Explaining semantic universals: Session 3

Milica Denić

UNIVERSITÉ
DE GENÈVE

Question

What explains which meanings are lexicalized across languages?

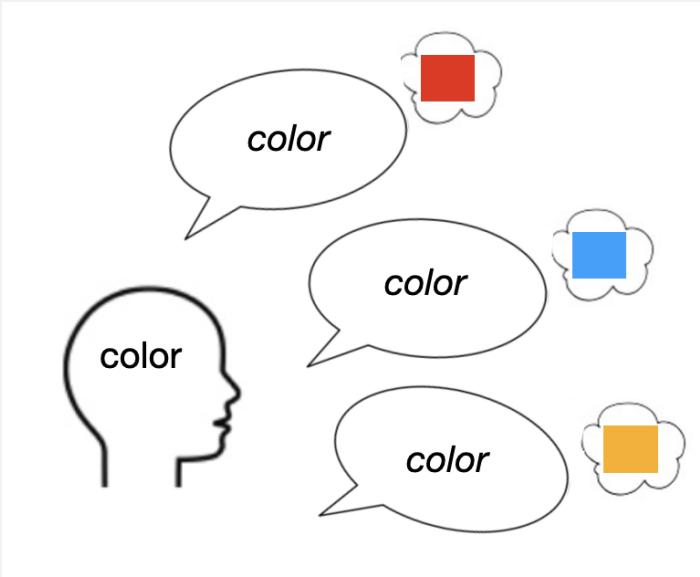
Simplicity of lexicon/informativeness trade-off

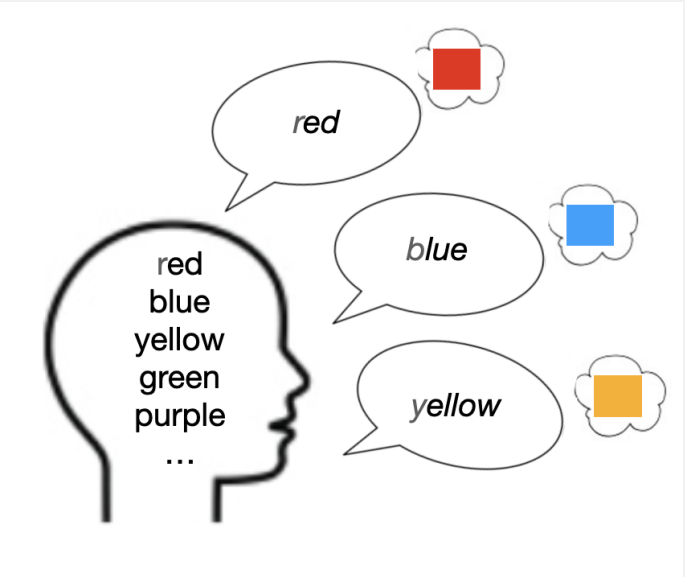
Kemp et Regier (2012), a.o.

Simplicity of lexicon/informativeness trade-off

Languages' lexicons are under two **competing pressures**:

(a) minimize **lexicon size** (b) maximize **informativeness**

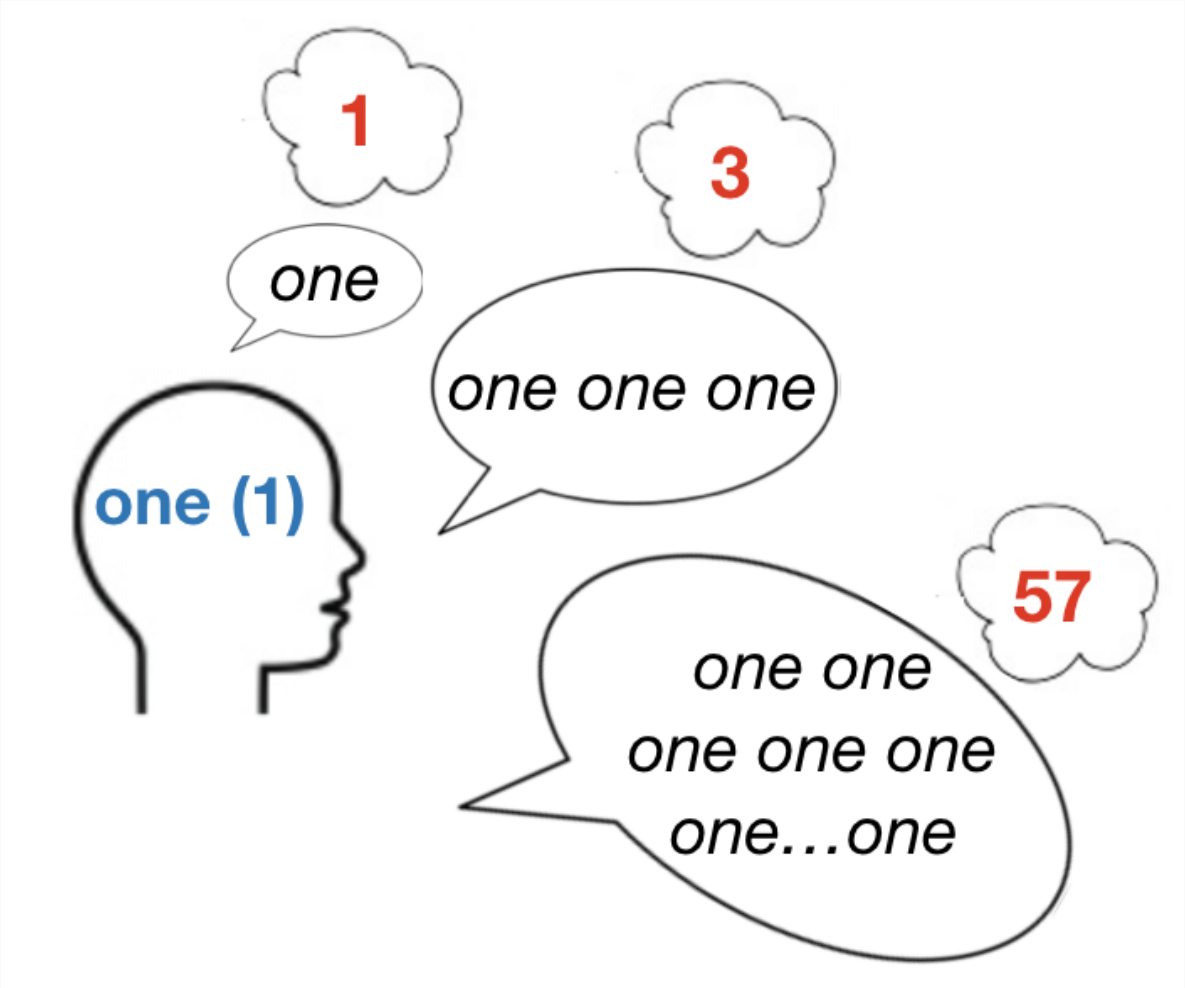




Languages **optimize the trade-off** between these two pressures.

Kemp et Regier (2012), a.o.

Yesterday: Simplicity of lexicon/informativeness trade-off optimization makes wrong predictions for numerals.



Yesterday: Pressures shaping lexicons

Not two, but (at least) three pressures:

1. **minimize lexicon size**
2. **maximize informativeness**
3. **minimize morphosyntactic complexity of utterances**

Today

We will discuss another class of expressions whose existence is surprising from the perspective of simplicity/informativeness trade-off.

(And doesn't seem to contribute to the minimization of morphosyntactic complexity either).

Negative polarity items

Negative polarity items (NPI)

Fauconnier (1975); Ladusaw (1979)

Negative polarity items (NPI)

Unacceptable NPIs

*John saw **any** bird.

Fauconnier (1975); Ladusaw (1979)

Negative polarity items (NPI)

Unacceptable NPIs	Acceptable NPIs
*John saw **any** bird.	John didn't see **any** bird.

Fauconnier (1975); Ladusaw (1979)

Negative polarity items (NPI)

Unacceptable NPIs	Acceptable NPIs
*John saw **any** bird.	John didn't see **any** bird.

A **entails** *B iff whenever A is true, B is true.*

Fauconnier (1975); Ladusaw (1979)

Negative polarity items (NPI)

Unacceptable NPIs

*John saw **any** bird.

Acceptable NPIs

John didn't see **any** bird.

Upward-entailing (UE)

John saw a bird.

$\Uparrow$

John saw a dove.

A **entails** *B iff whenever A is true, B is true.*

Fauconnier (1975); Ladusaw (1979)

Negative polarity items (NPI)

Unacceptable NPIs	Acceptable NPIs
*John saw **any** bird.	John didn't see **any** bird.
Upward-entailing (UE)	**Downward-entailing (DE)**
John saw a bird.	John didn't see a bird.
$\Uparrow$	$\Downarrow$
John saw a dove.	John didn't see a dove.

A **entails** *B iff whenever A is true, B is true.*

Fauconnier (1975); Ladusaw (1979)

NPIs are not acceptable in UE environments.

Everyone who has seen **any** bird is happy.

NPIs are not acceptable in UE environments.

Everyone who has seen **any** bird is happy.

NPIs are acceptable in DE environments.

*Everyone who is happy has seen **any** bird.

NPIs exist in very many (all?) languages.

When ‘any’ is acceptable, replacing it with a regular indefinite results in a truth-conditionally equivalent S (Chierchia, 2013).

When ‘any’ is acceptable, replacing it with a regular indefinite results in a truth-conditionally equivalent S (Chierchia, 2013).

