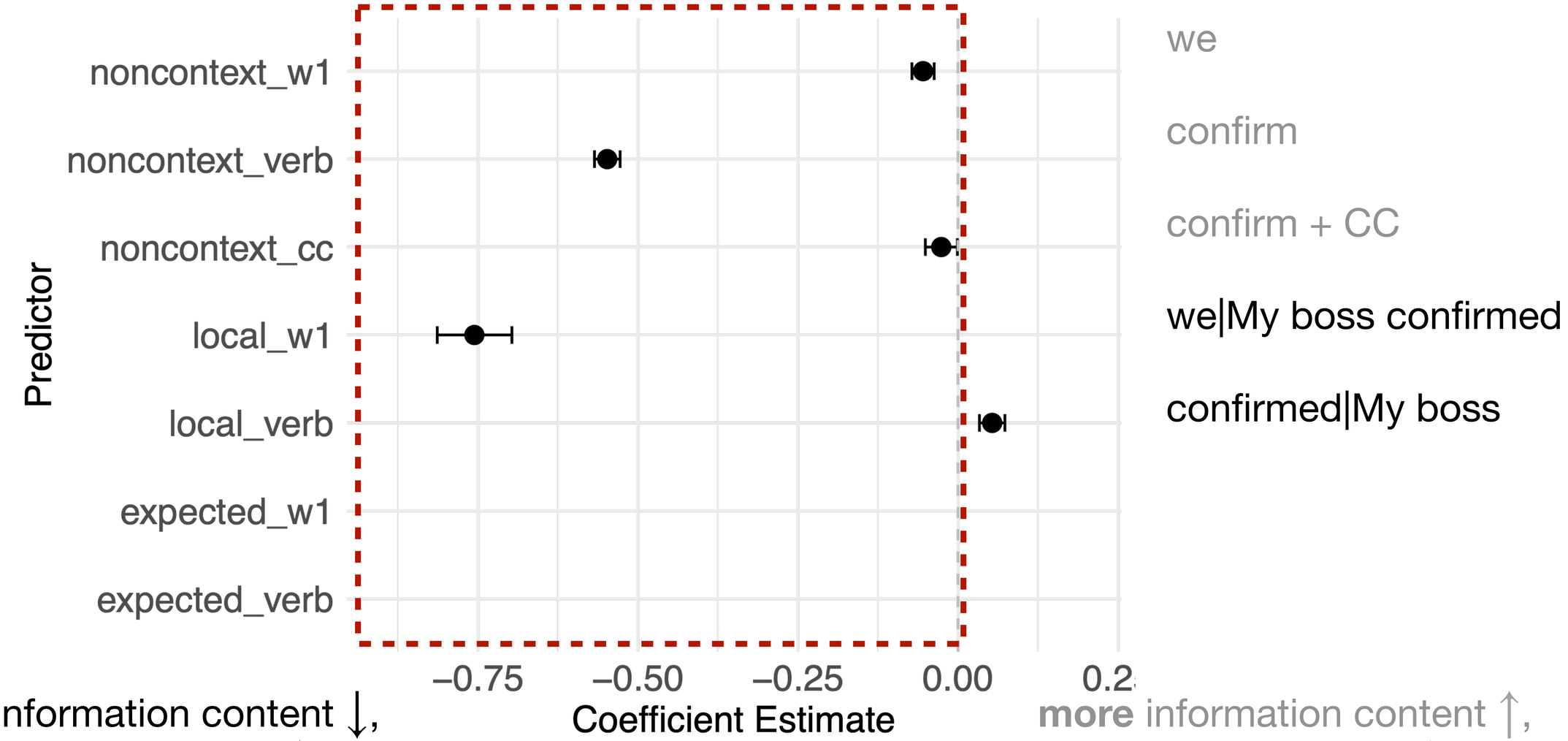
less information content ↓,
more dropping “that” ↑

more information content ↑,
more dropping “that” ↑

Results

My boss confirmed [CC we were absolutely crazy].

predicting the dropping of “that”



