Models of Grounded Cognition

Daniel Sabinasz – DFT summer school 22

What is grounded cognition?

- **Classical cognitive science:**
  - Higher cognitive competences (language, reasoning, planning, problem solving, ...) best explained as algorithmic processing of amodal symbols

What is grounded cognition?

- Example: Reasoning

The Porsche is parked to the left of the Dodge
The Ferrari is parked to the right of the Dodge

Therefore, the Dodge is parked to the left of the Ferrari

↓

$$\exists x \exists y \exists z \left(Porsche(x) \wedge Porsche(x) \wedge Dodge(y) \wedge Ferrari(z) \right.$$
$$\left. LeftOf(x,y) \wedge RightOf(z,y) \right) \rightarrow LeftOf(y,z)$$

What is grounded cognition?

- No empirical evidence for algorithmic processing of amodal symbols in the brain!
- Higher cognitive processes are *grounded* in sensory-motor regions of the brain
  - e.g., same brain regions involved in perception of objects of a given category also involved in reasoning (Pulvermüller, 2005)

What is grounded cognition?

- Inconsistencies with neural principles of computation (Richter et al., 2017)
  - function calls
  - random access memory
  - ...

What is grounded cognition?

- **Grounded cognition:**
  - Higher cognitive competences rely on perceptual/motor simulation using the same brain regions that are used in perception and motor action
  - (e.g., Barsalou, 2008)

What is grounded cognition?

- **Example** (Ragni & Knauff, 2013)

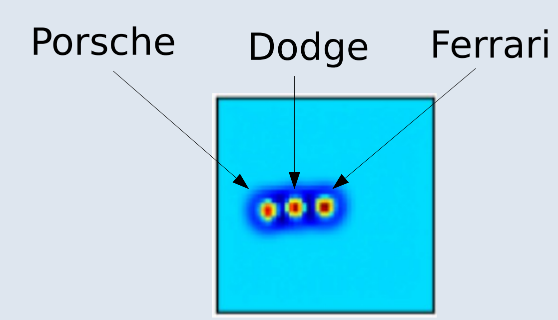
The Porsche is parked to the left of the Dodge
The Ferrari is parked to the right of the Dodge

Therefore, the Dodge is parked to the left of the Ferrari

Algorithmic proof systems

$$\exists x \exists y \exists z \left(Porsche(x) \wedge Porsche(x) \wedge Dodge(y) \wedge Ferrari(z) \right. \\ \left. LeftOf(x,y) \wedge RightOf(z,y) \right) \rightarrow LeftOf(y,z)$$

Spatial layout models



Ragni & Knauff (2013), Kounatidou, Richter, & Schöner (2018)

What is grounded cognition?

- **Example** (Ragni & Knauff, 2013)

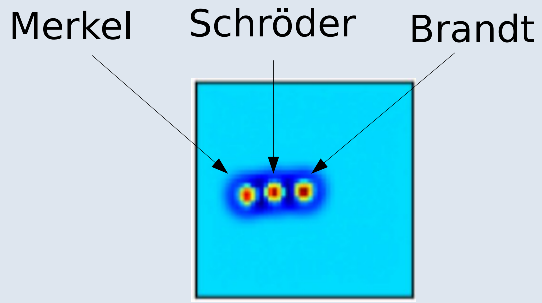
Willy Brandt was more popular than Gerhard Schröder
Gerhard Schröder was more popular than Angela Merkel

Therefore, Willy Brandt was more popular than Angela Merkel

Algorithmic proof systems

$$\exists x \exists y \exists z \left(Porsche(x) \wedge Porsche(x) \wedge Dodge(y) \wedge Ferrari(z) \right. \\ \left. LeftOf(x,y) \wedge RightOf(z,y) \right) \rightarrow LeftOf(y,z)$$

Spatial layout models



Ragni & Knauff (2013), Kounatidou, Richter, & Schöner (2018)

What is grounded cognition?

- Many of our abstract concepts are metaphorically related to more basic concepts (Lakoff and Johnson, 1980; Hofstadter and Sander, 2013)
  - e.g., up for happy
    down for sad

What is grounded cognition?

- **Hypothesis:** Many of our abstract concepts supervene on visuo-spatial concepts

- → Towards models of higher cognition (language understanding, reasoning, problem solving, ...) from models of visuo-spatial cognition

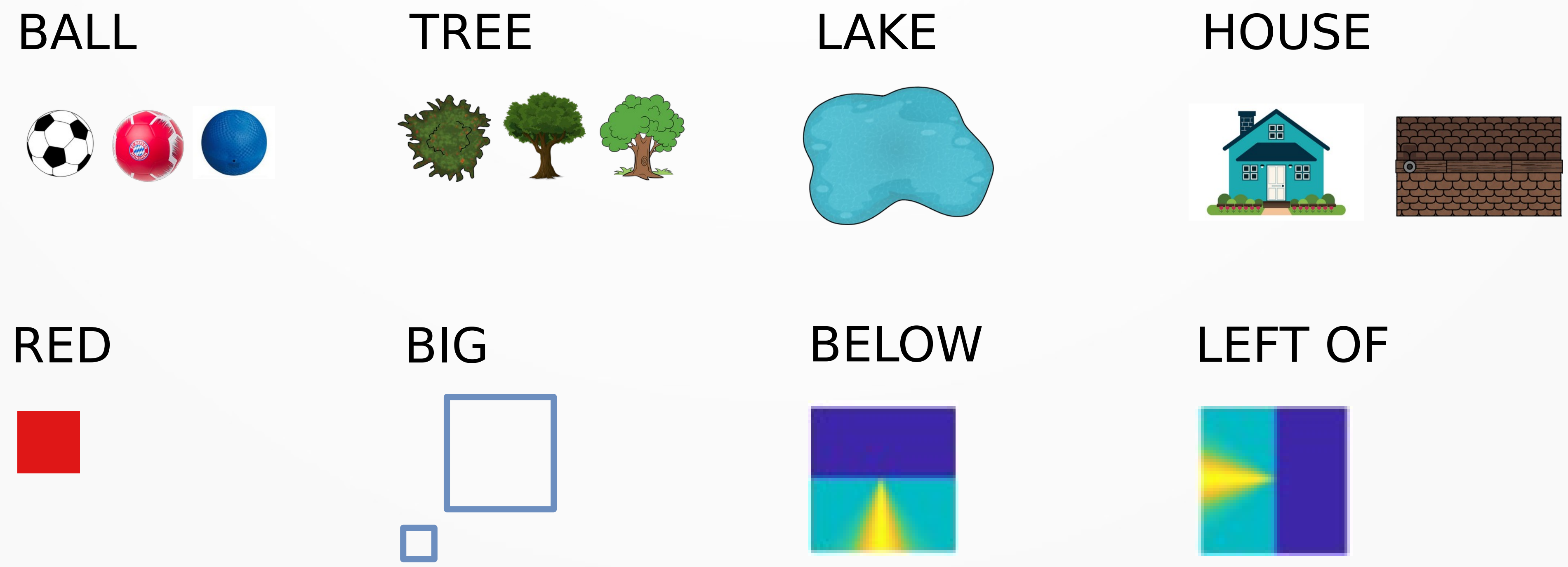
Background from cognitive linguistics

- **Cognitive linguistics** combines knowledge from
  - linguistics
  - psychology
  - neuroscience

  to infer the mechanisms that underlie language understanding / thought

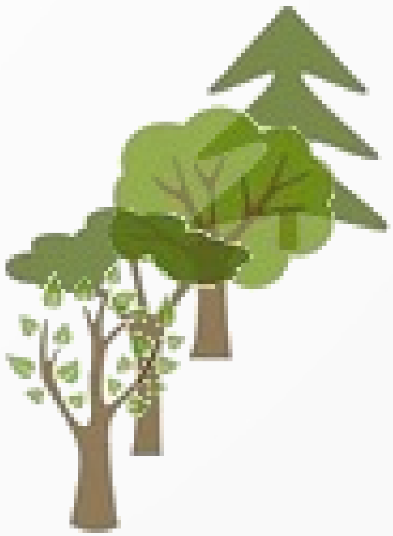
Background from cognitive linguistics

- **Conceptualist semantics:**
  Humans understand a word by virtue of possessing a concept denoted by the word



