

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

A Neural Dynamic Model Perceptually Grounds Nested Noun Phrases

Permalink

<https://escholarship.org/uc/item/88t5317c>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Sabinasz, Daniel
Schöner, Gregor

Publication Date

2022

Peer reviewed

A Neural Dynamic Model Perceptually Grounds Nested Noun Phrases

Daniel Sabinasz (daniel.sabinasz@ini.rub.de)

Gregor Schöner (gregor.schoener@ini.rub.de)

Institut für Neuroinformatik, Ruhr-Universität Bochum, 44780 Bochum, Germany

Abstract

We present a neural dynamic model that perceptually grounds nested noun phrases, i.e., noun phrases that contain further (possibly also nested) noun phrases as parts. The model receives input from the visual array and a representation of a noun phrase from language processing. It organizes a search for the denoted object in the visual scene. The model is a neural dynamic architecture of interacting neural populations which has clear interfaces with perceptual processes. It solves a set of theoretical challenges, including the problem of keeping a nested structure in short-term memory in a way that solves the problem of 2 and massive binding problem emphasized by Jackendoff (2002). The model organizes a search for the objects that are referenced in that structure. We motivate the model, demonstrate simulation results, and discuss how it differs from related models.

Keywords: neural process model; grounding; noun phrases; concepts; conceptual structure; dynamic field theory

Introduction

When hearing a complex linguistic expression, we understand its meaning by virtue of understanding the meanings of the individual words, often referred to as “concepts”, and combining those concepts in accord with the syntactic arrangement. What are the neural processes that bring this about? We ask this question committed to the embodiment stance that language understanding depends on perceptual and motor representations, and on the neural architecture that reflects an evolutionary and developmental history in which behavior is generated while the body is situated in an environmental context (Lakoff & Johnson, 1999; Barsalou, 1999). Yet we also aim to address the flexibility of the language faculty that has been described as its “recursive nature”, its “productivity” or its “creative aspect” (e.g., Chomsky, 1968; Fodor & Pylyshyn, 1988; Jackendoff, 2002), i.e., the ability to flexibly join atomic linguistic units into molecular linguistic units, and to join molecular linguistic units into more complex molecular linguistic units. This feature of language is evident in our ability to understand nested noun phrases (Figure 1 a-c). It raises challenges for neural process accounts that pertain to flexibly encoding items and relationships among items in short-term memory (STM), and to organizing a sequential search in accord with that encoding.

We propose a neural process model that can search the object referenced by a given nested noun phrase in the visual array. The account is based on Dynamic Field Theory (DFT; Schöner, Spencer, & the DFT Research Group, 2015),

a framework for building neural process models using recurrent neural networks. The model solves a set of challenges. First, such a model must account for how cognitive states are linked to sensory inputs and potentially to motor outputs. Second, it must be consistent with neural principles of computation, avoiding algorithmic elements that are not neurally realizable. Third, it must explain how a STM of the combinatorial structure of a phrase can be built and then read from step by step. That STM must encode descriptions of the objects and their relationships, as described in the noun phrase. Fourth, the model must organize a search for objects in the visual input in accordance with that structure. According to the grounded cognition stance (Barsalou, 2008), the neural processes supporting that search must overlap with the neural processes for perception. To the best of our knowledge, the model we present solves for the first time all of these challenges at once.

Hypotheses

Conceptual structure

The conceptual structure of a linguistic expression is a cognitive representation that characterizes the logical meaning of the expression as a combination of (ungrounded/symbolically characterized) concepts (Jackendoff, 2002). For nested noun phrases, it must specify (among other things) which objects there are, which concepts characterize them, and which relationships they bear to each other (Figure 1b).

Jackendoff hypothesizes that higher cognitive competences like reasoning and planning are underwritten by conceptual structure. This is given support by the fact that humans use conceptual combinations in non-linguistic problem-solving and goal-achievement (Barsalou, 2017). Grounding conceptual structures thus seems to be possible in non-verbal conceptual thinking, not only when guided by a linguistic expression. Both the processes of grounding linguistic expressions and of grounding non-linguistic conceptual structures must be capable of handling dependencies such as the one depicted in Figure 1b. We hypothesize that these processes are shared across the two domains, based on a single neural substrate. This is plausible if language input is first analyzed for its latent conceptual structure, as commonly assumed.

