







European Research Counci Established by the European Commission



ncil ALMA MATER STUDIORUM università di Bologna Funded by the European Union (GRANT AGREEMENT: ERC-2021-STG-101039777). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

## Word Specificity: A Measure to Investigate Semantic Abstraction and Linguistic Creativity

Giulia Rambelli

giulia.rambelli4@unibo.it

Hong Kong, August 7 2024





## Abstraction, a hallmark of human cognition, is the ability to pull off meaning from the experience





## WHAT IS ABSTRACTION?

Lower levels of abstraction (i.e., higher levels of concreteness) capture thoughts that are more specific, detailed, vivid, and imageable [...]. Higher levels of abstraction (i.e., lower levels of concreteness), on the other hand, include fewer readily observable characteristics and therefore capture thoughts that are less imageable<sub>[1]</sub> [Starting from the concrete notion of bridge], humans can easily understand extended and metaphorical notions such as "water bridges," "ant bridges," "the bridge of a song," "bridging the gender gap," "a bridge loan," "burning one's bridges," "water under the bridge," and so on. [...] One makes an abstraction of a concept when one extends that concept to more general instances, ones that are more removed from specific entities, as in the examples of "bridge"<sub>[2]</sub>

[1] Burgoon, Henderson, Markman(2013). There are many ways to see the forest for the trees: A tour guide for abstraction. Perspectives on Psych Science, 8, 501–520.

[2] Mitchell, M. (2021). Abstraction and Analogy-Making in Artificial Intelligence.

## **DISENTANGLING TWO VARIABLES**

#### (Abstractness) Concreteness

- the degree of perceptibility of a word referent
- a quality of concept (word) that has often exemplified by multiple interactive elements with low percetual similarity.

#### Specificity

- the degree of precision of a word meaning in terms of category inclusiveness
- taxonomic hierarchical organization





**ABSTRACTION** 





## VARIABLES OF ABSTRACTION



LOW SPECIFICITY

[3] Bolognesi, Burgers, & Caselli(2020). On abstraction: decoupling conceptual concreteness and categorical specificity. Cognitive Processing, 21(3), 365-381.



## ABSTRACTION IN LANGUAGE

Language is a lens through which we investigate conceptual representation.

Investigate how humans deal with abstraction from a data-driven perspective

- Corpus analyses and Distributional Semantic Models
- Linguistic capabilities of Large Language Models (LLMs)





#### **3 QUESTIONS ON WORD SPECIFICITY**

- 1. How does Specificity affect conceptual representation (through the lens of linguistic distributions)?
- 2. How does Specificity affect language comprehension in LLMs?
- 3. What role does word Specificity play in mechanisms of linguistic productivity and creativity?



## THE ROLE OF CONTEXT IN ABSTRACTION

#### **Context Availability Hypothesis**

there are differences in the availability and strength of contextual associations between concrete and abstract words [4]



Abstract words appear in various contexts, while concrete words are found in fewer contexts.

#### What about words varying in Specificity?

[4] Schwanenflugel et al. (1988). Context availability and lexical decisions for abstract and concrete words. Journal of Memory and Language, 27(5), 499-520.



## THE ROLE OF CONTEXT IN ABSTRACTION

What are the data-driven patterns highlighting differences between words with different degrees of abstraction?

Contextual Variability Depends on Categorical Specificity rather than Conceptual Concreteness: A Distributional Investigation on Italian data

The Contextual Variability of English Nouns: The Impact of Categorical Specificity beyond Conceptual Concreteness

Giulia Rambelli



Marianna M. Bolognesi



Giulia Rambelli, Marianna M. Bolognesi





#### COONTEXTUAL VARIABILITY

Contextual variability is the differences between the contexts of occurrence

- 1. Compute how a word and its contexts are similar using computational measures
- 2. Look how the patterns are the same for words varying in both Concreteness and Specificity



## **EXPERIMENTS OVERVIEW**

#### Material

- 676 English and 662 Italian nouns from ANEW
- Contexts extracted from ukWaC and ItWaC, used to create DSM

#### **Contextual Variability Measures**

• TNk / TCk: cosine similarity between a word and its k neighbors/contexts

2. Specificity

- NNk / CCk: cosine similarity of the k neighbors/contexts
- Entropy

#### **Regression Analyses**

- investigate how CV measures are explained in terms of
  - 1. Concreteness
    - $CV_i \sim Conc$   $CV_i \sim Spec$