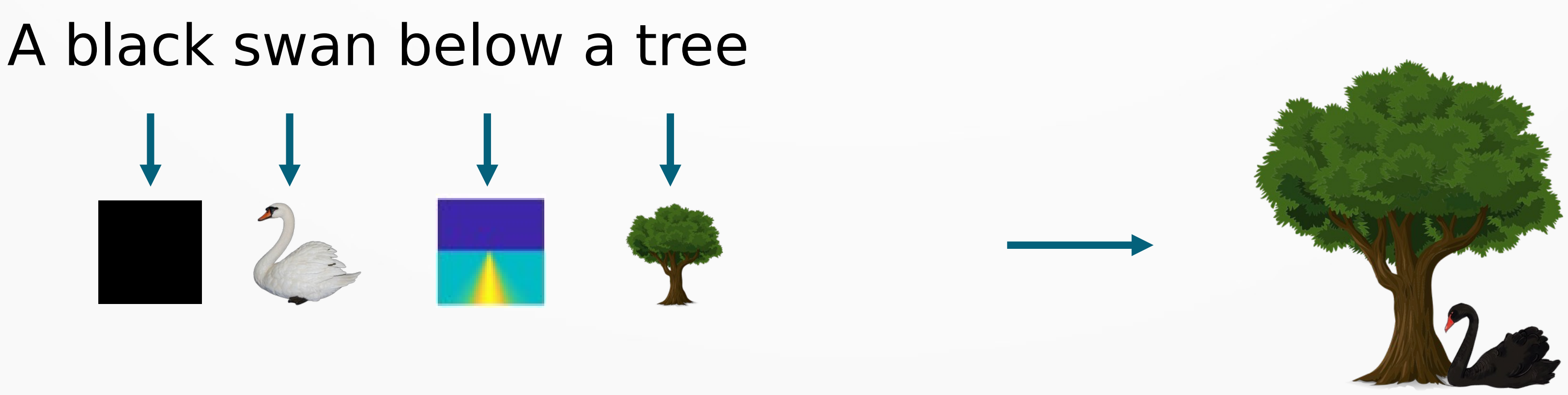
Background from cognitive linguistics

- **Prototype theory**: concepts are long-term memory representations of a prototypical instance of the category
- e.g., TREE prototype
  - Long-term memory representation
  - Involved in
    - understanding the word “tree“
    - categorizing something as a tree
    - imagining trees
    - reasoning about trees
    - behaving towards trees



Background from cognitive linguistics

- **Compositional semantics:**
  Humans understand a phrase by
  - activating the concepts denoted by the individual words
  - combining those concepts in accord with syntactic arrangement



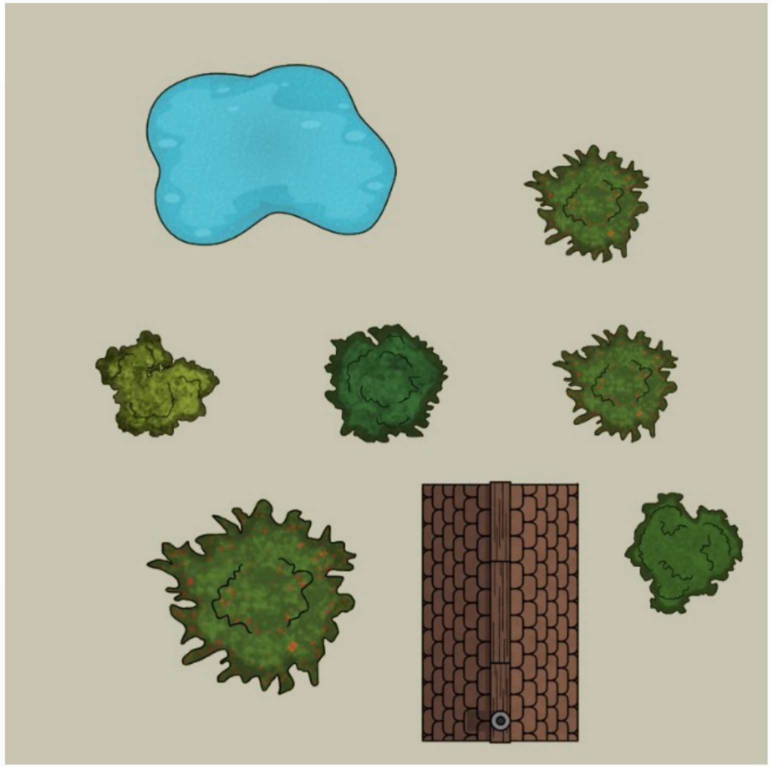
Background from cognitive linguistics

- **Noun phrase:** A phrase describing an object
- e.g.,
  - the tree
  - the small tree
  - the tree to the right of the house

prepositional phrase

Background from cognitive linguistics

- **Nested noun phrase**
- e.g.,
  - the tree below the lake
  - the tree to the right of the tree below the lake
  - the tree below the lake and above the house
  - ...



Background from cognitive linguistics

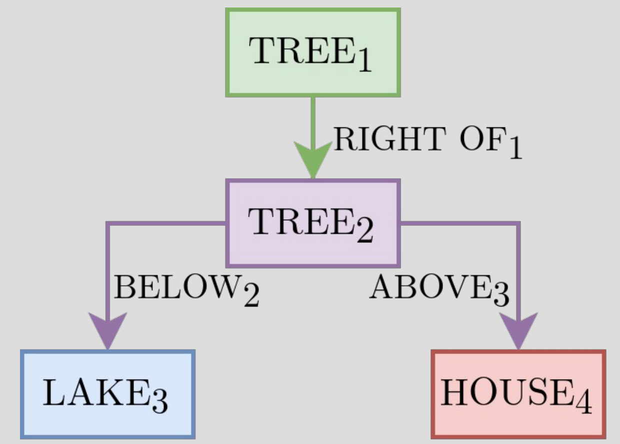
- **Conceptual structure** (Jackendoff, 2002)
  - cognitive representation
  - characterizes the meaning of a phrase as a combination of concepts

Phrase

“the tree to the right of the tree below the lake and above the house”

natural language processing →

Conceptual structure



Background from cognitive linguistics

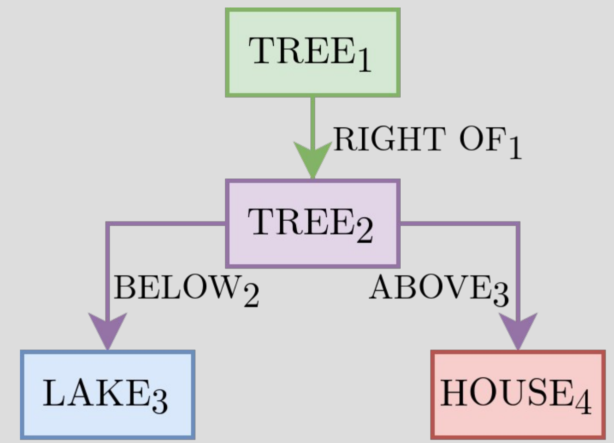
- CS of nested noun phrase must specify
  - which objects there are
  - which concepts characterize them
  - which relationships hold among the objects

Phrase

"the tree to the right of the tree below the lake and above the house"

