

Daniel Sabinasz

A Neural Dynamic Model for the Representation and Perceptual Grounding of Conceptual Structure

It is widely recognized that a crucial feature of human language is its combinatorial structure, and it is hypothesized by many that language understanding requires that the brain represents structure. Understanding how neural networks may represent, process, and manipulate structures is thus a crucial step towards accounting for the neural basis of language understanding and other higher cognitive competences. To make progress in this direction, we introduce a neural dynamic process model that perceptually grounds nested phrases, i.e., language expressions that describe the interrelationships among objects in a given visual scene, at potentially several layers of recursion. Examples for such expressions are “The tree to the left of the house is bigger than the tree to the right of the lake,” “The tree which is to the right of the house at the lake is smaller than the other tree,” “The ball approaches the tree, which is to the left of the lake and to the right of the house.” The model is based on Dynamic Field Theory (DFT; Schöner, Spencer, & the DFT Research Group, 2015), a framework for building dynamical systems models of the neural basis of various forms of cognition. It autonomously generates neural activation patterns that effectively search the objects represented in that structure, and tests all of the relationships represented in that structure, such as to decide whether the sentence represented in that structure is true. The model provides accounts for how Jackendoff’s (2002) challenges for representing structure may be solved in accord with DFT, for how concepts represented in the structure may be linked to their perceptual meanings, for how this may subserve visual search for the objects and the testing of the relationships with empirically corroborated neural mechanisms, for how the outcomes of a sequence of visual searches may be sustained in short-term memory such that they may be incorporated in other visual searches or the testing of relationships, and for how the various components of the model may be orchestrated such as to generate sequences of discrete processing steps that effectively enable the model to autonomously test the truth or falsehood of the entire nested phrase. We demonstrate simulation results that illustrate the model’s capacity to correctly determine the truth or falsehood of a large variety of nested phrases with a single set of parameters. We provide arguments and empirical evidence for various aspects of the model. We embed it in the literature on cognitive linguistics, the cognitive psychology of concepts, and the ongoing debate about systematicity and compositionality, which are widely believed to be explananda that models of higher cognition should work towards addressing. We compare the model to other approaches to the representation and perceptual grounding of structure, and conclude by highlighting limitations and next steps.

Es wird weithin angenommen dass eine wesentliche Eigenschaft der menschlichen Sprache ihre kombinatorische Struktur ist, und dass das menschliche Sprachverständnis erfordert, dass das Gehirn Struktur repräsentiert. Ein Verständnis, wie neuronale Netze Strukturen repräsentieren, verarbeiten und manipulieren können ist demnach ein wesentlicher Schritt, um die neuronale Basis des Sprachverständnisses zu verstehen. Um Fortschritt in diese Richtung zu machen, führen wir ein neuro-dynamisches Prozessmodell ein, welches rekursiv verschachtelte Phrasen perzeptuell grundieren kann, d.h., Phrasen die die Relationen zwischen Objekten in einer gegebenen visuellen Szene beschreiben, auf verschiedenen rekursiven Ebenen. Beispiele für solche Phrasen sind „Der Baum links vom Haus ist größer als der Baum rechts vom See“, „Der Baum rechts vom Haus am See ist kleiner als der andere Baum“, „Der Ball bewegt sich auf denjenigen Baum zu, welcher links vom See und rechts vom Haus ist.“ Das Modell basiert auf den Prinzipien der Dynamischen Feldtheorie (DFT; Schöner, Spencer, & DFT Research Group, 2015), einem Framework für das Modellieren der neuronalen Basis der Kognition mittels dynamischer Systeme. Es generiert autonom neuronale Aktivierungsmuster, die die Objekte suchen, welche in der Struktur repräsentiert sind, und testet alle Relationen, die in der Struktur repräsentiert sind, um letzten Endes zu entscheiden, ob die Phrase wahr ist. Das Modell erfasst, wie Jackendoffs (2002) Challenges für das Repräsentieren von Struktur mit DFT gelöst werden können, wie Konzepte die in der Struktur repräsentiert sind mit ihren perzeptuellen Bedeutungen assoziiert werden können, wie dies visuelle Suche nach Objekten und das Testen von Relationen mit empirisch korroborierten neuronalen Mechanismen ermöglicht, wie die Ergebnisse einer Reihe von visuellen Suchen im Arbeitsgedächtnis gehalten werden können, damit sie in zukünftigen visuellen Suchen oder dem Testen von Relationen berücksichtigt werden können, und wie die diversen Komponenten des Modells orchestriert werden können um Sequenzen von diskreten Arbeitsschritten zu generieren die letztlich dem Modell ermöglichen, autonom die Wahrheit oder Falschheit der Phrase zu überprüfen. Wir demonstrieren Simulationsergebnisse, die zeigen, dass das Modell korrekt die Wahrheit oder Falschheit einer großen Bandbreite von verschachtelten Phrasen mit einem einzelnen Parametersatz überprüfen kann. Wir liefern Argumente und empirische Evidenz für diverse Aspekte des Modells. Wir betten es in die Literatur über kognitive Linguistik, die kognitive Psychologie der Konzepte, und die fortwährende Debatte über Systematizität und Kompositionalität ein. Wir vergleichen das Modell mit anderen Ansätzen zur Repräsentation und perzeptuellen Grundierung von Strukturen, und konkludieren indem wir Grenzen des Modells und nächste Schritte beschreiben.