(1) John didn’t see {any / a} bird.

When ‘any’ is acceptable, replacing it with a regular indefinite results in a truth-conditionally equivalent S (Chierchia, 2013).

(1) John didn’t see {any / a} bird.

NPIs increase the complexity of the lexicon, but they don’t seem to increase the informativeness of a language.

Why do languages have NPIs?

Hypothesis 1: Scope disambiguation (Barker, 2018)

(2) John didn't see any bird. $\neg \gg \exists$, $^*\exists \gg \neg$

(3) John didn't see a bird. $\neg \gg \exists$, $\exists \gg \neg$

NPIs help with disambiguation.

Hypothesis 1: Scope disambiguation (Barker, 2018)

(2) John didn't see any bird. $\neg \gg \exists$, $^*\exists \gg \neg$

(3) John didn't see a bird. $\neg \gg \exists$, $\exists \gg \neg$

NPIs help with disambiguation.

But:

- Why should we only care about ambiguities in DE environments?
  **Everyone bought any cookies.*

Hypothesis 1: Scope disambiguation (Barker, 2018)

(2) John didn't see any bird. $\neg \gg \exists$, $^*\exists \gg \neg$

(3) John didn't see a bird. $\neg \gg \exists$, $\exists \gg \neg$

NPIs help with disambiguation.

But:

- Why should we only care about ambiguities in DE environments?
  **Everyone bought any cookies.*
- Languages have existential elements (e.g., bare plurals in English) which don't create scope ambiguities and are not NPIs.

  John didn't buy cookies. $\neg \gg \exists$, $^*\exists \gg \neg$

  Everyone bought cookies. $\forall \gg \exists$, $^*\exists \gg \forall$

Hypothesis 2: Emphasis (Chierchia, 2013)

(4) John didn't see ANY bird.

(5) ?John didn't see A bird.

NPIs can be used in an emphatic way.

Hypothesis 2: Emphasis (Chierchia, 2013)

(4) John didn't see ANY bird.

(5) ?John didn't see A bird.

NPIs can be used in an emphatic way.

But:

- Most of the time, NPIs are not used with emphasis.
  John didn't see any bird.
- Languages have existential-like elements (e.g., numeral *one* in English) which can be used with emphasis and are not NPIs.

  John didn't buy ONE cookie.

  John bought ONE cookie.

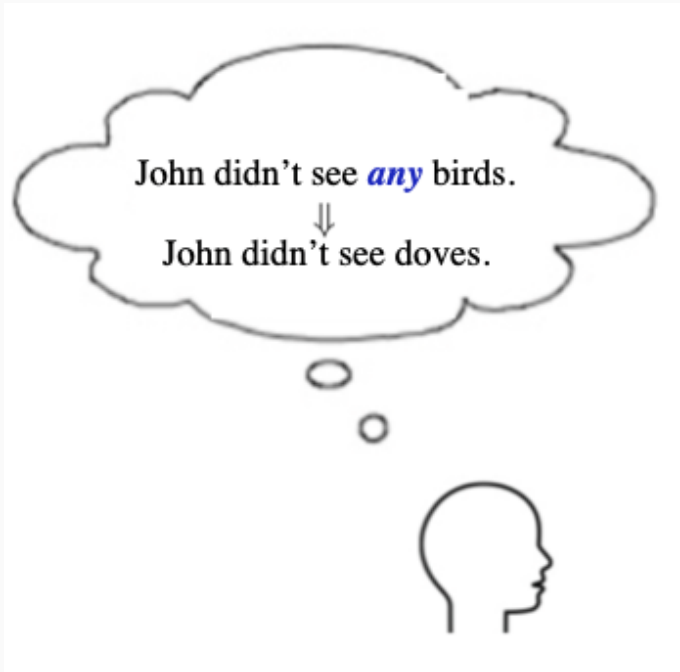
Hypothesis 3: Reasoning (Dowty, 1994; Szabolcsi et al., 2008)

NPIs could **signal that, according to the speaker, an environment is DE**, and exist due to pressure to **communicate not only the denotation but also potential inferences**.

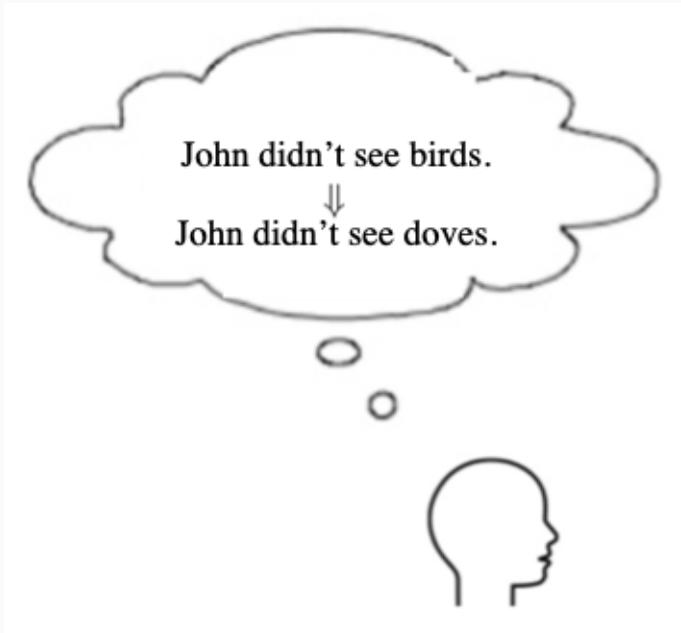
Hypothesis 3: Reasoning (Dowty, 1994; Szabolcsi et al., 2008)

NPIs could **signal that, according to the speaker, an environment is DE**, and exist due to pressure to **communicate not only the denotation but also potential inferences**.

More

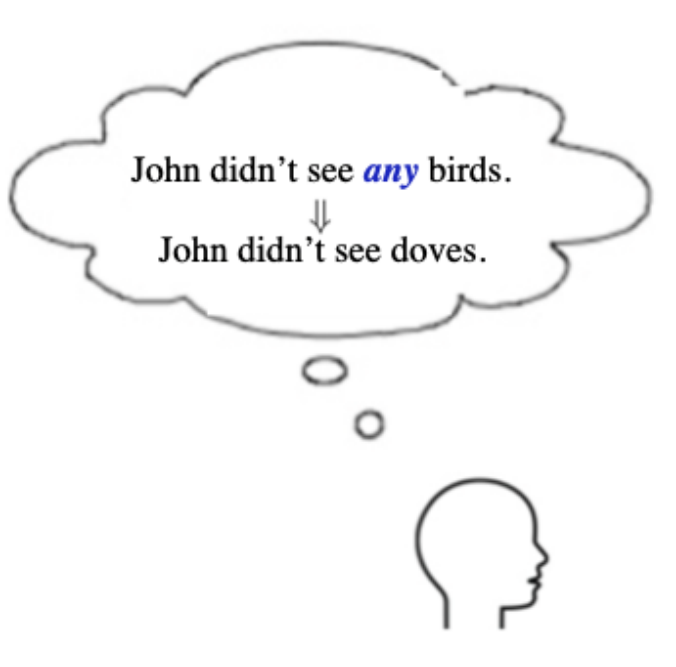


Less



Do NPIs actually influence reasoning?

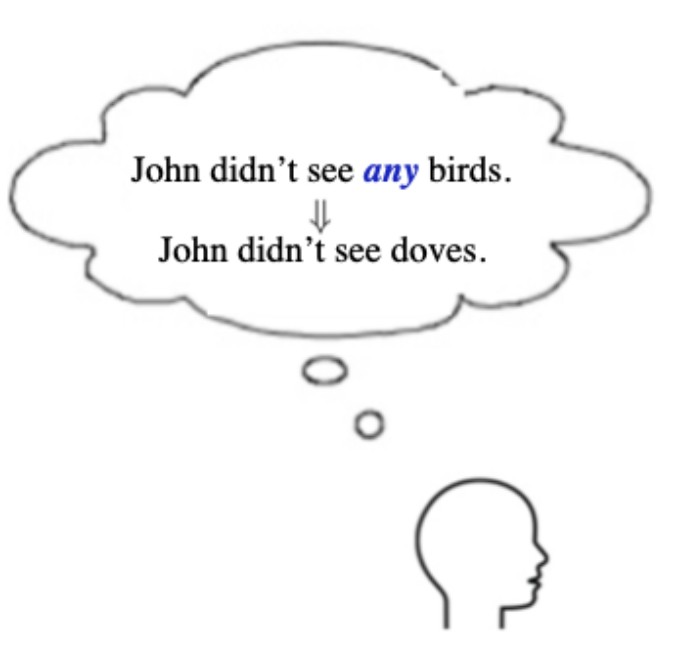
More



Less

Do NPIs actually influence reasoning?