natural language processing →

Conceptual structure

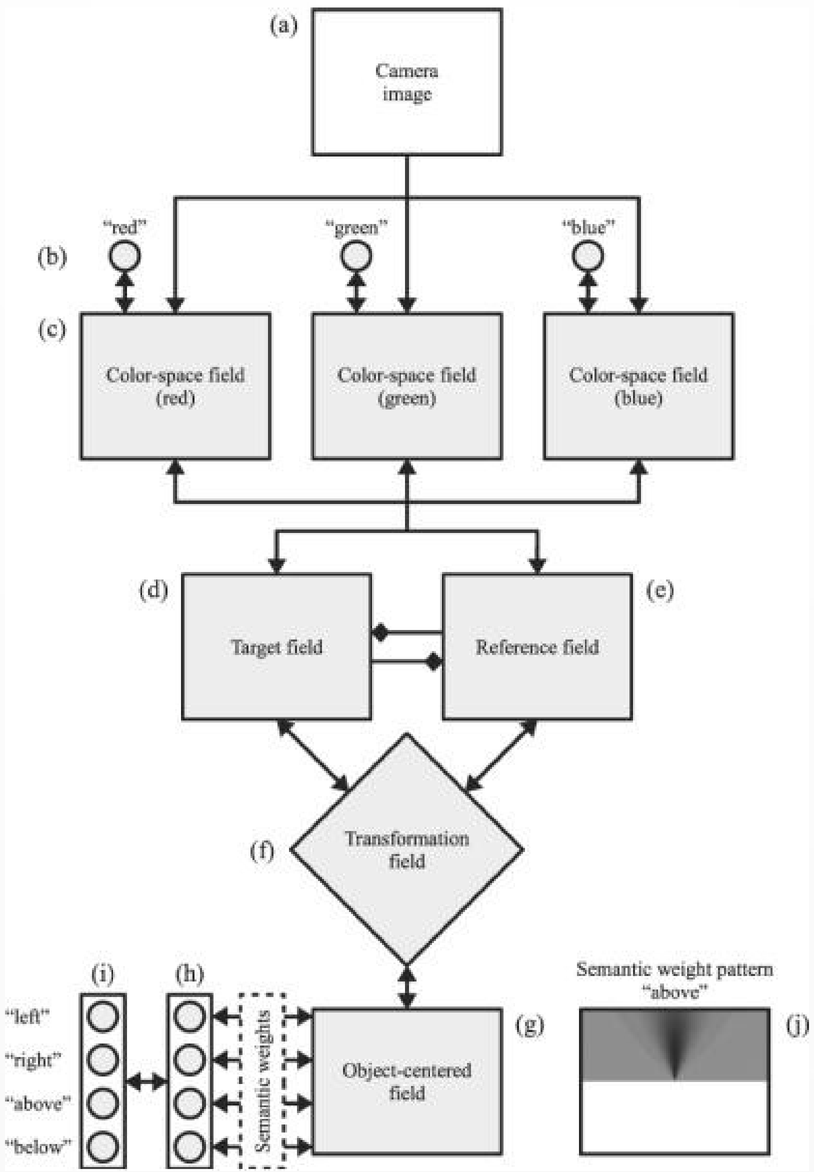


Models of grounded cognition

- Lipinski et al. (2012)
  - Where is the green object relative to the red object?
    → to the right

  - Which object is above the blue object?
    → the red object

Models of grounded cognition

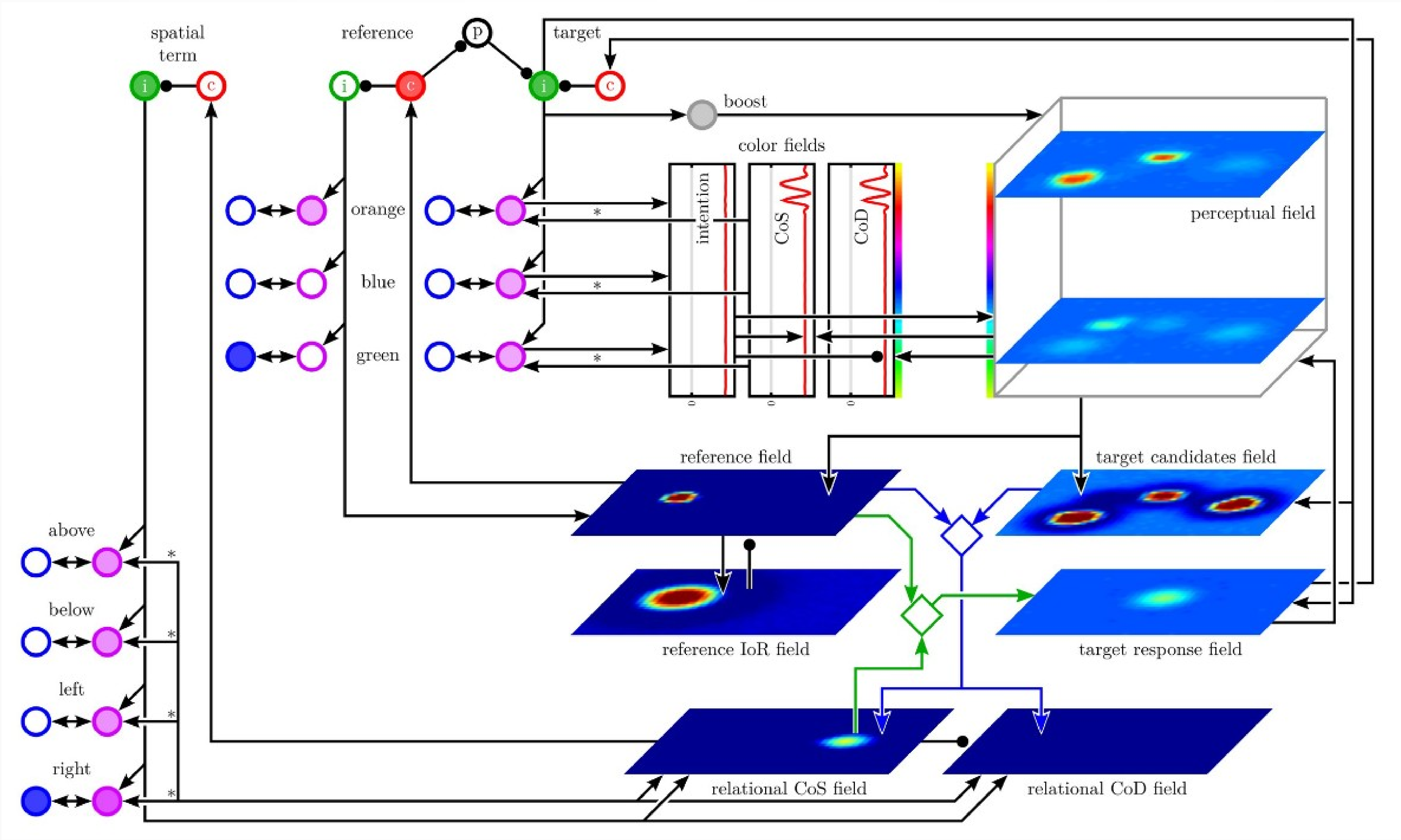


Models of grounded cognition

- Richter et al. (2014)
  - **Grounding** a noun phrase with a single prepositional phrase:
  - e.g., “the red object to the left of the green object“

  - Requires autonomous hypothesis testing!

Models of grounded cognition



Sabinasz & Schöner (2022)

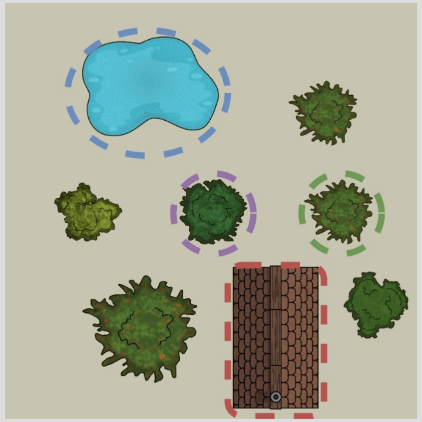
- Neural process model that can search the object referenced by a given nested noun phrase in the visual input

Phrase

“the tree to the right of the tree below the lake and above the house”