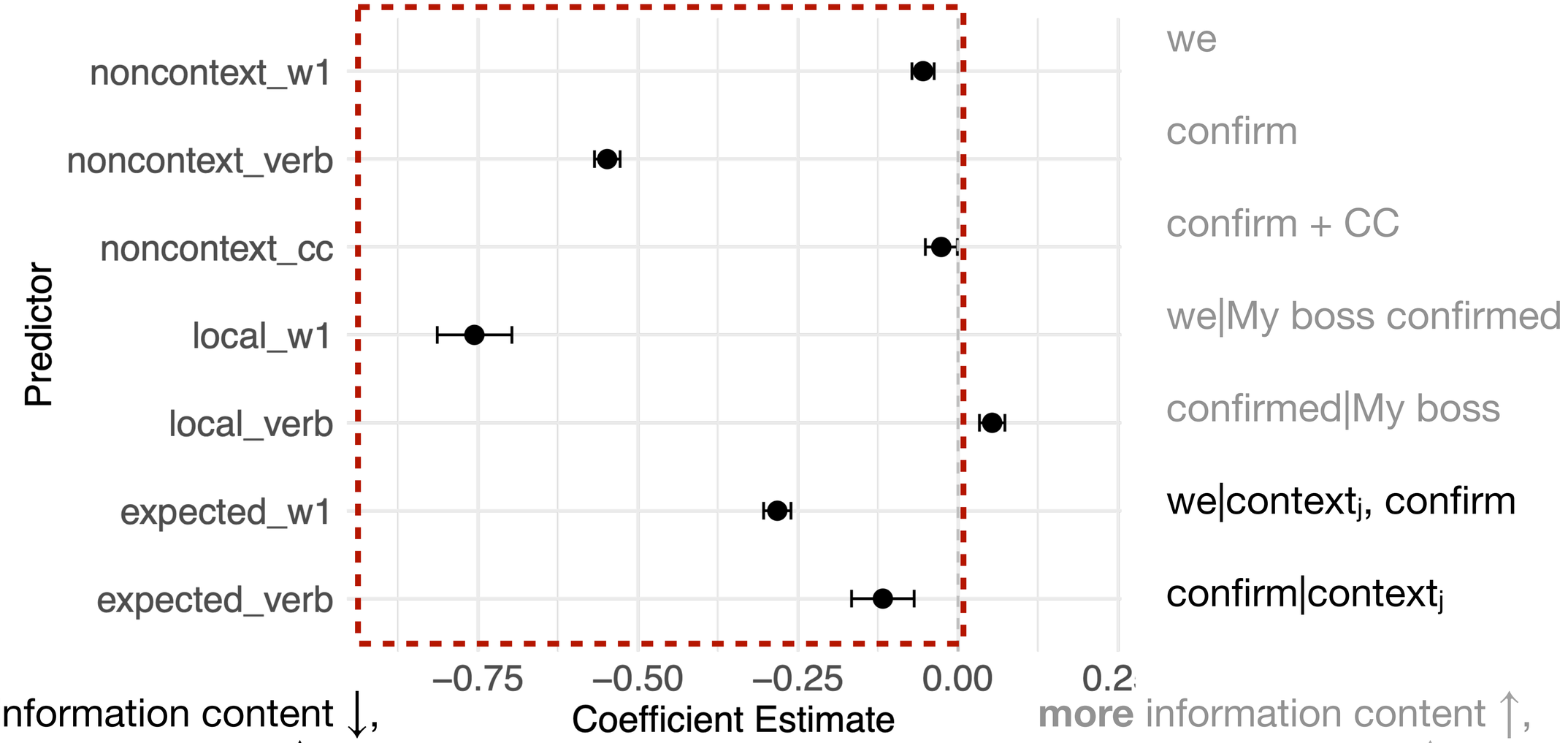
less information content ↓,
more dropping “that” ↑

more information content ↑,
more dropping “that” ↑

Results

My boss confirmed [$_{CC}$ we were absolutely crazy].