More



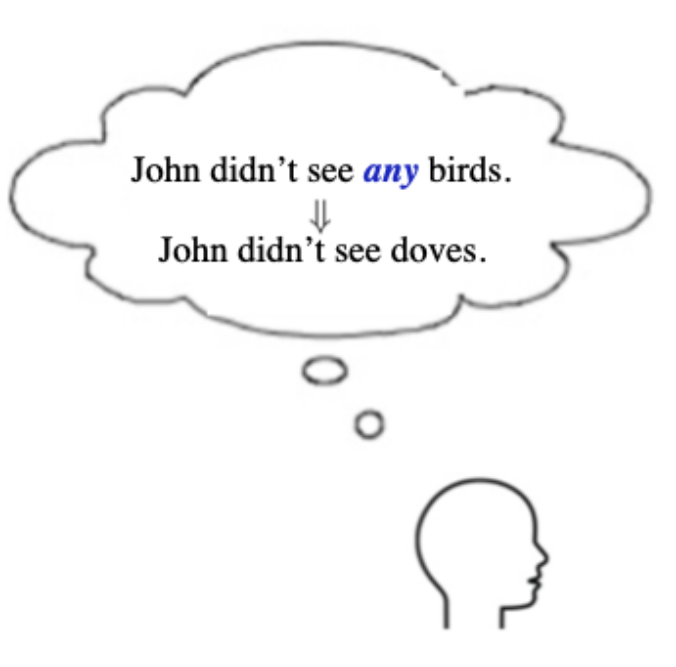
Less

Szabolcsi et al. (2008): no difference.

Very high inference acceptance rate in their experiment!

Do NPIs actually influence reasoning?

More



Less

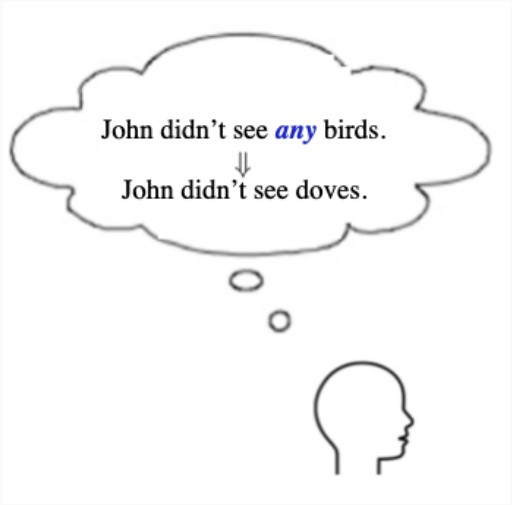
Szabolcsi et al. (2008): no difference.

Very high inference acceptance rate in their experiment!

→ *Is the influence of the NPI not detected due to the* **ceiling effect**?

Ceiling effect and how to avoid it

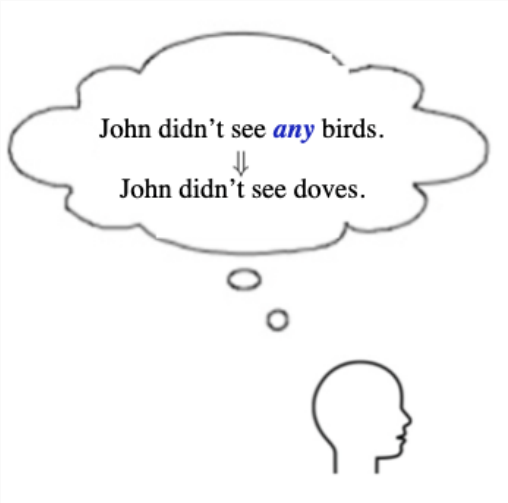
More



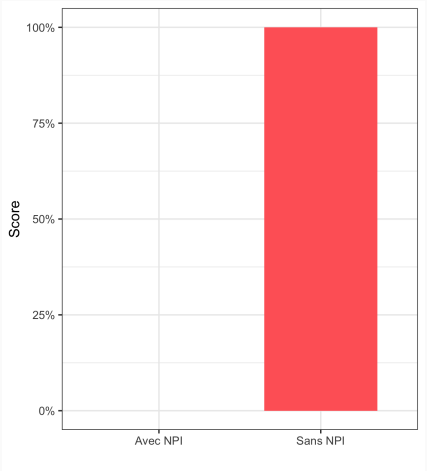
Less

Ceiling effect and how to avoid it

More

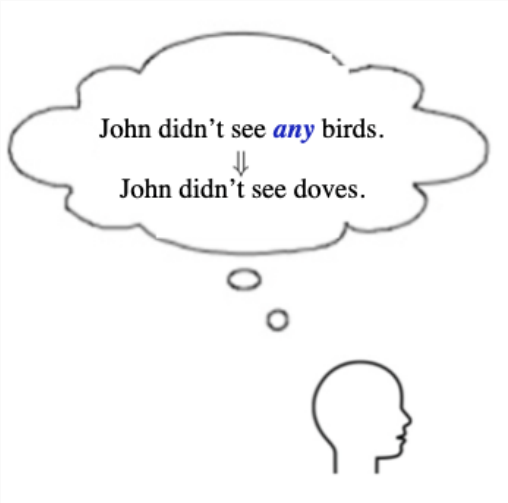


Less

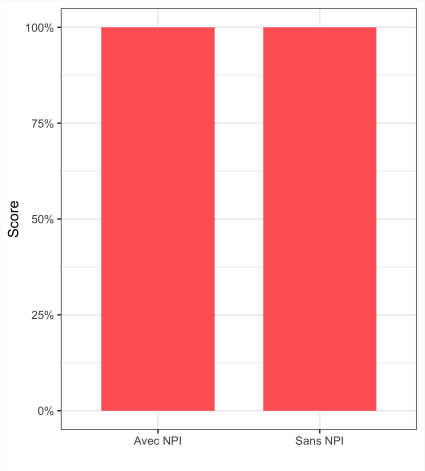


Ceiling effect and how to avoid it

More

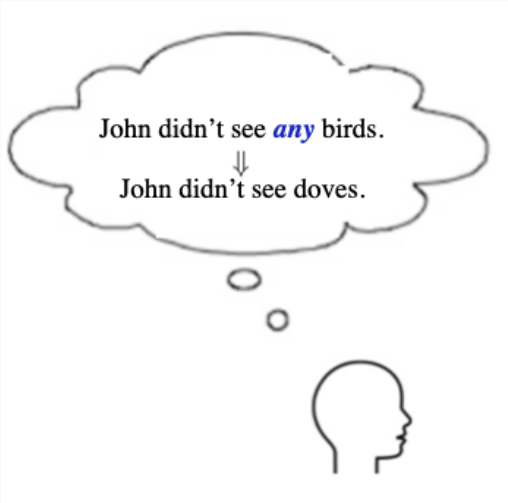


Less

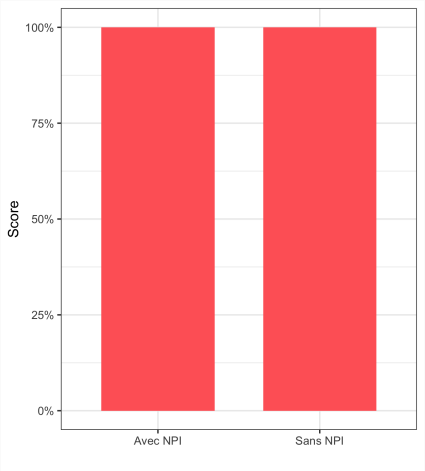


Ceiling effect and how to avoid it

More



Less



We will examine the influence of NPIs on inferences in an **environment where people have reasoning difficulties**.

Experiment

Denić et al. (2021), *Cognition*

A candidate: non-monotonic (NM) environments

A candidate: non-monotonic (NM) environments

Exactly 12 people saw a bird.

$\not\Downarrow$ $\not\Uparrow$

Exactly 12 people saw a dove.

A candidate: non-monotonic (NM) environments

Exactly 12 people saw a bird.

$\not\Downarrow$ $\not\Uparrow$

Exactly 12 people saw a dove.

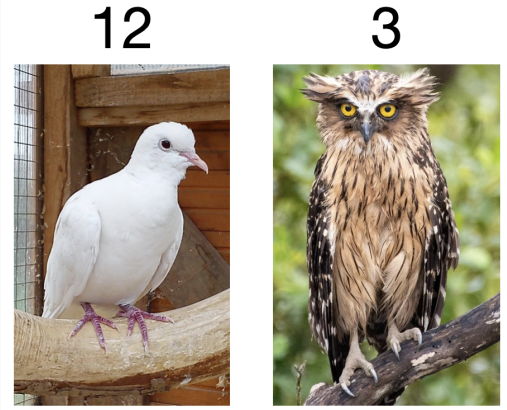
6 6

A candidate: non-monotonic (NM) environments

Exactly 12 people saw a bird.

$\nVdash$ $\nvDash$

Exactly 12 people saw a dove.



A candidate: non-monotonic (NM) environments

Exactly 12 people saw a bird.

$\not\Downarrow$ $\not\Uparrow$

Exactly 12 people saw a dove.

Chemla et al. (2011):

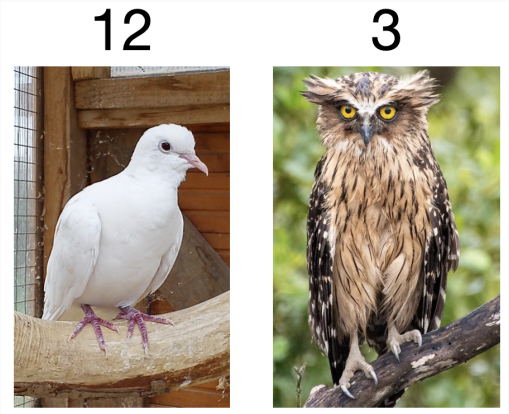
- People sometimes **mistakenly** think that these environments are DE.

A candidate: non-monotonic (NM) environments

Exactly 12 people saw a bird.

⇂ ↾

Exactly 12 people saw a dove.