3. Their interaction  $CV_i \sim Conc * Spec$ 



## **R1: CONCRETENESS EFFECTS**

	EN	IT
TN5	4.74%	2.10%
NN5	9.88%	2.40%
TC5	15.80%	4.70%
CC5	3%	4%
ENTROPY	0.13%	2.10%

- Higher  $R^2$ : TC and NN
- High correlation between Concreteness and TC10 (.44) and NN5 (.42)



### **R2: SPECIFICITY EFFECTS**

	EN	IT	
TN5	4.74%	2.10%	
NN5	9.88%	2.40%	
TC5	15.80%	4.70%	
CC5	3%	4%	
ENTROPY	0.13%	2.10%	

	EN	IT
TN5	2.83%	13%
NN5	6.90%	11.80%
TC5	11.18%	26.50%
CC5	2.53%	0%
ENTROPY	14.96%	34.30%

- Higher  $R^2$ : TC and NN
- High correlation between Concreteness and TC10 (.44) and NN5 (.42)

- Higher R<sup>2</sup> : Entropy and TC
- Negative correlation between Entropy and Specificity



#### R2: SPECIFICITY EFFECTS

	PASTA	FOOD
Concreteness	4.86	4.8
Specificity	4.23	1.52
TC10	.65	.49
Top contexts	dish (.66), sauce(.81), bread (.59), rice (.59) , food (.49), salad (.78)	eat (.64), find (.29), drink (.61), chain (.35), animal (.51), fast (.22)

- Generic words occur in various contexts not closely tied to the target word
- Specific words exhibit a stronger association with similar contexts
  - The same for abstract words (thus less evident in English)



#### **R3: INTERACTION EFFECTS**

	EN	IT
TN5	4.54%	13.10%
NN5	10.09%	11.80%
TC5	16.81%	27.70%
CC5	3.76%	6.70%
DCR	1.49%	4.60%
ENTROPY	26.80%	36.60%

#### Entropy

Same trend for EN and IT

- 1. generic words have high entropy (top)
- specific words have low entropy (bottom)





#### TAKEHOME MESSAGE

- **Specific words** have well-defined, similar contexts.
- Generic words, whether abstract or concrete, have broader and more diverse contexts.
- **Concreteness** is more significant in explaining noun contextual variability in English than in Italian.
- The **interaction** between Concreteness and Specificity accounts for a significant portion of the variation in the regression analyses.
- Entropy is cross-linguistically reliable, while measures computed using similar neighbors or syntagmatic contexts are correlated but more language-dependent.



#### **3 QUESTIONS ON WORD SPECIFICITY**

- 1. How does Specificity affect conceptual representation (through the lens of linguistic distributions)?
- 2. How does Specificity affect language comprehension in LLMs?
- 3. What role does word Specificity play in mechanisms of linguistic productivity and creativity?



## **CONCEPTUAL COMBINATION**

Two lexical concepts are often used together as phrases to represent a combined concept of **greater specificity**[5].



cactus fish

chocolate crocodile



LEXICALIZED COMPOUNDS

NOVEL COMPOUNDS

apple pie

[5] Dunbar & Myers (1988). Concept combination and the characterization of lexical concepts. In Hüllen, Werner; Schulze, Rainer (eds.). Understanding the lexicon: meaning, sense and world knowledge in lexical semantics, 292–302.



## CASE STUDY: COMPOUND INTERPRETATION

Noun-noun compounds have three semantic components:

a head that determines the category,

a modifier that determines how the subcategory is different from other subcategories,

and a relation between modifier and head.[5]

#### chocolate cake ≠ birthday cake ≠ coffee cake ≠ marble cake

[6] Krott, A. (2009). The Role of Analogy for Compound Words. In Blevins, J. P., Blevins, J. (Eds.), Analogy in Grammar: Form and Acquisition, 118-136.
[7] Jackendoff R. (2016). English noun-noun compounds in Conceptual Semantics. In: ten Hacken P, ed. The Semantics of Compounding, 15-37.



#### UNDERSTAND CONCEPTUAL COMBINATIONS

Do LLMs Grasp Semantic Relations in Lexicalized Noun Compounds?