A neural mechanism for representing conceptual structure must address known challenges. Jackendoff’s “problem of

2” is exemplified by the phrase “the small tree above the big tree” in which the word “tree” occurs twice, once combined with “small”, once combined with “big”. How may a neural representation of conceptual structure encode that there are two trees, one of them small, the other one big? Relatedly, in the phrase “the lake above the tree above the house”, the word “tree” is the object of the first preposition but is itself further described by the second preposition. How may a neural representation encode that there is a tree that is above the house, and above which there is the lake? This exemplifies Jackendoff’s “massiveness of the binding problem”, as it requires the ability to flexibly bind the representation of an object to multiple other representations.

Time structure of compositional search

Perceptually grounding a noun phrase entails a search for objects in the visual array that are characterized by certain concepts and by certain relations to other objects, as specified in the conceptual structure. The time structure of this search is subject to a number of constraints which suggest that humans take the conceptual structure into account when they order the search process. Experimental studies demonstrate that objects in the visual array are sometimes attended in the order in which they are mentioned, but not necessarily so (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Altmann & Kamide, 1999; Burigo & Knoefler, 2015). A reordering may occur. Some orders are more likely than others, although the criteria for choosing an order are not well understood. Some orders may lead to more efficient search than others, and such efficiency considerations may plausibly affect the ordering. In the example of Figure 1a, a possible efficient search strategy would select a candidate object only once the objects to which it is related have been found and memorized (e.g., find the lake and the house first, then find a tree that is below the lake and above the house, then find another tree to the right of that tree). Such influence of conceptual structure on search would make it plausible that conceptual structure is explicitly represented in the brain and affects the organization of the search.

Short-term memory of conceptual structure

To control the search process, the conceptual structure must be represented as a STM that is stable on the time scale over which multiple candidate selections and relation evaluations take place. Otherwise, a syntactic re-analysis of the phonological loop would have to occur every time a new object is attended, which would predict, implausibly, that the time between two successive attentional selections scales with the length of the noun phrase. Stability of STM can be achieved through recurrent self-excitatory and other-inhibitory neural interactions (Grossberg et al., 1978).

Search processes may start before the phrase has been completely encoded in STM. Humans start searching referents even after hearing only the first few words of a phrase (Tanenhaus et al., 1995) while simultaneously building a syntactic interpretation that they gradually refine as more words

are processed (Frazier & Rayner, 1982; Ferreira & Henderson, 1991; Meng & Bader, 2000). Thus, the STM of conceptual structure must enable simultaneous reading and writing.

Dynamic Field Theory

Our model is built from dynamic neural fields (Amari, 1977; Schöner et al., 2015), each of which models the activation of a population of neurons at time t as an activation function $u(x, t)$ defined over a feature dimension x . In the present model, x is a discrete dimension, and u evolves according to

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

τ is the time scale, h the negative resting level, s the input. σ is a sigmoidal transfer function. The second line formalizes self-excitation of a field location with strength c_{exc} and pairwise inhibition between field locations with strength c_{inh} .

Initially, the activation is at resting level h . When small input is supplied, the activation tracks an attractor at $h + s(x, t) < 0$. When sufficiently large input is supplied at some location, that attractor becomes unstable and the field forms a peak of positive activation there. Depending on the strength of self-excitation and pairwise inhibition, the field may either allow only for a single peak (in which case different field locations with sufficient input compete for selection) or for the co-existence of multiple peaks. Moreover, sufficient self-excitation may make fields self-sustained, so that peaks remain when the inducing input disappears – a model of STM.

Peaks are the units of representation of DFT, since they yield a non-zero output $\sigma(u(x, t))$ which can be passed on to other fields as input s . This allows to build architectures.

Model

Our model is depicted in Figure 1 d,e. In the following, we describe each of its components.

Short-term memory of conceptual structure

Objects A solution to Jackendoff’s challenges requires a unique binding agent for each object. To encode that a given object is characterized by a given set of concepts (e.g., *big*, *red*, and *house*), the binding agent of that object needs to be bound to those concepts. To encode that multiple different objects are characterized by a given concept (e.g., if there are multiple trees), the different binding agents of those objects need to be independently bound to the concept.