compositional search →

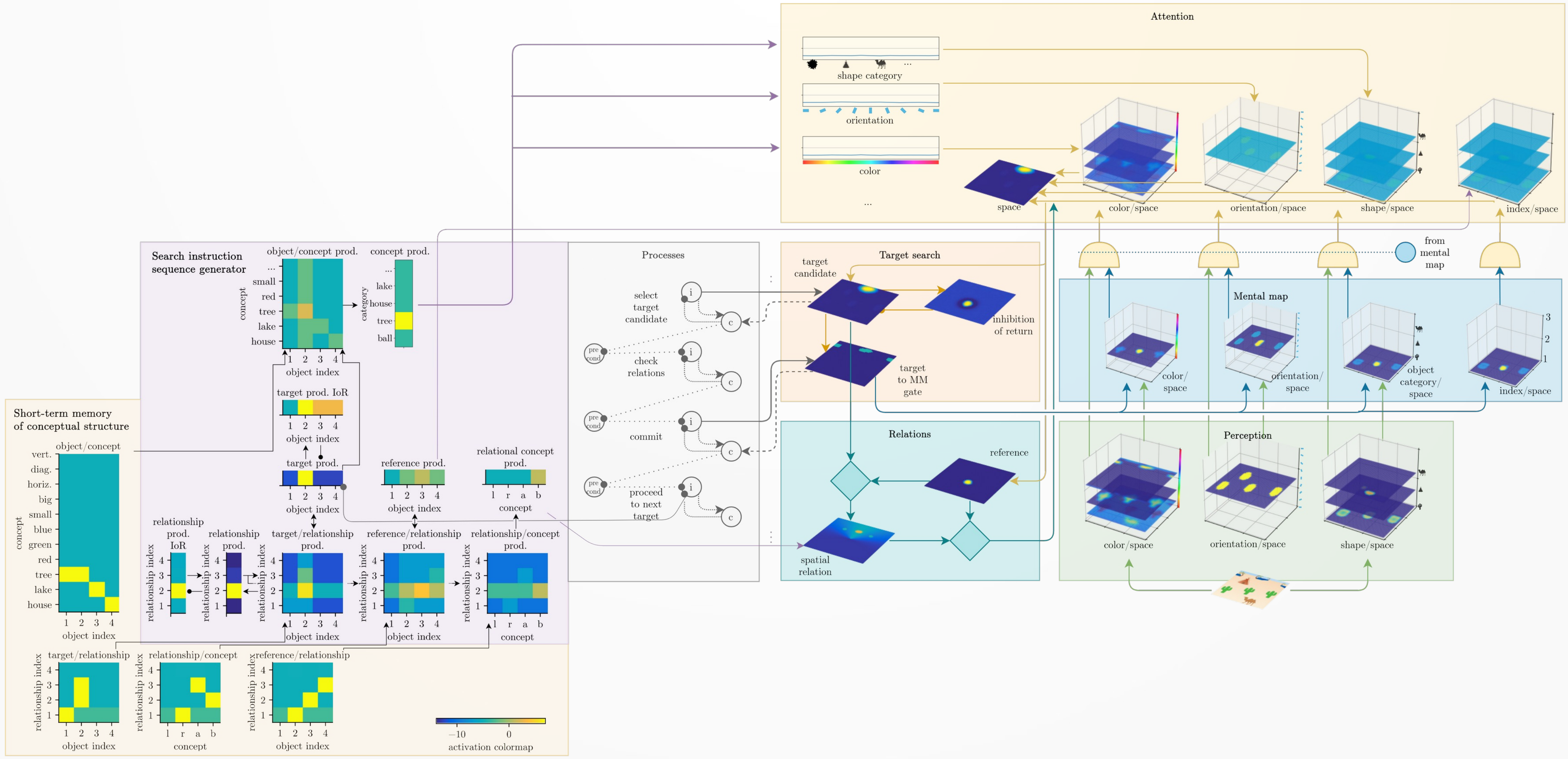
Visual input



Motivation

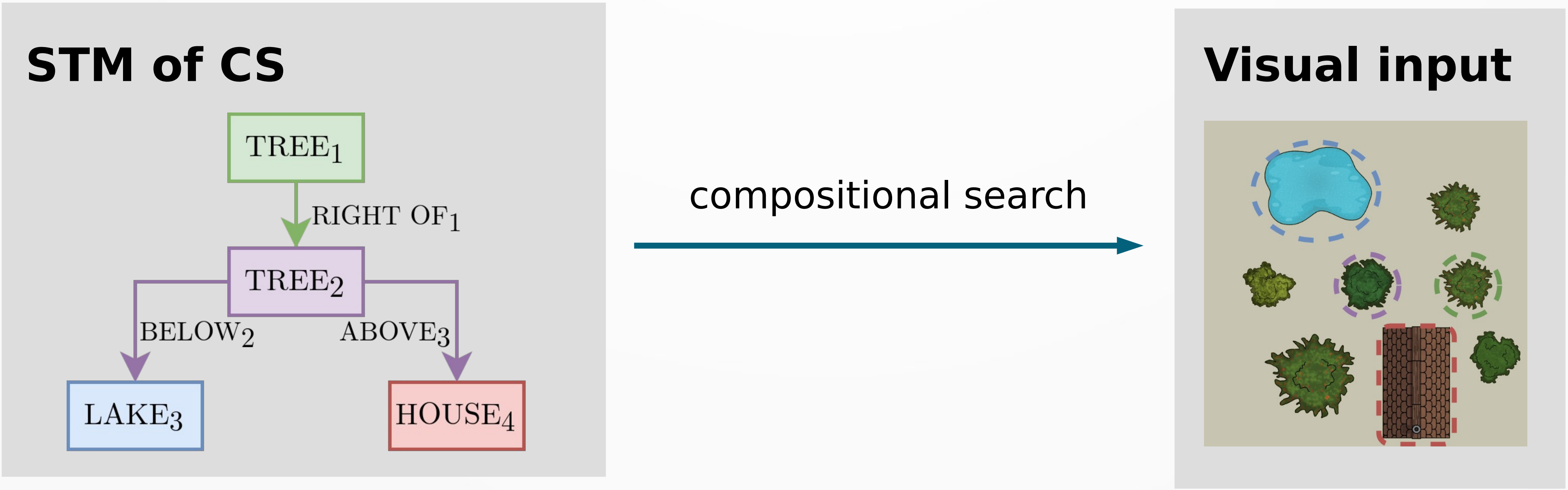
- May serve as a blueprint for models that understand other grammatically complex language
- Solves many challenges
  - Link to sensory input
  - Neural representations of concepts
  - Neural processes for combining concepts
  - Neural processes for relational reasoning
  - Neural short-term memory of conceptual structure

Sabinasz & Schöner (2022)



STM of conceptual structure

- Assume that conceptual structure is represented as a short-term memory and directs the visual search process



STM of conceptual structure

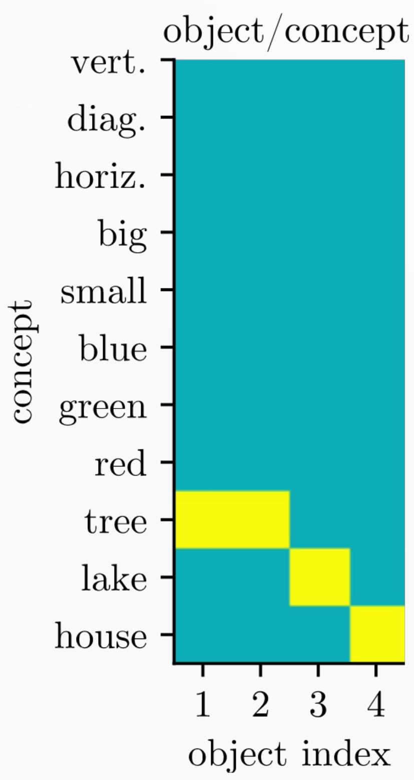
- Any neural STM of conceptual structure must address **Jackendoff's challenges**
  - The problem of 2:
    e.g., "the small tree above the big tree"
  - The massiveness of the binding problem:
    e.g., "the lake above the tree above the house"

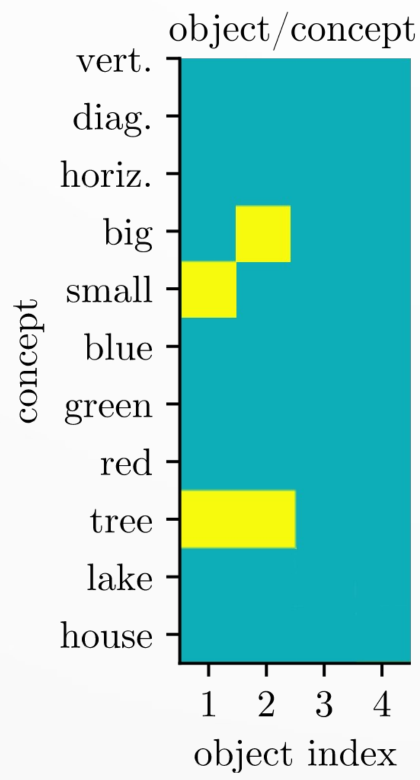
STM of conceptual structure

- Addressing Jackendoff‘s challenges
  - Assume that language pre-processing embeds objects into a discrete index dimension
  - “the tree 1 right of the tree 2 below the lake 3 and above the house 4”

STM of conceptual structure

- The index dimension may serve as a binding agent, enabling a neural STM to encode which concepts characterize a given object





Analogous to **feature integration theory**, in which space serves as a binding agent (Treisman & Gelade, 1980)

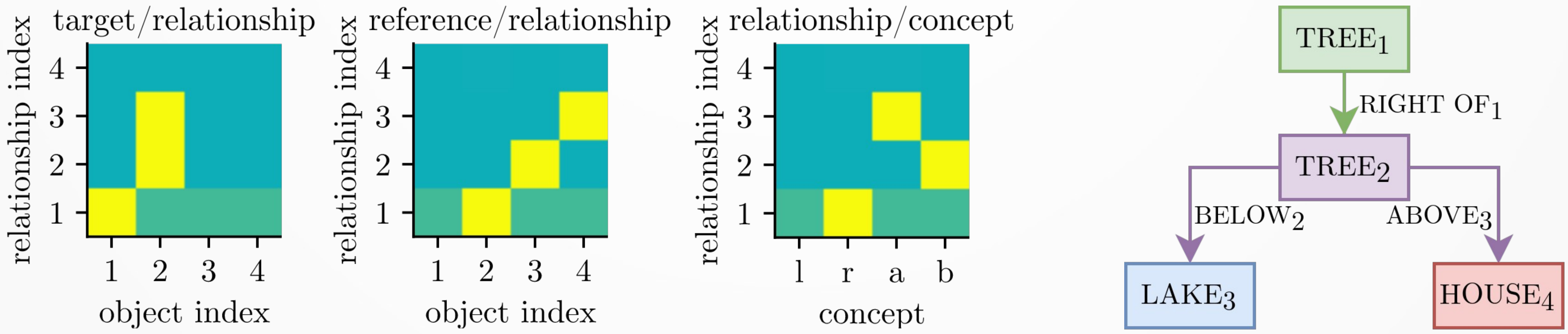
Discrete neural field

$$\tau \dot{u}(x,t) = -u(x,t) + h + s(x,t) + c_{\text{exc}} \cdot \sigma(u(x,t)) - \sum_{x' \neq x} c_{\text{inh}} \cdot \sigma(u(x',t))$$