predicting the dropping of “that”



less information content ↓,
more dropping “that” ↑

more information content ↑,
more dropping “that” ↑

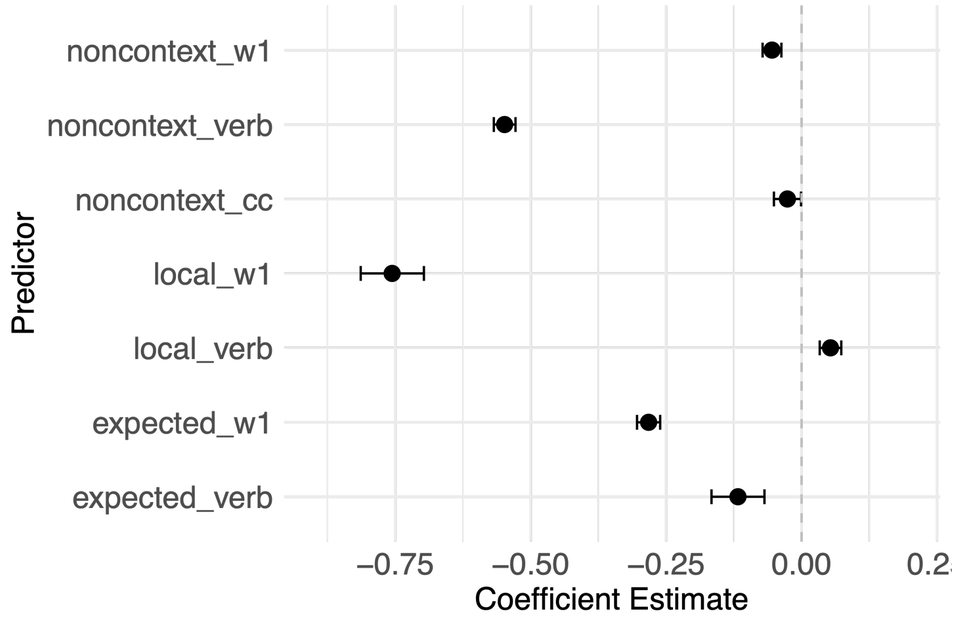
Results

- How is the information density computed?
  - Is the information density computed online during processing or based on accumulated contextual experience or some non-contextual heuristics (e.g., frequency)?

  both non-contextual, contextual local and expected informativity

- What is the context window for the uniform information density?
  - Is it before, at, or after the possible reduction site?

  …at the onset of the first word in the CC



Limitations and future directions

- Include the local informativity and the expected informativity of the syntactic information of the CC.
  - Future work: training a classifier to predict CC (Hao & Kaiser, 2025).
- Circularity issue with calculating the local informativity of w_1
  - Future work: training a language model on a corpus with no overt complementizer “that”.

Limitations and future directions

- Include the local informativity and the expected informativity of the syntactic information of the CC.
  - Future work: training a classifier to predict CC (Hao & Kaiser, 2025).
- Circularity issue with calculating the local informativity of w_1
  - Future work: training a language model on a corpus with no overt complementizer “that”.
- Interaction between local and expected information density
  - The local informativity of w_1 might have different effects on verbs with higher expected informativity or grammaticalized constructions than verbs with lower expected informativity.
- Cross modality comparison with spoken corpora or transcribed speech.

Acknowledgements

Thanks to

- Vic Ferreira and members of the Language and Cognition Lab and LeM🍋N Lab at UCSD for discussion.

https://github.com/pennydy/comp_drop

UC San Diego

Additional slides

Expected informativity
phonological reduction

$$\mathbb{E}_j[I(V_k \mid context_j)]$$
$$\approx \sum_{j:V_j=V_k} P(context_j \mid V_j) I(V_j \mid context_j)$$

Cohen Priva (2015)

Expected informativity
phonological reduction

$$\mathbb{E}_j[I(V_k \mid context_j)]$$
$$\approx \sum_{j:V_j=V_k} P(context_j \mid V_j) I(V_j \mid context_j)$$

- Segments that are usually unpredictable, i.e., having higher informativity across contexts, have longer duration and are less likely to delete even when they are predictable in context.

/n/ in *explanation* is not reduced, even though it is highly predictable in after [ʃə]

- Segments that are usually predictable, i.e., globally uninformative, can be reduced even when it is locally unpredictable.

Cohen Priva (2015)

Expected informativity
phonological reduction

$$\mathbb{E}_j[I(V_k \mid context_j)]$$
$$\approx \sum_{j:V_j=V_k} P(context_j \mid V_j) I(V_j \mid context_j)$$

- Segments that are usually unpredictable, i.e., having higher informativity across contexts, have longer duration and are less likely to delete even when they are predictable in context.

/n/ in *explanation* is not reduced, even though it is highly predictable in after [ʃə]

- Segments that are usually predictable, i.e., globally uninformative, can be reduced even when it is locally unpredictable.