We propose that the required binding agent is realized as an object index that is assigned to each object upon processing its linguistic description. For the example in Figure 1a, the indices would be assigned as follows: “the tree(1) right of the tree(2) (which is) below the lake(3) and above the house(4)”.

The STM that specifies which concepts characterize an object with a given index is modeled by the *object/concept field*. It is defined over the discrete object index dimension and the discrete concept dimension. A peak at location (O, C) reflects that the object with index O is characterized by concept C .

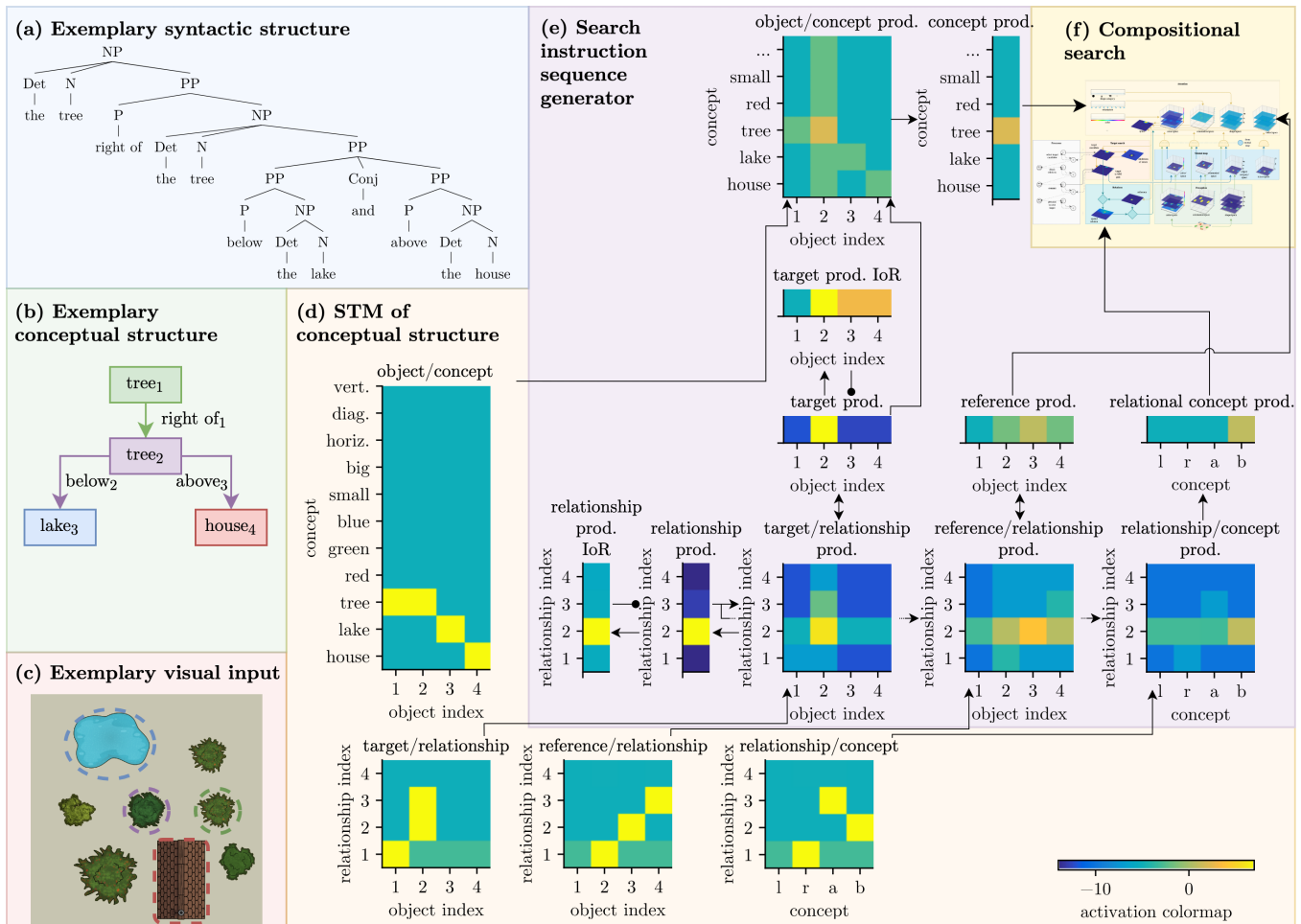


Figure 1: (a) Syntactic structure of an exemplary nested noun phrase. (b) Conceptual structure for that example. (c) Visual input in which the referent of the example phrase can be found. (d) The model for representing conceptual structure. (e) The model for generating a search instruction sequence. (f) Link to a model of compositional search (Sabinasz et al., 2020).