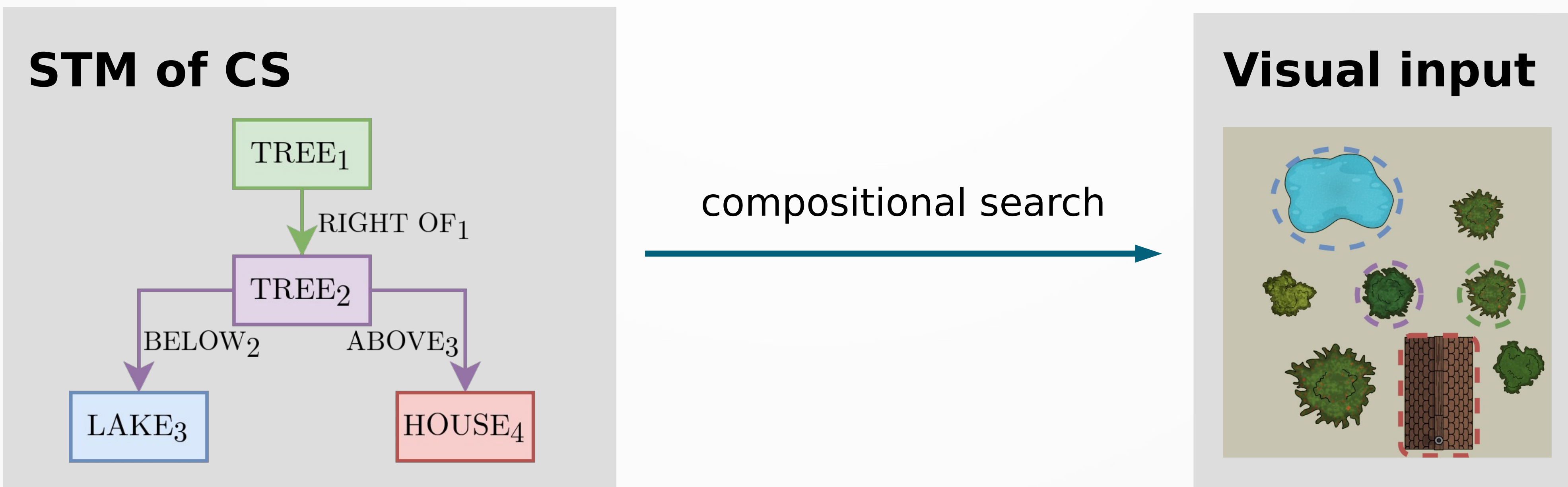
STM of conceptual structure

- Also assume that language pre-processing embeds relationships into an index dimension
  - "the tree right of 1 the tree below 2 the lake and above 3 the house"