/d/ in *sudden* tend to be deleted → [səṇ]

Cohen Priva (2015)

Corpus study

Bayesian mixed-effects logistic regression

position of the verb, distance between verb and CC,
length of CC onset, length of the remaining CC

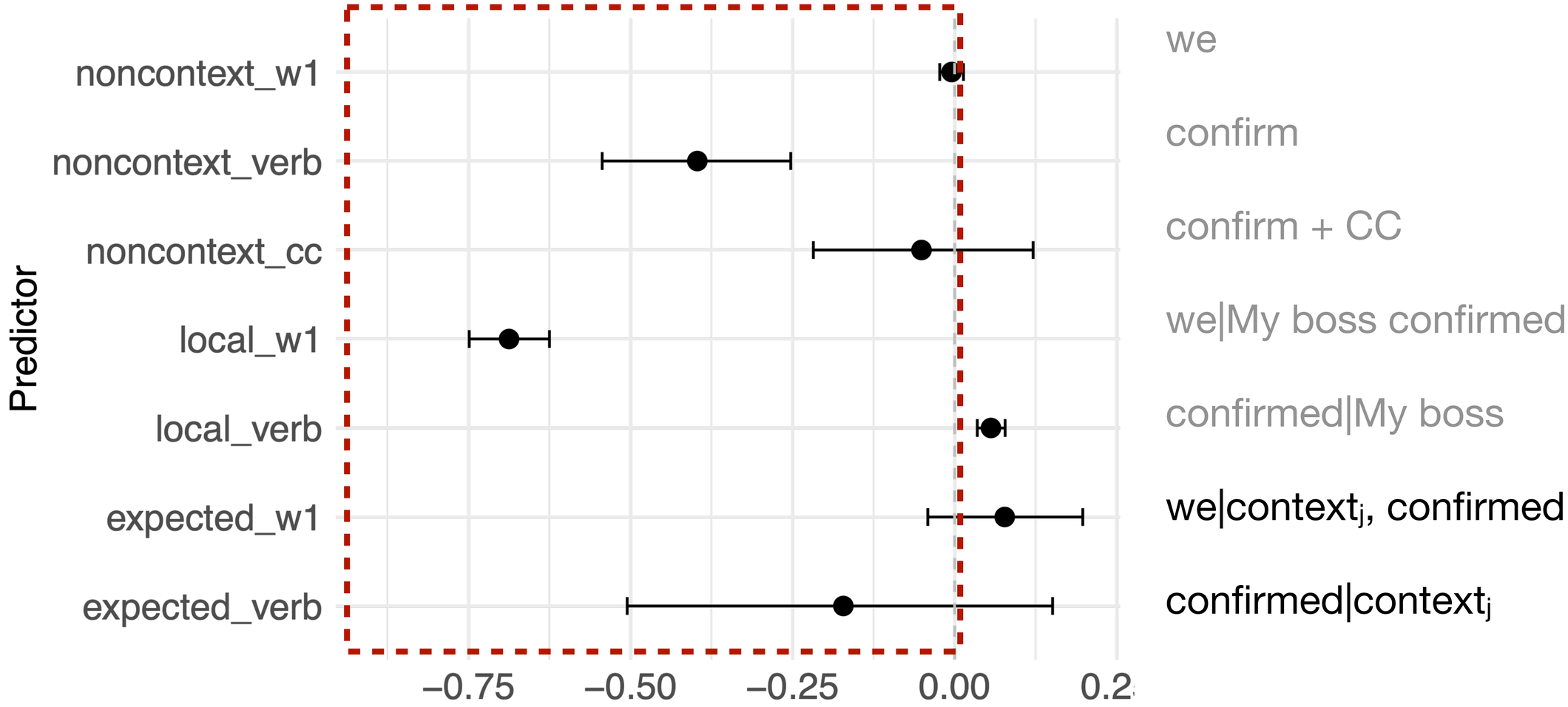
~ control variables + **main effects** + (1|verb)

	V_k	CC	w_1
non-contextual	**noncontext_verb**	**noncontext_cc**	**noncontext_w1**
local contextual	**local_verb**		**local_w1**
expected	**expected_verb**		**expected_w1**

Results

My boss confirmed [CC we were absolutely crazy].