Relationships A solution to Jackendoff’s challenges also requires a binding agent that is unique to each relationship. To encode that a given object occurs in multiple different relationships (e.g., “the tree below the lake and above the house”), the binding agent of the object needs to be independently bound to the different binding agents of the relationships. This binding also has to specify in which relational role the object occurs. In the present paper, we limit ourselves to spatial relationships like *left of*, *right of*, *above*, and *below* with two relational roles – *target* and *reference*. For example, in the phrase “the tree below the lake”, the tree is the target and the lake is the reference.

The binding agent for relationships is realized in the form of a relationship index assigned to each relationship upon processing its linguistic description. For our example, the indices would be assigned as follows: “the tree right of (1) the tree (which is) below (2) the lake and above (3) the house”.

The STM that allows binding objects to relationships and roles is modeled by two fields, the *target/relationship field* and the *reference/relationship field*. They are both defined

over the object index dimension and the relationship index dimension. A peak at some location (O, R) carries the information that the object with index O is, respectively, the target or reference of the relation with index R . As an example, consider the activation snapshots in Figure 1d. They encode the relationships of Figure 1b.

Every relationship is additionally characterized by a relational concept. This is modeled by the *relationship/concept field*, which is defined over the discrete relational concept dimension and the relationship index. A peak in that field at some location (C, R) reflects that the relationship with index R is characterized by the relational concept C .

The described fields are filled by a language pre-processing system that performs the index assignments using two neural mechanisms of index nodes that get activated in sequence (Sandamirskaya & Schöner, 2010), with a novel mechanism that allows to also go backwards in the sequence, which is necessary to refer back to a previous object when encountering a new prepositional phrase. In the present paper, we do not describe the details of this mechanism.

Search instruction sequence generator

The search instruction sequence generator guides the compositional search system such as to find the configuration of objects described in the conceptual structure.

Target production The model contains a *target production field* defined over the object index dimension. A peak in that field at some location O reflects a decision to find the target object with index O next. The field is self-sustained, so that a peak remains until it is actively inhibited. Furthermore, the locations mutually inhibit each other, so that only one of them can be active at a time. The objects stored in the conceptual structure therefore compete for selection as the object that is to be searched next. This is necessary because the compositional search system can only find one object at a time.

Each object should only be searched once. For this purpose, the model contains a *target production inhibition-of-return (IoR) field*, which receives input from the target production field and is self-sustained. It inhibits the target production field, preventing a previously searched-for object to win the next competition.

When the target production field forms a peak for a certain object index, a search should be triggered which takes into account all of the concepts specified for that object index in the object/concept field. To achieve this, the model contains an *object/concept production field*. That field receives subthreshold input from the object/concept field. Additionally, it receives subthreshold input from the target production field along the shared object index dimension. The field forms peaks wherever these two inputs overlap. Thus, a peak at some location (O, C) carries the information that the selected target object O is characterized by the concept C .

Further, the model contains a *concept production field* defined over the discrete concept dimension. That field receives input from the object/concept production field, which is summed along the object index dimension. In effect, it forms a peak at a concept whenever the object/concept production field contains a peak at that concept, which can serve as input to the search system.

When the search system has successfully found the object, it can temporarily inhibit the target production field, thereby deleting the peak. After inhibition is released, the target indices compete for activation again, leading to a new decision about which target object to find next.

Relationship production Analogously to the target production field, the model contains a *relationship production field* defined over the relationship index dimension. A peak in that field signals that the respective relationship is to be processed by the compositional search system next. As before, the field is self-sustained, mutual inhibition leads to competition, an IoR field prevents a relationship from being selected more than once, and the search system can delete the peak after processing the relationship.