Chemla et al. (2011):

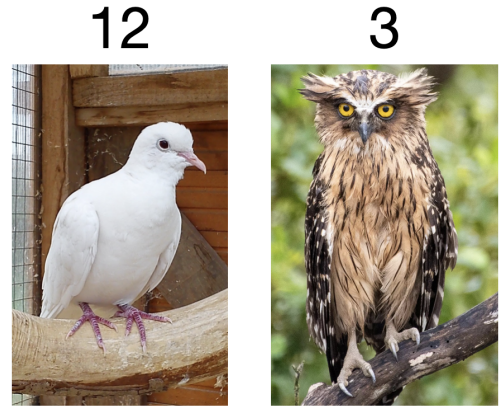
- People sometimes **mistakenly** think that these environments are DE.
- **Consequently, NPIs in NM environments are moderately acceptable**:

A candidate: non-monotonic (NM) environments

Exactly 12 people saw a bird.

⇃ ↾

Exactly 12 people saw a dove.



Chemla et al. (2011):

- People sometimes **mistakenly** think that these environments are DE.
- **Consequently, NPIs in NM environments are moderately acceptable**:

  ?Exactly 12 people saw **any** birds.

Research questions (more specifically)

Does the presence of an NPI **increase** *the inference rate from a general term to a specific term in NM environments?*

Denić et al. (2021), *Cognition*

Research questions (more specifically)

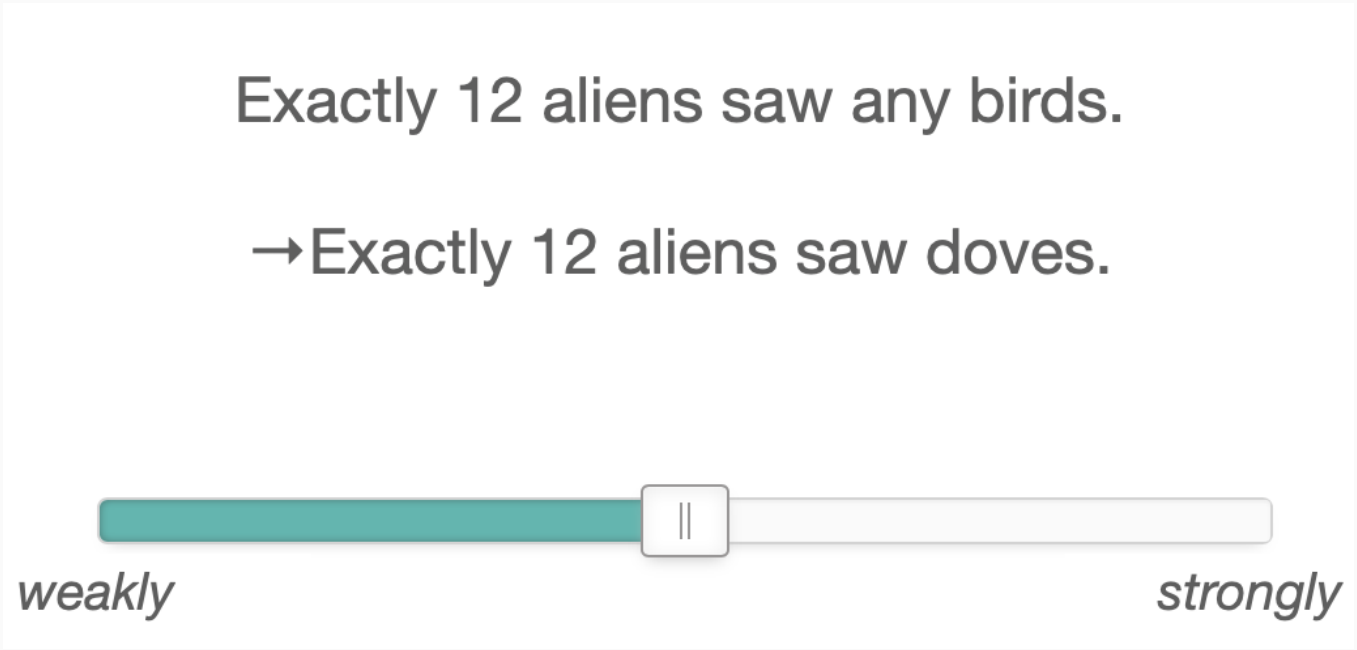
Does the presence of an NPI **increase** *the inference rate from a general term to a specific term in NM environments?*

Does the presence of an NPI **decrease** *the inference rate from a specific term to a general term in NM environments?*

Denić et al. (2021), *Cognition*

Inferential Judgment Task: Example Trial

Imagine that you hear the first sentence and indicate to what extent you would conclude that the second sentence is true.



We manipulate NPI presence and inference direction.

Inference direction 1: from a general term to a specific term.

With NPI

We manipulate NPI presence and inference direction.

Inference direction 1: from a general term to a specific term.

With NPI

Without NPI

We manipulate NPI presence and inference direction.

Inference direction 1: from a general term to a specific term.

With NPI

Without NPI

Prediction :

We manipulate NPI presence and inference direction.

Inference direction 1: from a general term to a specific term.

With NPI

Without NPI

Prediction : With NPI $>$ Without NPI

We manipulate NPI presence and inference direction.

Inference direction 2: from a specific term to a general term.

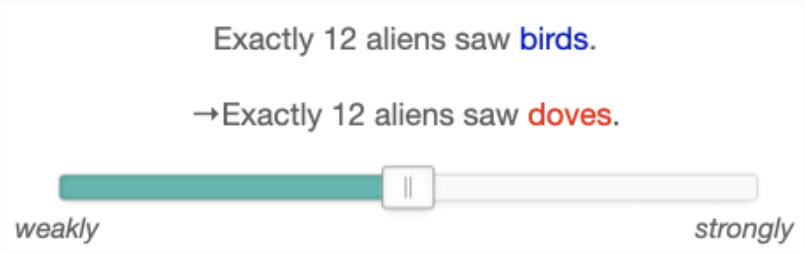
With NPI

We manipulate NPI presence and inference direction.

Inference direction 2: from a specific term to a general term.

With NPI

Without NPI

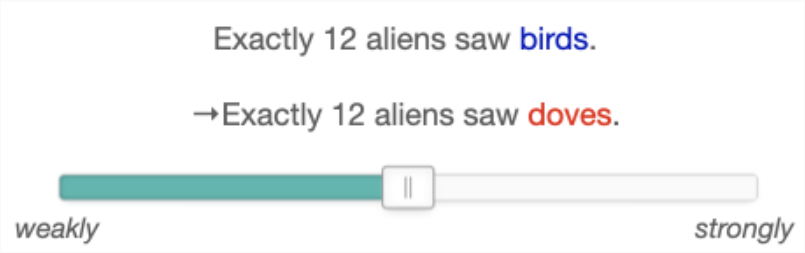


We manipulate NPI presence and inference direction.

Inference direction 2: from a specific term to a general term.

With NPI

Without NPI



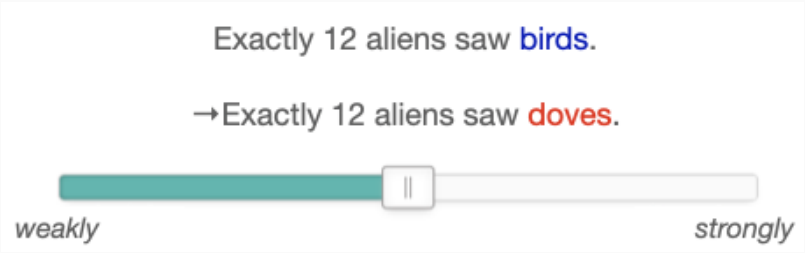
Prediction :

We manipulate NPI presence and inference direction.

Inference direction 2: from a specific term to a general term.

With NPI

Without NPI



Prediction : With NPI $<$ Without NPI

Summary of predictions

Summary of predictions

Inference direction 1: from a general term to a specific term.

Prediction: With NPI $>$ Without NPI

Summary of predictions

Inference direction 1: from a general term to a specific term.

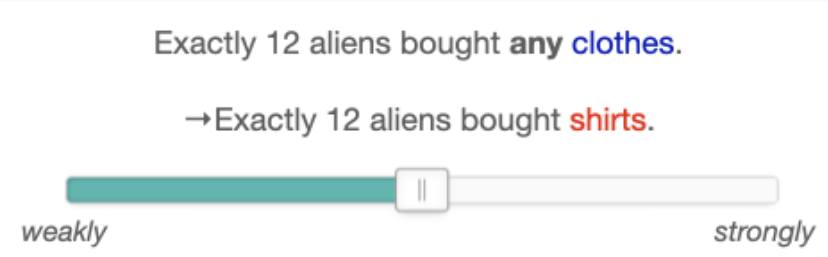
Prediction: With NPI $>$ Without NPI

Inference direction 2: from a specific term to a general term.

Prediction: With NPI $<$ Without NPI

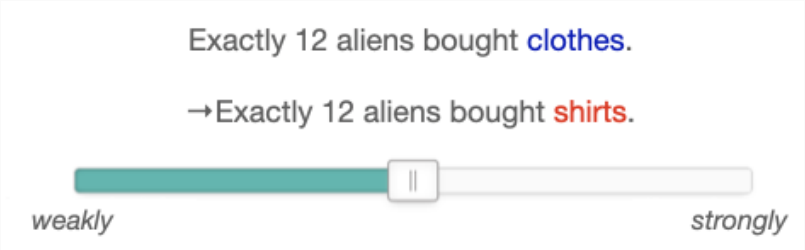
There are 8 items in each of the four conditions.