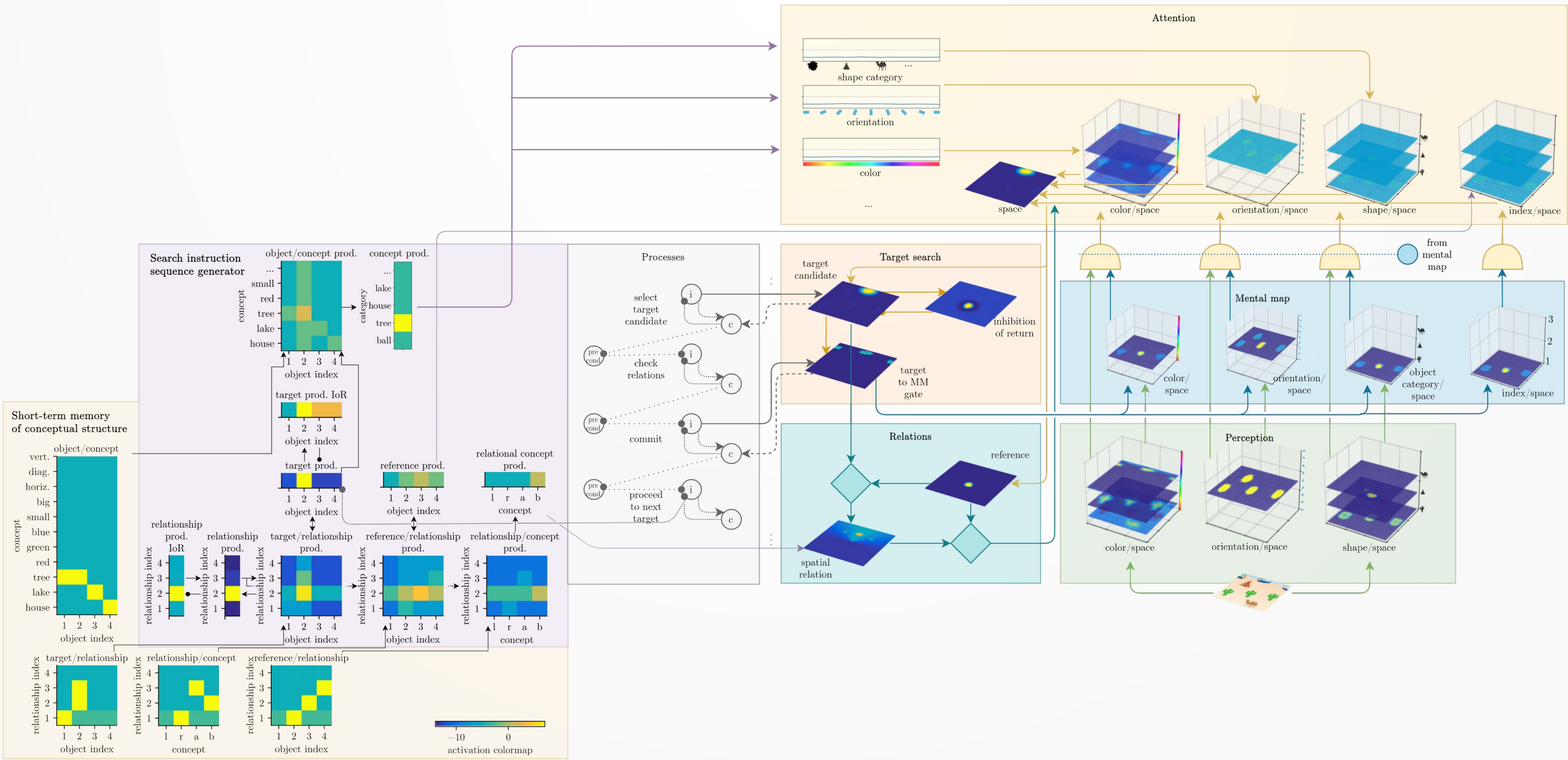
STM of conceptual structure



STM of conceptual structure

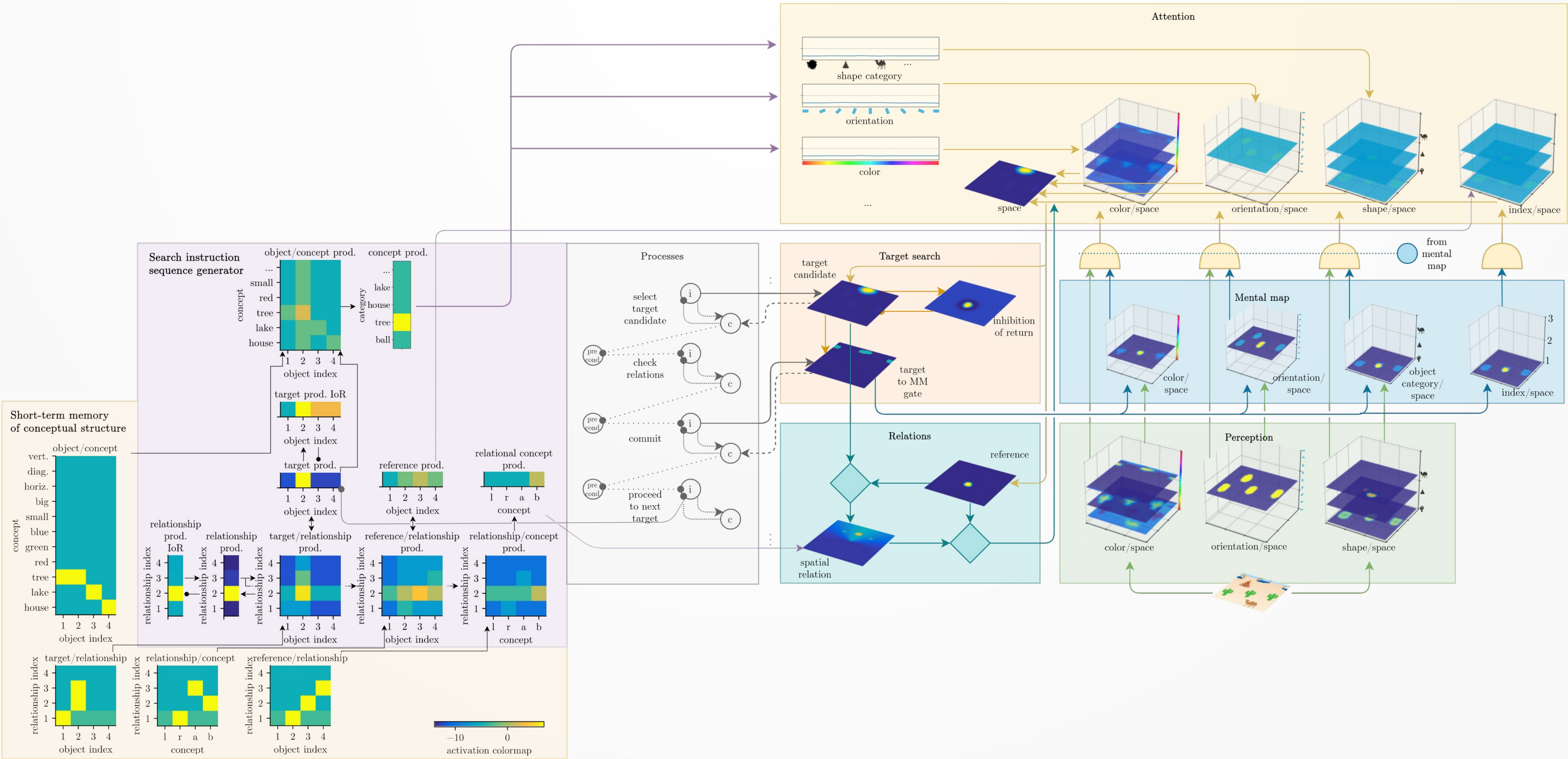


Sabinasz & Schöner (2022)

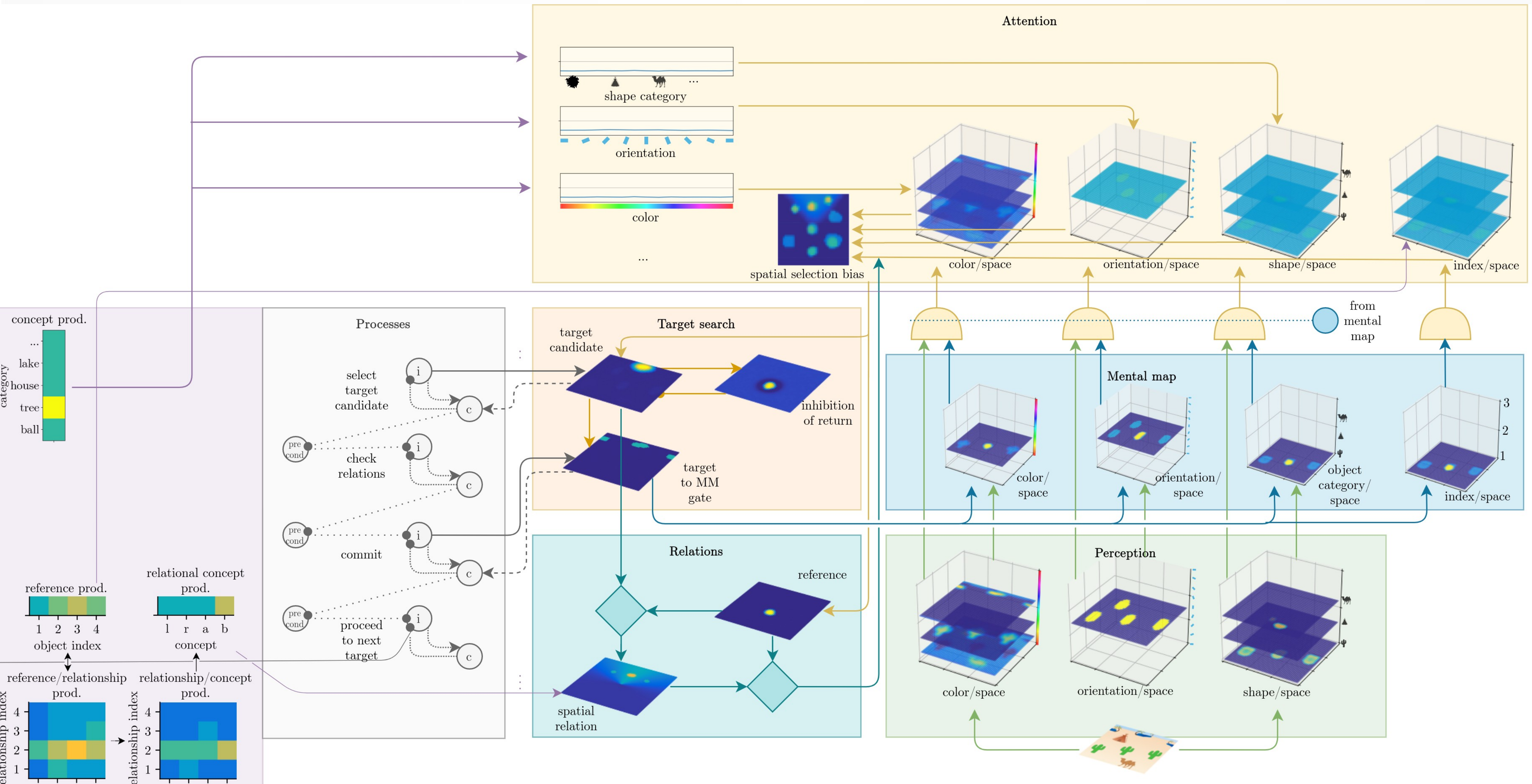


Video: Interfacing conceptual structure with compositional search

Compositional search



Compositional search

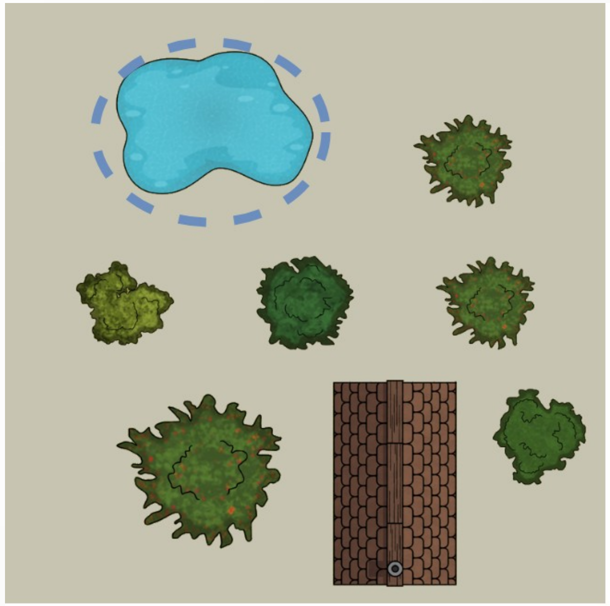


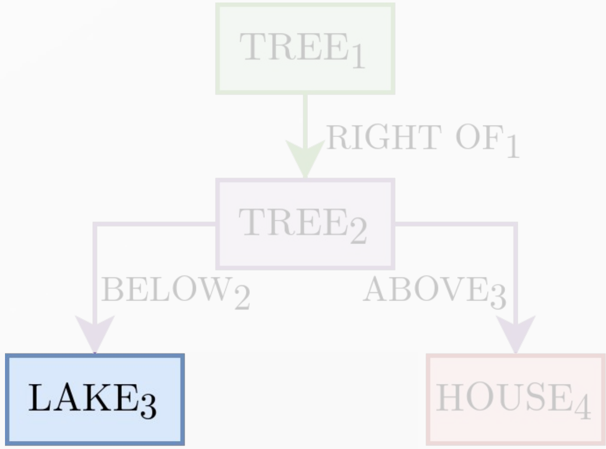
Time structure of compositional search

- Objects sometimes attended in order of mention, but not necessarily (Tanenhaus et al., 1995; Burigo & Knoeferle, 2015)
- Reordering may occur
- Plausibly guided by efficiency considerations
- e.g., select an object only once the related objects have been found and memorized