A relationship should only become active if it contains the target object that is currently searched for in its target

role. This is achieved by means of the *target/relationship production field*. It receives subthreshold input from the target/relationship field, as well as from the target production field along the shared object index dimension. It forms peaks where these two inputs overlap. Thus, if object O has been selected in the target production field, the target/relationship production field forms a peak at all locations (O, R) for all relationships R that contain O in their target role. The relationship production field receives input from that field, which is summed along the object index dimension. In effect, the competition is biased strongly towards relationships that contain the currently active target object in their target role.

Searching for a target that is characterized by a relationship to a reference object requires reading out the reference object index and the relational concept specified in the conceptual structure. The former is achieved by the *reference/relationship production field*, which receives subthreshold input from the reference/relationship field, as well as from the relationship production field along the shared relationship index dimension. When these inputs overlap, the field forms a peak. A peak at location (O, R) carries the information that the selected relationship R has object O as its reference object. Additionally, there is a *reference production field* defined over the object index dimension. That field receives input from the reference/relationship production field, which is summed along the object index dimension. Effectively, the reference production field forms a peak on object O when the reference/relationship production field contains a peak at O .

The reading-out of the relational concept is achieved in a completely analogous fashion by the *relationship/concept production field* and a *relational concept production field* defined over the discrete relational concept dimension.

Competitive advantages Recall that the order in which objects are searched can in principle be arbitrary, but certain more efficient orders are more likely to be employed. A parsimonious way to account for this is by assuming that the order emerges due to dynamic interactions that bias the competition in the target production field and relationship production field to favor the selection of some objects or relationships over others. In the present incarnation of the model, a negative bias is provided to each object index in the target production field, which is proportional to the number of relationships with unsaturated reference objects that the respective object occurs in as a target. This is achieved by the spreading of subthreshold activation. The details of this mechanism are beyond the scope of the present paper. In effect, targets whose reference objects have already been found are preferentially selected.

Compositional search

The production fields for *concept*, *reference*, and *relational concept* project to a model of compositional search (an improved version of Sabinasz et al., 2020), which enables, e.g., to search for “a tree below 3 and above 4”. A *target candidate* is identified based on a matching first relation (e.g., “a tree below 3”), and then it is checked whether the additional

relations also match (e.g., “above 4”). If all relations match, the target is stored in a *mental map field* which is defined over space and discrete index (Figure 2, bottom row). This enables referring back to the object’s location in future searches.

Results

To test the model’s behavior, we simulated it using the DFT software framework *cedar*. Figure 2 shows activation snapshots of relevant fields as it grounds the phrase from Figure 1a. This example involves two trees and therefore probes the model’s ability to solve the problem of 2. Further, the second tree is the reference object of the first relationship and the target object of the other two relationships, probing the model’s ability to flexibly bind object descriptions to multiple different relationships.

Selecting a house: At time t_1 , the target production field has high subthreshold activation at indices 3 and 4, since these objects do not occur as a target in any relationships with unsaturated reference objects. By t_2 , object index 4 has won the competition and has been selected for search. This has resulted in subthreshold slice input to column 4 of the object/concept production field, which has led to the formation of a peak at location (4, house). In turn, the concept production field has formed a peak for *house*. This has triggered the compositional search system to visually search for houses¹, which has resulted in a selection of the location of a house in the target candidate field. That location is committed to index 4 in the mental map field.

Selecting a lake: At t_3 , the target production field is biased away from object index 4 by the IoR field, resulting in a selection of object index 3 by time t_4 . By the same mechanisms as before, a lake is selected in the target candidate field and stored at index 3 in the mental map field.

Selecting a tree below the lake: At t_5 , the target production field has higher activation at index 2 than at the other remaining index 1, since it doesn’t occur as a target in any relationship with an unsaturated reference object. By t_6 , it has won the competition. By the same mechanisms as before, the tree concept becomes active in the concept production field.

Meanwhile, relationship 2 has won the competition in the relationship production field, since it is one of the relationships that contain the selected target object 2 in their target role (providing one source of bias), and since its reference object has been found shortly before (providing another source of bias). In effect, the reference production field has selected object 3 as the reference object due to the coupling through the reference/relationship production field, and the relational concept production field has selected *below* due to the coupling through the relationship/concept production field.