Inf. direction 1, With NPI



...

Inf. direction 1, Without NPI



...

Items and participants

Within-subject experiment

Items and participants

Within-subject experiment

128 items, presented in a random order, of which:

Items and participants

Within-subject experiment

128 items, presented in a random order, of which:

- 32 target items (8 per condition);

Items and participants

Within-subject experiment

128 items, presented in a random order, of which:

- 32 target items (8 per condition);
- 96 control items, evaluating reasoning in DE and UE environments.

Items and participants

Within-subject experiment

128 items, presented in a random order, of which:

- 32 target items (8 per condition);
- 96 control items, evaluating reasoning in DE and UE environments.

66 participants (74 recruited, 8 excluded)

Results 1: from a general to a specific term

Results 1: from a general to a specific term

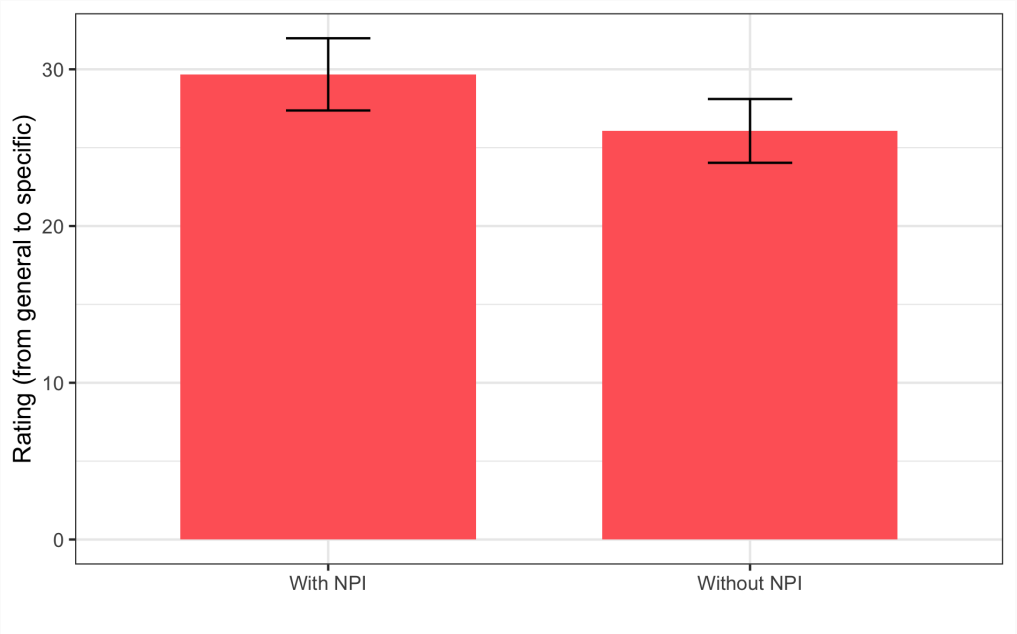
With NPI

Without NPI

Results 1: from a general to a specific term

With NPI

Without NPI



$(\chi^2(1) = 5.9, p = .01)$

Results 2: from a specific to a general term

Results 2: from a specific to a general term

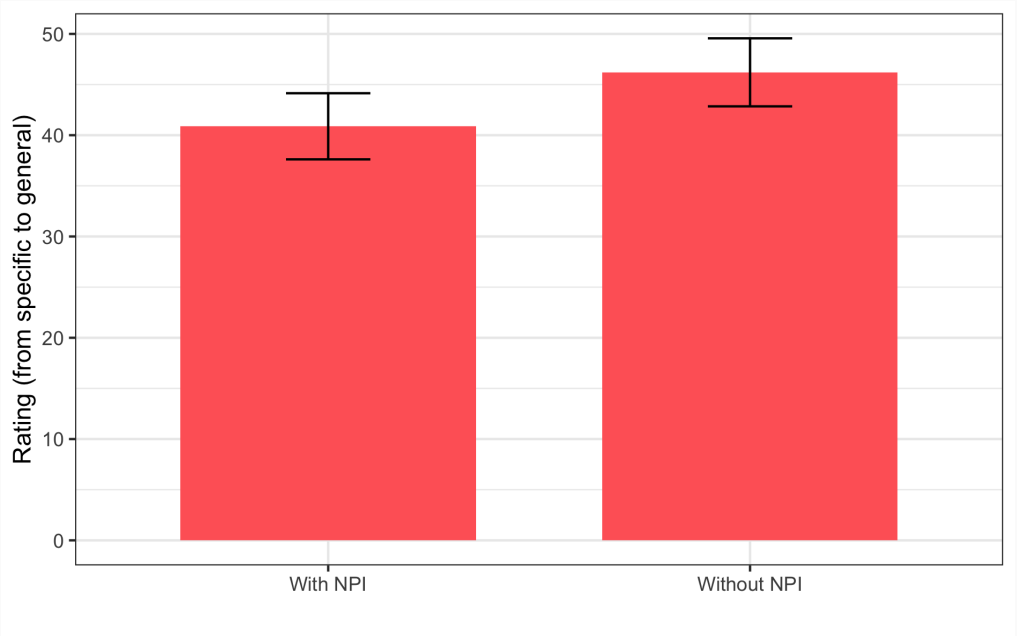
With NPI

Without NPI

Results 2: from a specific to a general term

With NPI

Without NPI



$(\chi^2(1) = 11.2, p < .001)$

Summary of the findings

Summary of the findings

Inference direction 1: from a general term to a specific term.

Result: With NPI $>$ Without NPI

Summary of the findings

Inference direction 1: from a general term to a specific term.

Result: With NPI $>$ Without NPI

Inference direction 2: from a specific term to a general term.

Result: With NPI $<$ Without NPI

Study conclusions

The problem:

> NPIs increase the complexity of the lexicon, but they don't seem to increase the informativeness of a language.

The result:

> **But NPIs signal that (according to the speaker) their environment validates general $\rightarrow$ specific inferences!**

$\rightarrow$ They may serve as a **semantic processing aid**.

Study conclusions

NPIs contribute to the overall informativeness of language (maximizing the probability that the speaker and listener communicate successfully) by reducing (inference computation) noise.

More generally, to what extent can such (near-)duplicates in functional lexicon be shown to be efficient from language processing perspective?

A related example?

A related example?

English complementizers for embedded declaratives: $C_1 = \emptyset$, $C_2 = \textit{that}$

(6) Mary believes (that) John came to the party.

A related example?

English complementizers for embedded declaratives: $C_1 = \emptyset$, $C_2 = \textit{that}$

(6) Mary believes (that) John came to the party.

People tend to produce overt *that* more often when the sentence is (informationally) difficult to process. (Jaeger et Levy, 2006)

A related example?

English complementizers for embedded declaratives: $C_1 = \emptyset$, $C_2 = \textit{that}$

(6) Mary believes (that) John came to the party.

People tend to produce overt *that* more often when the sentence is (informationally) difficult to process. (Jaeger et Levy, 2006)

What about other (near-)duplicates in functional vocabulary?

Open questions and future directions for NPIs

Why don't we have items specializing for NM or UE?

NB: Positive polarity items arguably don't specialize for UE.

Open questions and future directions for NPIs

Why do we see NPIs of only some syntactic categories?

Determiner: any
Quantifier: anyone
Adverb: yet
Verb: sleep a wink

But no adjective or noun NPIs?

Zooming out

This tutorial/research seminar

Question: Which meanings are lexicalized across languages?

We zoomed into one line of work within this rich literature: **the simplicity/informativeness trade-off hypothesis** (Kemp et Regier, 2012 and others).

We explored both some of its **successes** and some of its **limitations**.

This tutorial/research seminar

Question: Which meanings are lexicalized across languages?

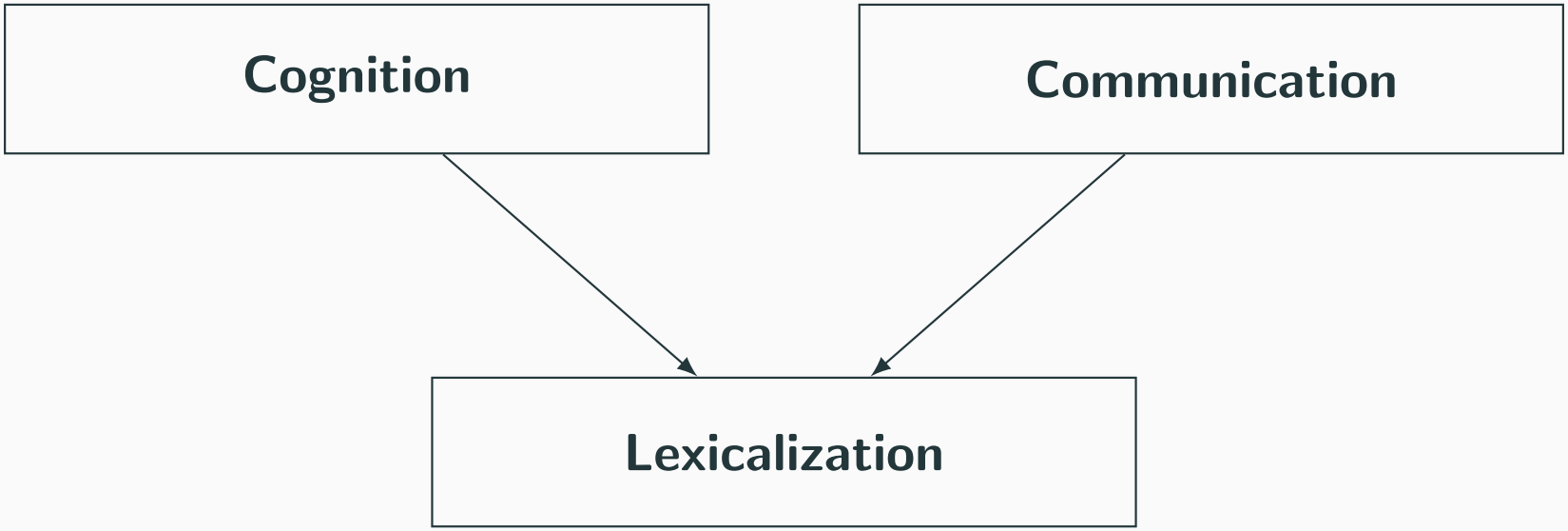
We zoomed into one line of work within this rich literature: **the simplicity/informativeness trade-off hypothesis** (Kemp et Regier, 2012 and others).