Time structure of compositional search

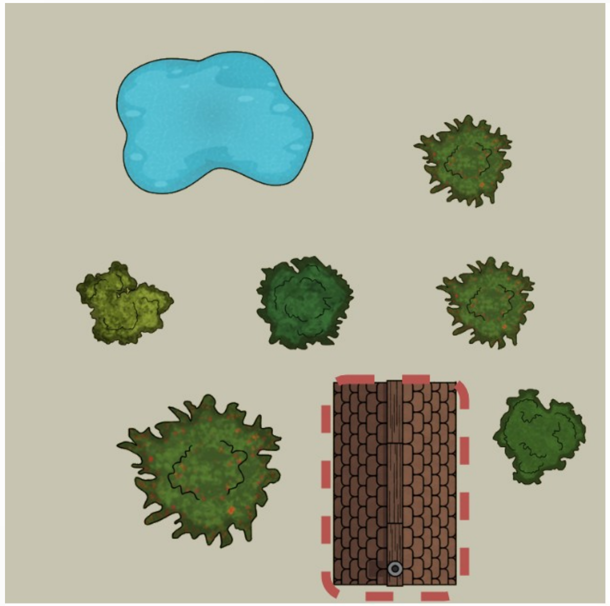
- find the lake

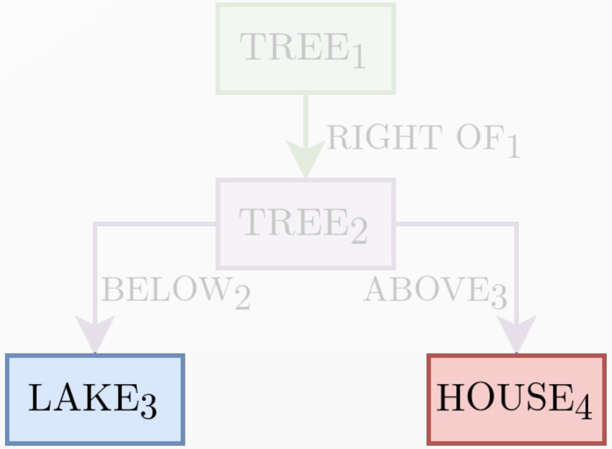




Time structure of compositional search

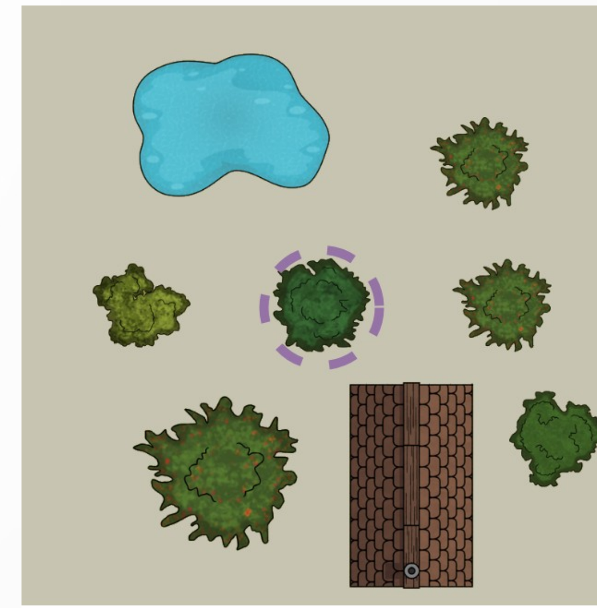
- find the lake
- find the house

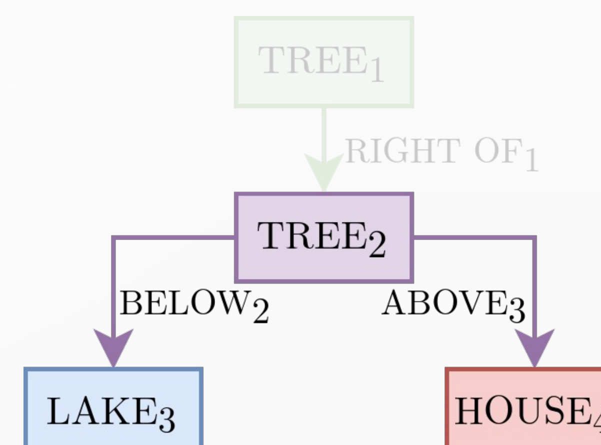




Time structure of compositional search

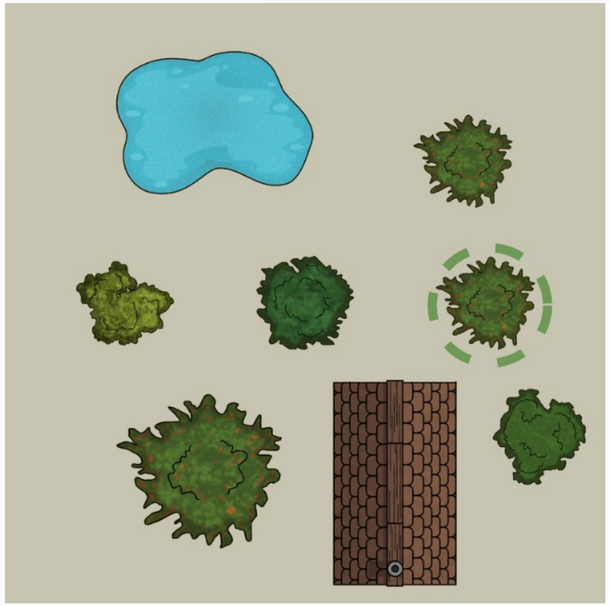
- find the lake
- find the house
- find the tree below the lake and above the house

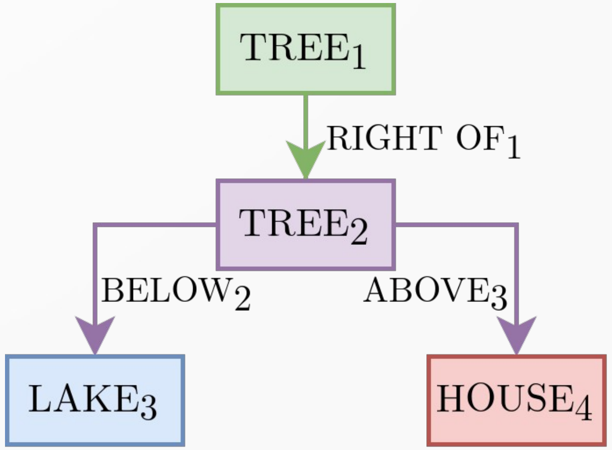




Time structure of compositional search

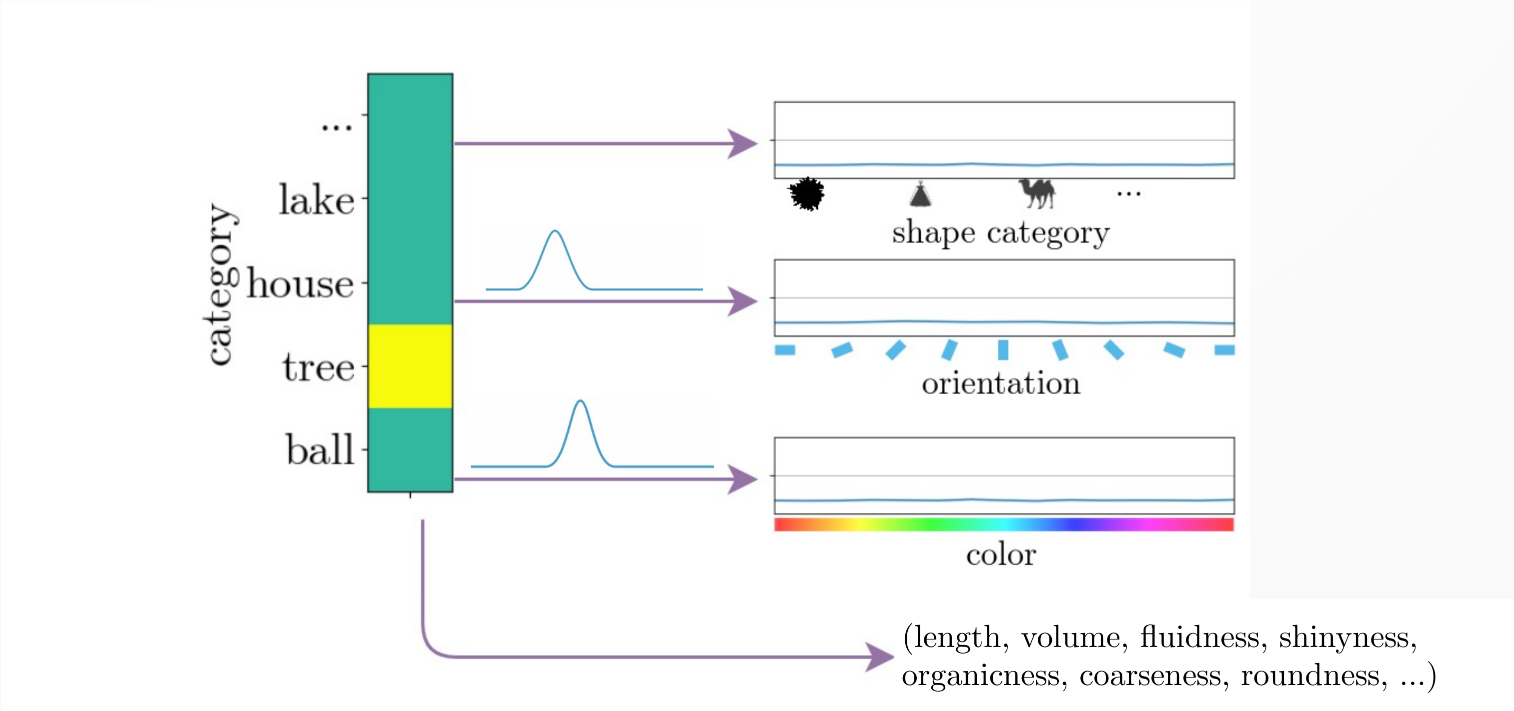
- find the lake
- find the house
- find the tree below the lake and above the house
- find the tree to the right of that tree