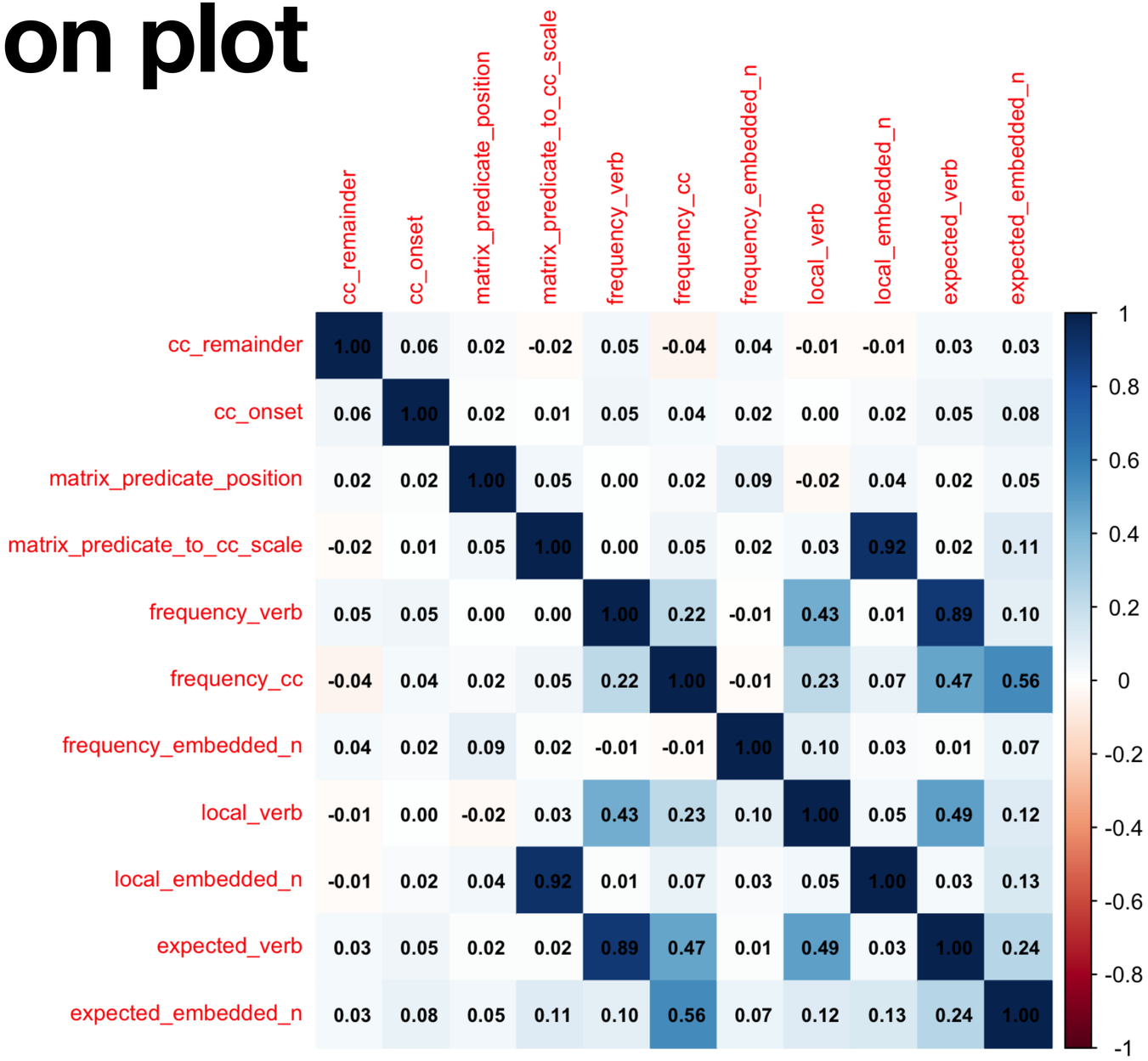
predicting the dropping of “that”



the **less** information content, the more likely “that” will be dropped

the more information content, the more likely “that” will be dropped

Correlation plot



heaviness?

semantic light/heavy verbs

→ My boss whispered [$_{CC}$ that we were absolutely crazy].

?My boss whispered [$_{CC}$ we were absolutely crazy].

My boss said [$_{CC}$ that we were absolutely crazy].

→ My boss said [$_{CC}$ we were absolutely crazy].