Can Large Language Models Interpret Noun-Noun Compounds? A Linguistically-Motivated Study on Lexicalized and Novel Compounds

> **Giulia Rambelli** University of Bologna **Claudia Collacciani** University of Bologna

Emmanuele Chersoni The Hong Kong Polytechnic University Marianna Bolognesi University of Bologna





## **EXPERIMENT OVERVIEW: DATASET**

#### Data 668 compositional and lexicalized compounds

**ABSTRACTION** 

	compound	coarse-grained	fine-grained	Hatcher-Bourque	paraphrase (Papper 2021)
<ul> <li>concreteness ratings<sub>[9]</sub></li> </ul>	- 11630 	(11atz, 2011)	(11atz, 2011)	(Fepper, 2022)	(Fepper, 2021)
• 9 semantic relations converted	plastic bag	containment	SUBSTANCE -MATERIAL- INGREDIENT	COMP(OSITION)-P	a bag that is composed of plastic
Into paraphrases [8 10]	trash bag	containment	CONTAIN	CONT(AINMENT)-R	a bag that contains trash
	supermarket shelf	loc_part_whole	LOCATION	LOCATION	a shelf that is located in a supermarket
<ul> <li>Each compound has a correct</li> </ul>			WHOLE+		
naranhrase + 8 distractors	car deer	loc_part_whole	PART_OR	PARTONOMY	a door that is part of a car
paraprirase + o distractors			_MEMBER_OF		
plastic bag	food company	purpose	CREATE- PROVIDE- GENERATE- SELL	PRODUCTION	a company that produces food
X "a bag that produces plastic"	bank loan	causal	CREATOR- PROVIDER- CAUSE_OF	PROD(UCTION)-R	a loan that a bank produces
	research group	purpose	PERFORM& ENGAGE_IN	PURPOSE	a group intended for research
	art class	topical	TOPIC	TOPIC-R	a class that is about art
	wind turbine	topical	MEAN	US(A)G(E)-R	a turbine that uses wind

[8] Tratz, S. (2011). Semantically-enriched Parsing for Natural Language Understanding. Ph.D. thesis

[9] Muraki, et al. (2023). Concreteness Ratings for 62,000 English Multiword Expressions. Behavior Research Methods

[10] Pepper, S. (2022). Hatcher-Bourque: Towards a Reusable Classification of Semantic Relations. In Binominal Lexemes in Cross-Linguistic Perspective.



### EXPERIMENT OVERVIEW: METHODS

Models recent LLMs(base + instruct)

Task Compound interpretation as multiple-choice task:

- Choose the correct interpretation among several paraphrases (to avoid "parroting").
- Surprisal of sentences

 $S_{good}$  = "olive oil is an oil composed of olives"  $S_{bad}$  = "olive oil is an oil intended for olives"  $S(S_{good}) < S(S_{bad})$ 

$$S(w_i) = -\log P(w_i|w_0..w_{i-1})$$
$$S(sent) = \sum_{i=0}^{N} S(w_i)$$

#### • Metalinguistic prompting

🚫 LLaMA 🖟

Which is the most likely description
of "olive oil"?
1. an oil that uses olives;
2. an oil that is part of olives;
...
9. an oil that is composed of olives

MISTRA

Falcon LLM



## **OVERALL RESULTS**

#### SURPRISAL

#### PROMPTING

model	ACCURACY
BERT-large	0.262
GPT2-x1	0.338
Llama-2	0.401
Falcon	0.433
Mistral	0.403
Llama-2-7B-chat-hf	0.448
Falcon-7B-Instruct	0.38
Mistral-7B-Instruct	0.428

model	1-shot	3-shot
Llama-2-7B-chat-hf	.41	.18
Mistral-7B-Instruct	.59	.56
Falcon-7B-Instruct	.15	.14



#### SURPRISAL RESULTS



- **COMP(OSITION)-R** and **PRODUCTION** are almost perfect
  - PURPOSE, PROD-R, and TOPIC-R are mostly mistaken
- Compounds characterized by higher concreteness are interpreted more accurately (*concreteness effect*)

Rela	ation	Count	Mean Conc
CO	MP-R	85	4.47
COI	NT-R	54	4.49
LO	CATION	107	4.15
PAF	RTONOMY	16	4.58
PRC	)D-R	13	3.18
PRO	DUCTION	47	4.34
PUE	RPOSE	270	4.01
TOF	PIC-R	66	3.30
USC	G-R	10	4.24

24



#### TAKEHOME MESSAGE

- The task of interpreting lexicalized compounds is hard for LLMs
- The complexity of the task depends on the concreteness of the compound
  - Specific concrete instances are easy to decouple that specific abstract ones
- What about people?