We explored both some of its **successes** and some of its **limitations**.

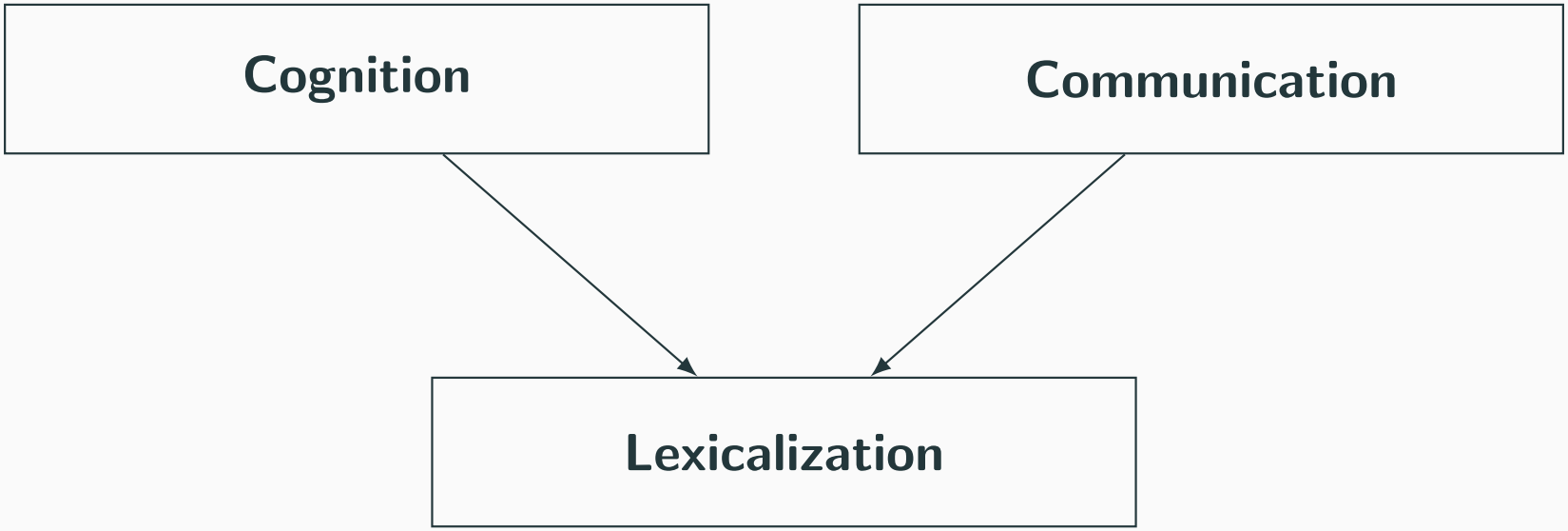
The limitations led us to several discoveries:

- an additional pressure shaping lexicons
- a better understanding of how informativeness shapes lexicons
- hypotheses about cognitive biases shaping language (e.g., negation, division cognitively complex)

But there is much more work to be done in disentangling cognitive and communicative pressures shaping lexicons...



But there is much more work to be done in disentangling cognitive and communicative pressures shaping lexicons...



What's explained and what remains to be explained (across semantic domains)? What can we learn from what remains to be explained?

References

Chris Barker : Negative polarity as scope marking. Linguistics and philosophy, 41(5):483–510, 2018.

Emmanuel Chemla, Vincent Homer et Daniel Rothschild : Modularity and intuitions in formal semantics: The case of polarity items. Linguistics and Philosophy, 34(6):537–570, 2011.

Gennaro Chierchia : Logic in Grammar: Polarity, Free choice, and Intervention. Oxford University Press, 2013.

Milica Denić, Vincent Homer, Daniel Rothschild et Emmanuel Chemla : The influence of polarity items on inferential judgments. Cognition, 215:104791, 2021.

References ii

David Dowty : The role of negative polarity and concord marking in natural language reasoning. In Semantics and Linguistic Theory, pages 114–144, 1994.

Gilles Fauconnier : Polarity and the scale principle. Chicago Linguistics Society, 11:107–126, 1975.

T Jaeger et Roger Levy : Speakers optimize information density through syntactic reduction. Advances in neural information processing systems, 19, 2006.

Charles Kemp et Terry Regier : Kinship categories across languages reflect general communicative principles. Science, 336 (6084):1049–1054, 2012.

A. W. Ladusaw : Polarity sensitivity as inherent scope relations. Thèse de doctorat, University of Texas at Austin, 1979.

References iii

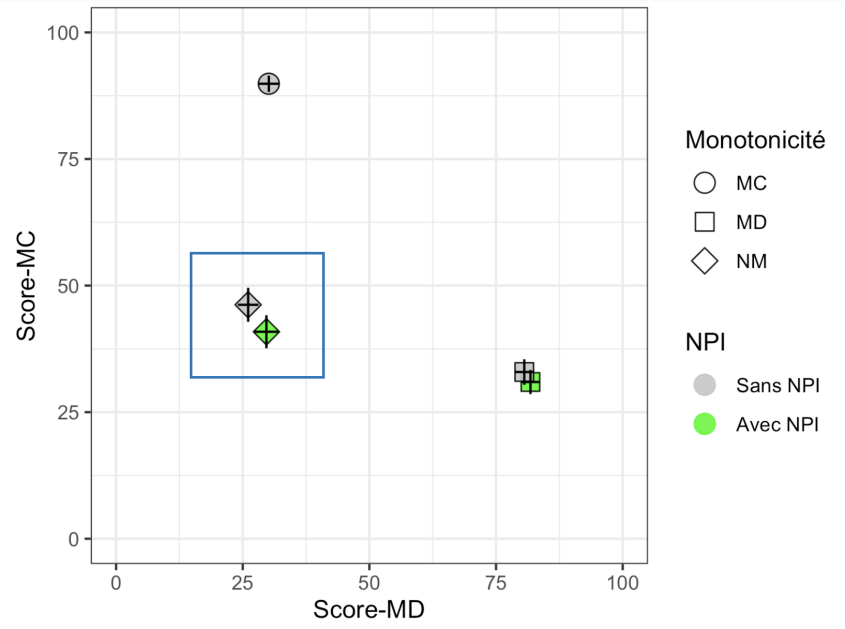
Anna Szabolcsi, Lewis Bott et Brian McElree : The effect of negative polarity items on inference verification. Journal of semantics, 25(4):411–450, 2008.

Other parsing aids across languages?

For questions: either a semantic analysis of questions that preserves the DE analysis;

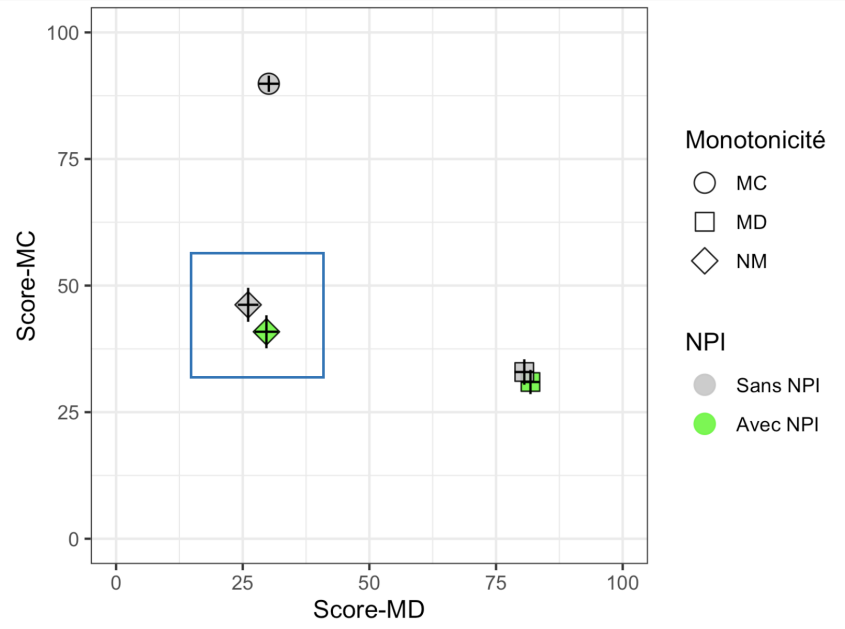
Or another property licenses NPIs that somehow strongly correlates with downward monotonicity...