DFT models of grounded cognition

- Object concepts are synaptic weight patterns which encode prototype distributions in feature spaces (Johnson, Spencer, & Schöner, 2008; Sabinasz, 2019)



Discussion

- Existing methods implement algorithmic tree traversal and, therefore, make use of pointers and recursive function calls (e.g., Brown, Buntschuh, & Wilpon, 1992; Nagao & Rekimoto, 1995; Gorniak & Roy, 2004)
- Not clear how this could be realized by neural processes

Discussion

- Search order emerges from interactions that bias competitive selection in favor of objects whose reference objects have already been found
- The relational dependency structure can thus affect the order without requiring algorithmic tree traversal methods

Discussion

- Neural dynamic implementation of vector-symbolic architectures (Stewart & Eliasmith, 2012) also address Jackendoff's challenges and enable coupling to perceptual and motor processes (Eliasmith, 2013)
- Short-term memory not stable against noise-induced drift

Conclusion

- Presented neural dynamic process model that can perceptually ground a nested noun phrase
- Consistent with neural principles formalized in DFT
- STM of conceptual structure
  - Filled by language system
  - Provides input to neural process that generates a sequence of searches that together successfully and efficiently find the described object

Thanks for your attention!

References

- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. Cognition, 73(3), 247–264.
- Amari, S.-i. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. Biological Cybernetics, 27(2), 77–87.
- Barsalou, L. W. (1999). Perceptual symbol systems. Behavioral and Brain Sciences, 22(4), 577–609.
- Barsalou, L. W. (2008). Grounded cognition. Annual Review of Psychology, 59, 617–645.
- Barsalou, L. W. (2017). Cognitively plausible theories of concept composition. In J. A. Hampton & Y. Winter (Eds.),
- Compositionality and concepts in linguistics and psychology (pp. 9–30). Springer International Publishing.

References

- Brown, M. K., Buntschuh, B. M., & Wilpon, J. G. (1992). Sam: A perceptive spoken language-understanding robot. IEEE Transactions on Systems, Man, and Cybernetics, 22(6), 1390–1402.
- Burigo, M., & Knoeferle, P. (2015). Visual attention during spatial language comprehension. PLoS ONE, 10(1).
- Chomsky, N. (1968). Language and mind. New York: Harcourt Brace & World.
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: a new methodology for the real-time investigation of speech perception, memory, and language processing. Cognitive Psychology, 6, 84–107.
- Eliasmith, C. (2013). How to build a brain: A neural architecture for biological cognition. Oxford University Press.

References

- Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden-path sentences. Journal of Memory and Language, 30(6), 725–745.
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. Cognition, 28(1-2), 3–71.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. Cognitive Psychology, 14(2), 178–210.
- Gayler, R. W. (2003). Vector symbolic architectures answer jackendoff's challenges for cognitive neuroscience. ICCS/ASCS International Conference on Cognitive Science, 133–138.
- Gorniak, P., & Roy, D. (2004). Grounded semantic composition for visual scenes. Journal of Artificial Intelligence Research, 21, 429–470.

References

- Grossberg, S., et al. (1978). Competition, decision, and consensus. Journal of Mathematical Analysis and Applications, 66(2), 470–493.
- Jackendoff, R. (2002). Foundations of language: Brain, meaning, grammar, evolution. Oxford University Press.
- Johnson, J. S., Simmering, V. R., & Buss, A. T. (2014). Beyond slots and resources: Grounding cognitive concepts in neural dynamics. Attention, Perception, & Psychophysics, 76, 1630-1654.
- Kounatidou, P., Richter, M., & Schöner, G. (2018). A neural dynamic architecture that autonomously builds mental models. In CogSci.
- Lakoff, G., & Johnson, M. (1999). Philosophy in the flesh: The embodied mind and its challenge to western thought. New York: Basic Books.
- Machery, E. (2009). Doing without concepts. Oxford University Press.
- Meng, M., & Bader, M. (2000). Ungrammaticality detection and garden path strength: Evidence for serial parsing. Language and Cognitive processes, 15(6), 615–666.

References

- Grossberg, S., et al. (1978). Competition, decision, and consensus. Journal of Mathematical Analysis and Applications, 66(2), 470–493.
- Jackendoff, R. (2002). Foundations of language: Brain, meaning, grammar, evolution. Oxford University Press.
- Johnson, J. S., Simmering, V. R., & Buss, A. T. (2014). Beyond slots and resources: Grounding cognitive concepts in neural dynamics. Attention, Perception, & Psychophysics, 76, 1630-1654.
- Kounatidou, P., Richter, M., & Schöner, G. (2018). A neural dynamic architecture that autonomously builds mental models. In CogSci.
- Lakoff, G., & Johnson, M. (1999). Philosophy in the flesh: The embodied mind and its challenge to western thought. New York: Basic Books.
- Machery, E. (2009). Doing without concepts. Oxford University Press.

References

- Meng, M., & Bader, M. (2000). Ungrammaticality detection and garden path strength: Evidence for serial parsing. Language and Cognitive processes, 15(6), 615–666.
- Nagao, K., & Rekimoto, J. (1995). Ubiquitous talker: Spoken language interaction with real world objects. In Proceedings of the international joint conference on artificial intelligence (pp. 1284–1290).
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. Nature reviews neuroscience, 6(7), 576-582.
- Ragni, M., & Knauff, M. (2013). A theory and a computational model of spatial reasoning with preferred mental models. Psychological review, 120(3), 561.
- Richter, M., Lins, J., & Schöner, G. (2017). A Neural Dynamic Model Generates Descriptions of Object-Oriented Actions. Topics in Cognitive Science, 9(1), 35-47.

References

- Sabinasz, D. (2019). A neural dynamic model for the perceptual grounding of combinatorial concepts. Unpublished master’s thesis, Ruhr University Bochum, Germany. (preprint accessible via Open Science Foundation at https://osf.io/mra26)
- Sabinasz, D., Richter, M., Lins, J., & Schöner, G. (2020). Grounding spatial language in perception by combining concepts in a neural dynamic architecture. In Proceedings of the 42th annual conference of the cognitive science society. Austin, TX: Cognitive Science Society.
- Sabinasz, D., & Schöner, G. (2022). A Neural Dynamic Model Perceptually Grounds Nested Noun Phrases. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 44, No. 44).
- Sandamirskaya, Y., & Schöner, G. (2010). An embodied account of serial order: How instabilities drive sequence generation. Neural Networks, 23(10), 1164–1179.

References

- Schöner, G., Spencer, J. P., & the DFT Research Group. (2015). Dynamic Thinking: A Primer on Dynamic Field Theory. New York, NY: Oxford University Press.
- Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. Artificial Intelligence, 46(1-2), 159–217.
- Stewart, T., & Eliasmith, C. (2012). Compositionality and biologically plausible models. In M. Werning, W. Hinzen, & E. Machery (Eds.), The Oxford Handbook of Compositionality. Oxford: Oxford University Press.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. Science, 268(5217), 1632–1634.
- Van der Velde, F., & De Kamps, M. (2006). Neural blackboard architectures of combinatorial structures in cognition. Behavioral and Brain Sciences, 29(1), 37–70.