What linguistic properties make compounds more or less difficult for humans and LLMs to interpret?



#### **3 QUESTIONS ON WORD SPECIFICITY**

- 1. How does Specificity affect conceptual representation (through the lens of linguistic distributions)?
- 2. How does Specificity affect language comprehension in LLMs?
- 3. What role does word Specificity play in mechanisms of linguistic productivity and creativity?



## **INTERPRETATION OF NOVEL CONCEPTS**

#### AVOCADO CHAIR

"A chair shaped like an avocado"

"A chair for avocados"





**RELATIONAL AMBIGUITY** 



## CASE STUDY: NOVEL COMPOUND INTERPRETATION

Interpreting a novel compound involves the conceptual and lexical systems:

- access the concepts denoted by the words and
- select a relation to form a unified conceptual representation

Hypothesis

 The on-line interpretative processing of novel nominal compounds is affected by analogous lexicalized compounds<sub>[11,12]</sub>

#### mud man

*milk man* 'A man who delivers mud' *garbage man* 'A man who collects mud'

[11] Gagné & Spalding (2006). Using Conceptual Combination Research to Better Understand Novel Compound Words. In SKASE
 Journal of Theoretical Linguistics, 3(2),9-16.
 [12] Rambelli et al. (2022). Compositionality as an Analogical Process: Introducing ANNE. In Proceedings of CogALex 2022.



## **CONCEPTUAL COMBINATIONS IN LLMs**

Are LLMs able to generalize semantic relations over novel compounds?

Can Large Language Models Interpret Noun-Noun Compounds? A Linguistically-Motivated Study on Lexicalized and Novel Compounds

> **Giulia Rambelli** University of Bologna **Claudia Collacciani** University of Bologna

Emmanuele Chersoni The Hong Kong Polytechnic University Marianna Bolognesi University of Bologna







#### EXPERIMENT OVERVIEW

**Hypothesis** People generalize (i.e., can they abstract) an implicit semantic relation that ties the two constituents of a conventional compound and transfer it to a semantically similar but novel compound

**RQ** Do LLMs generalize the semantic relation as well?

Diagnostic data 64 novel compounds

Head/modifier substituted with a hypernym from WordNet<sub>[13]</sub>



EQUIPMENT BOX GLOVE BOX GLOVE COMPARTMENT

[13] Fellbaum, C (1998) WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press.



## E2: RESULTS



SURPRISAL

	same	Head	sameMod	
model	1	3	1	3
	shot	shot	shot	shot
Llama-2-7B-chat-hf	.156	.172	.141	.219
Mistral-7B-Instruct	.578	.531	.469	.30
Falcon-7B-Instruct	.047	.063	.079	.047

- Changing the modifier is less problematic than changing the head
  - Supoptimal solution: choose PURPOSE relation
- *equipment box* -> "a box that contains equipment"
   *glove container* -> "a container intended for gloves"





#### TAKEHOME MESSAGE

- Changing the modifier is less problematic than changing the head
  - As observed in behavioral experiments
- The way we create novel compounds can tell us more about the reasons why LLMs fail
  - Instead of taking novel compounds, we manipulated lexicalized ones by varying the specificity of components (new!)
  - Evaluate human and LLMs performance!
- More studies about the productivity of language in LLMs
  - "to identify processes and mechanisms within our repertoire of computational algorithms and representations" [14]





## CONCLUSIONS

- **Specificity** is an important linguistic variable that could affect language processing.
- Stimuli for psycholinguistic and computational tasks should be balanced considering this variable
  - But there are no ratings for Specificity!



# Word Ladders is a free, educational mobile application for Android and iOS.







GET IT ON

Google Play

Download on the App Store





https://www.abstractionproject.eu/ https://site.unibo.it/abstraction/en



@Abstraction\_ERC



Marianna Bolognesi



Emmanuele Chersoni



Andrea E. Ravelli



Caterina Vllani







European Research Council



giulia.rambelli4@unibo.it