Alternative explanations and why not



This result can't be due to:

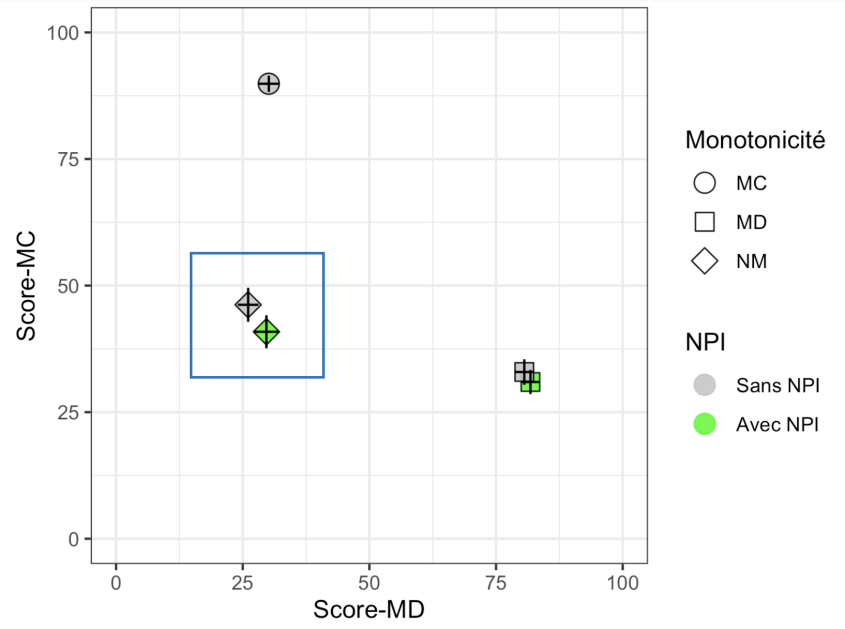
Alternative explanations and why not



This result can't be due to:

- a *yes*-response bias

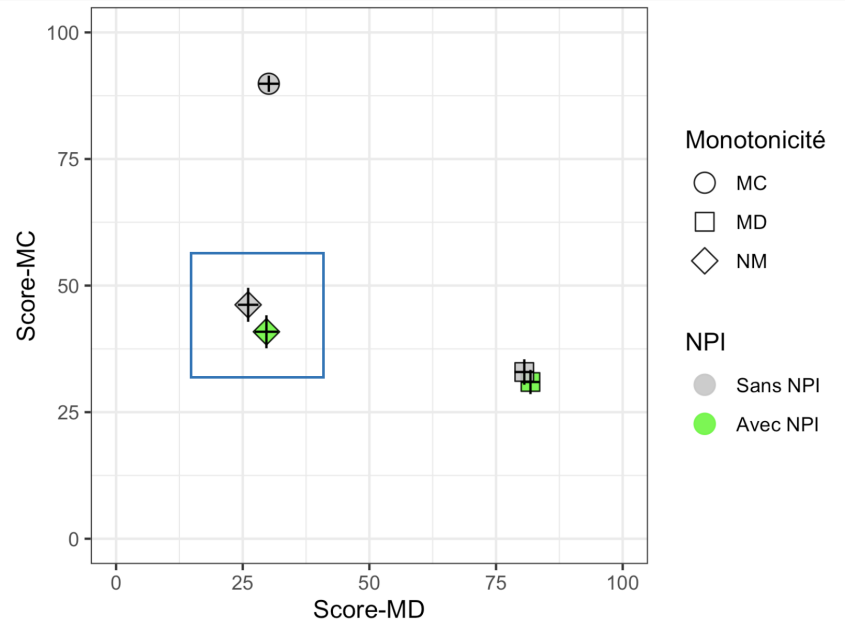
Alternative explanations and why not



This result can't be due to:

- a *yes*-response bias
- a *no*-response bias

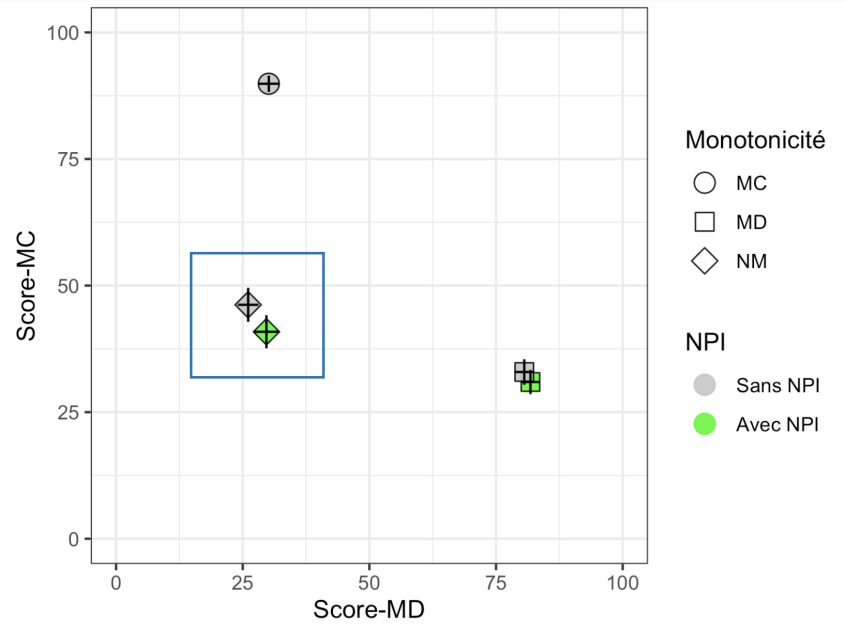
Alternative explanations and why not



This result can't be due to:

- a *yes*-response bias
- a *no*-response bias
- *noise*

Alternative explanations and why not



This result can't be due to:

- a *yes*-response bias
- a *no*-response bias
- *noise*

Résultats

Résultats

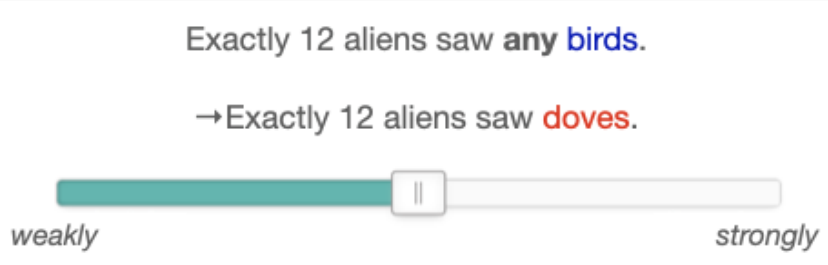
Inférence d'un terme général à un terme spécifique

Avec NPI > **Sans NPI**

Résultats

Inférence d'un terme général à un terme spécifique

Avec NPI > **Sans NPI**

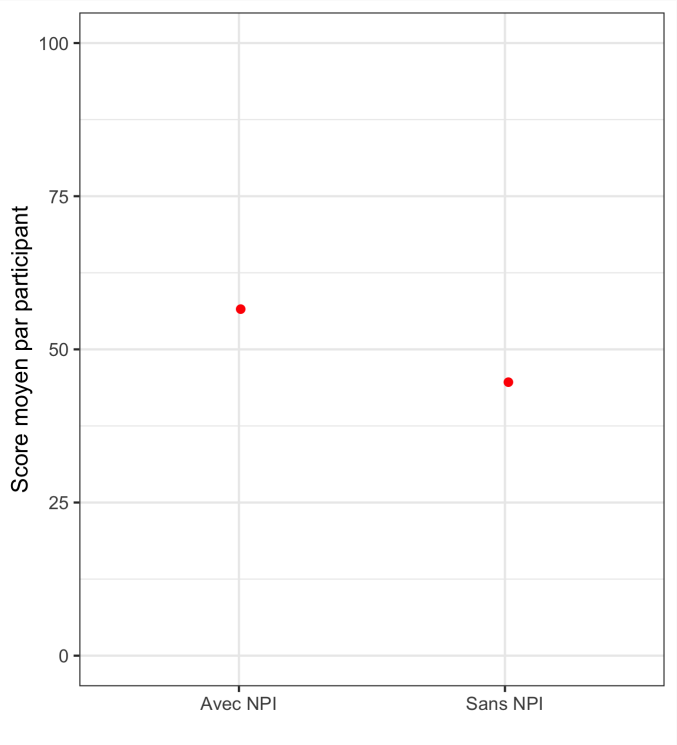


participant	Score moyen 'Avec NPI'	Score moyen 'Sans NPI'
participant1	52%	41%
participant2	33%	26%
participant3	17%	13%
...	...	...