Thus, the compositional search system has been triggered to search for a tree that is below reference object 3 (the lake), whose position can be read from the mental map. By t_6 , such

an object has been selected in the target candidate field.

Checking if that tree is also above the house: At t_7 , relationship 3 is selected in the relationship production field. By the same mechanisms as before, this leads to selection of 4 in the reference production field and *above* in the relational concept production field, which in turn leads the compositional search system to check whether the target candidate is above object 4 from the mental map.

Since all the relationships matched, the target candidate is committed to index 2 in the mental map field at t_8 .

From t_9 to t_{10} , analogous mechanisms as before lead the model to find object 1 (the tree to the right of 2).

As noted before, during online language comprehension, objects are often attended in the order in which they are mentioned. Our model does not do this for this example because the STM of conceptual structure is already filled before the search is started, which enables the search from bottom to top. If the STM were filled in the order in which objects are mentioned, while the search is already going on, then our model would also attend objects in the order of mention.

Discussion

We have presented a neural dynamic process model that can perceptually ground a nested noun phrase in the visual array, to which it is dynamically coupled through the sensory surface. While not explicitly mapped onto areas on the brain, the model is consistent with the neural principles formalized in DFT that characterize neural populations by strong internal interaction and enable their coupling into neural dynamic architectures. The model contains a STM of conceptual structure that can represent the structure of relational dependency between objects. The STM can be filled by the language system while simultaneously providing input to a neural process that generates a sequence of searches that together successfully and efficiently find the described object. Thus, the model solves the challenges identified in the introduction.

The search order emerges from local interactions that bias competitive selection (e.g., in favor of objects whose reference objects have already been found). Effectively, the conceptual structure tree can thus affect the order (e.g., by leading to an emergent processing from the leaves to the root) without requiring algorithmic tree traversal methods.

A number of algorithmic models that resolve noun phrases or conceptual structures in perceptual representations have been proposed. Some make use of pointers and recursive function calls to implement the tree traversal of conceptual structures (e.g., Brown, Buntschuh, & Wilpon, 1992; Nagao & Rekimoto, 1995; Gorniak & Roy, 2004). It is not obvious how such methods would be realized by neural processes.

Van der Velde and De Kamps (2006) propose a neural model addressing Jackendoff’s challenges at the level of syntactic phrase structure. That model uses neural assemblies to represent noun phrases (implicitly representing object descriptions), and prepositional phrases (implicitly representing relationship descriptions), together with a mechanism that

¹This is enabled using a field defined over two-dimensional space and a discrete object category dimension, which is fed by the feature maps of a 2-layer CNN with hand-crafted weights (Sabinasz, 2019).