Résultats

Inférence d'un terme général à un terme spécifique

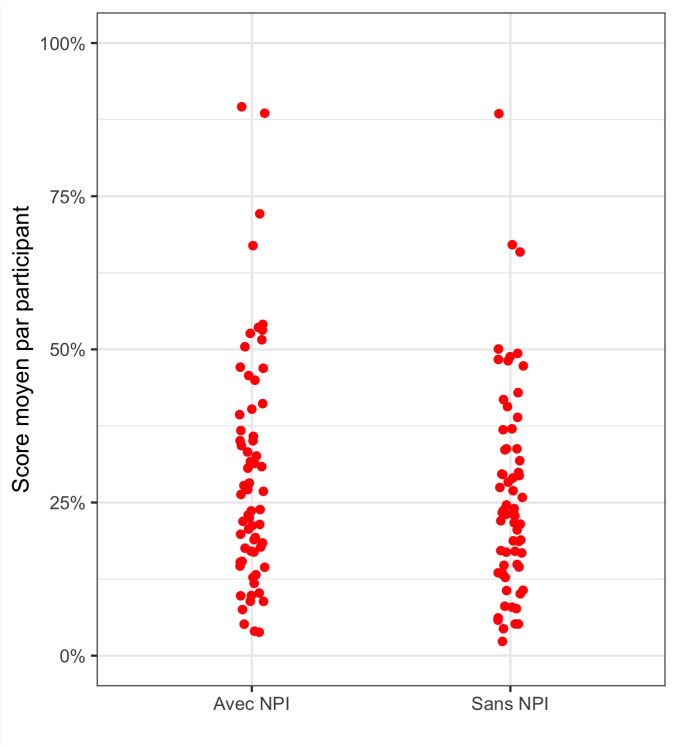
Avec NPI > **Sans NPI**



Résultats

Inférence d'un terme général à un terme spécifique

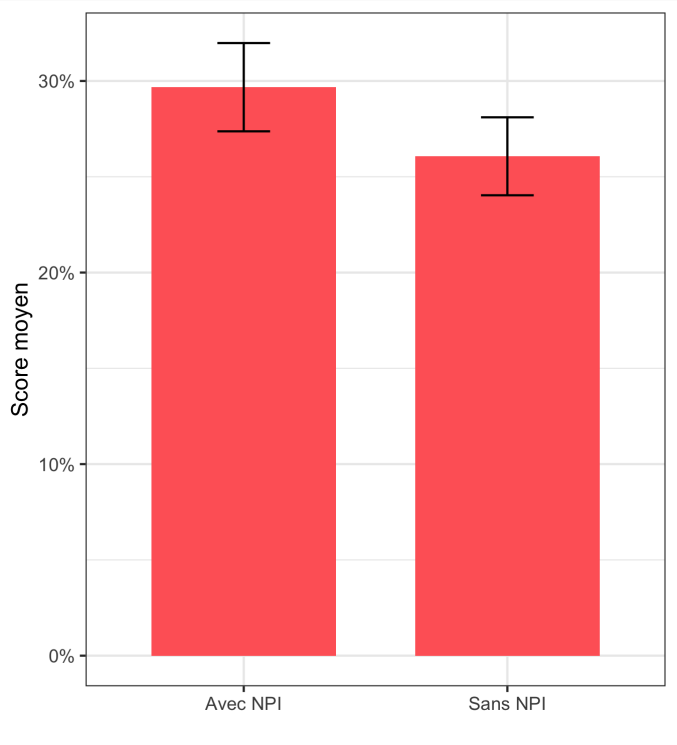
Avec NPI > **Sans NPI**



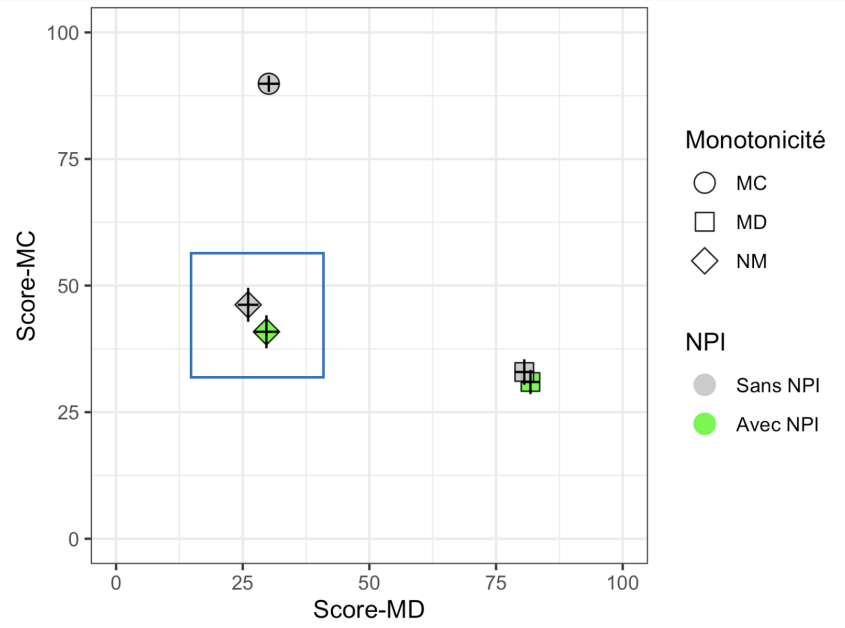
Résultats

Inférence d'un terme général à un terme spécifique

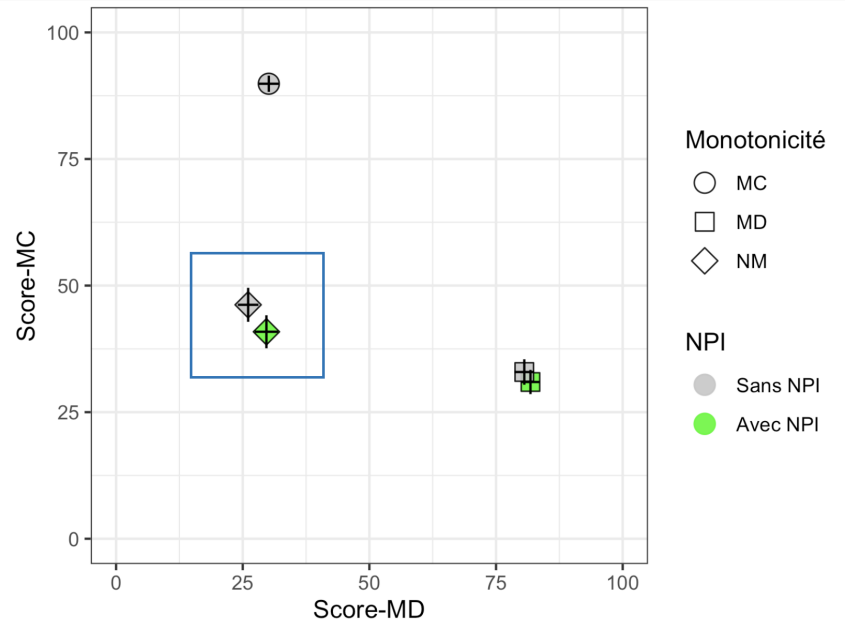
Avec NPI > **Sans NPI**



Les NPI augmentent la mesure dans laquelle un environnement est perçu comme MD et comme non-MC.



Les NPI augmentent la mesure dans laquelle un environnement est perçu comme MD et comme non-MC.

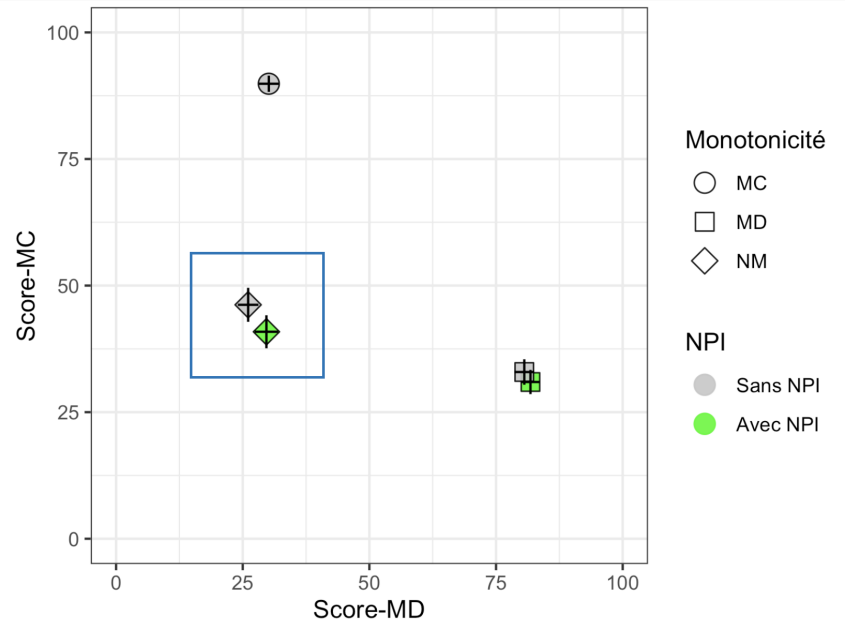


Réponses alignées : Score-MD

$$100 - \text{Score-MC}$$

Réponses alignées $\sim$ **NPI**

Les NPI augmentent la mesure dans laquelle un environnement est perçu comme MD et comme non-MC.



Réponses alignées : Score-MD

$$100 - \text{Score-MC}$$

Réponses alignées $\sim$ **NPI**

$$(\chi^2(1) = 18.6, p < .001)$$

Contrôles

	Environnement MC	**Environnement MD**
Score-MC $\Uparrow$	Tout le monde a vu une colombe. $\rightarrow$ Tout le monde a vu un oiseau.	Peu de gens ont vu une colombe. $\rightarrow$ Peu de gens ont vu un oiseau.
Score-MD $\Downarrow$	Tout le monde a vu un oiseau. $\rightarrow$ Tout le monde a vu une colombe.	Peu de gens ont vu un oiseau. $\rightarrow$ Peu de gens ont vu une colombe