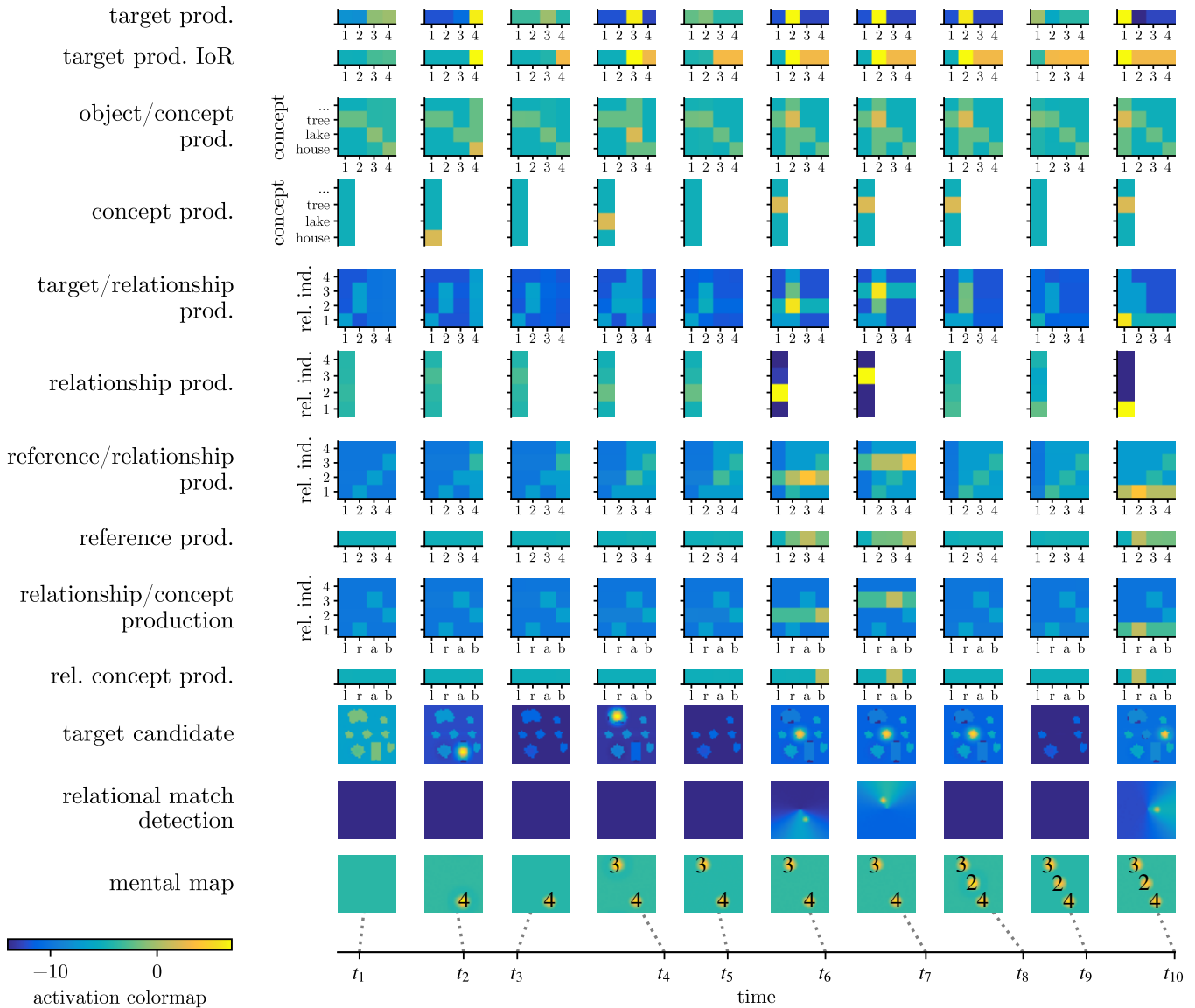


Figure 2: Activation snapshots of the architecture while it generates a search instruction sequence for the example phrase from Figure 1a. Field activations are shown as color-coded snapshots at discrete moments in time (t_1, \dots, t_{10}).

flexibly binds assemblies. This emphasizes syntactic rather than conceptual structure. The neural mechanisms employed differ from ours. For instance, in the present model, interactions within populations stabilize STM against decay, which we find to be important to solve the task. This leads to a limit on the number of objects and relationships that can be held in the STM of conceptual structure (Johnson, Simmering, & Buss, 2014), much smaller than the number of objects that can, in principle, be mentioned in a given sentence or noun phrase. For example, even though humans are able to find the referent of a chain like “the lion next to the zebra on the mat next to the house at the lake”, we contend that not all object and relationships are held in the STM of conceptual structure at the same time. When the number of mentioned objects (or relationships) exceeds a limit, newly mentioned objects

(or relationships) replace old ones in the STM of conceptual structure. This prevents the kind of combinatorial explosions discussed by Stewart and Eliasmith (2012).

Vector-symbolic architectures (VSAs) offer an alternative neurally inspired framework which addresses Jackendoff’s challenges for STM representations of conceptual structure (Smolensky, 1990; Gayler, 2003). An implementation in spiking neural networks (Stewart & Eliasmith, 2012) has been hypothesized to enable coupling to perceptual and motor processes (Eliasmith, 2013). The extent to which these accounts are compatible with the neural principles postulated in DFT needs further study, however.

Our ongoing research explores how the architecture may scale to large vocabularies and more complex grammatical structures.

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.
- 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.
- 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.
- 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.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic Books.
- 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).
- 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.
- Sandamirskaya, Y., & Schöner, G. (2010). An embodied account of serial order: How instabilities drive sequence generation. *Neural Networks*, 23(10), 1164–1179.
- 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.