

# The role of language in broader human cognition: evidence from neuroscience

by

Anna Alexandrovna Ivanova

B.S. in Neuroscience & Computer Science  
University of Miami, 2017

Submitted to the Department of Brain and Cognitive Sciences  
in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy in Brain and Cognitive Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2022

© 2022 Massachusetts Institute of Technology. All rights reserved

Signature of Author .....

Department of Brain and Cognitive Sciences  
August 23, 2022

Certified by .....

Evelina Fedorenko  
Frederick A. (1971) and Carole J. Middleton Career  
Development Associate Professor of Neuroscience  
Thesis Supervisor

Accepted by.....

Mark Harnett  
Graduate Officer, Department of Brain and Cognitive Sciences

# The role of language in broader human cognition: evidence from neuroscience

by

Anna Alexandrovna Ivanova

Submitted to the Department of Brain and Cognitive Sciences  
on August 23<sup>rd</sup>, 2022 in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in Brain and Cognitive Sciences

## Abstract

Many philosophers, psychologists, biologists, computer scientists, and linguists have argued that language processing serves as a foundation for human cognition. However, evidence from neuroscience has shown that language might rely on specialized cognitive mechanisms that are distinct from many aspects of human thought. In this thesis, I use cognitive neuroscience to test the limits of the brain's functional specialization for language processing. In **Chapter 1**, I describe how evidence from neuroscience can illuminate the relationship between language and other cognitive functions. In **Chapter 2**, I investigate activity in the brain's language network in response to computer code, an input that shares many structural similarities with natural language. I find that, despite these similarities, the language network responds weakly or not at all during computer code comprehension; instead, this process elicits responses in brain areas of a distinct, domain-general multiple demand network. In **Chapter 3** and **Chapter 4**, I study the language network's responses to pictures of objects and events during semantic tasks, which, like language comprehension, require access to conceptual information. I show that the language network does not respond during an object semantics task and that its responses to event semantics are not causally important for performing the task. In **Chapter 5**, I describe a set of brain regions that respond to semantic demand regardless of stimulus type (sentences vs. pictures) and show that they are distinct from both the language network and the domain-general multiple demand network. Finally, in **Chapter 6**, I discuss the implications of my work for a neuroscience-informed account of the mechanisms underlying human cognition and language use. My work establishes that language processing mechanisms are largely distinct from mechanisms that support the processing of non-linguistic structure and meaning, even for closely matched inputs, and helps further delineate the functional architecture of the human mind.

Thesis Supervisor: Evelina Fedorenko

Title: Frederick A. (1971) and Carole J. Middleton Career Development Associate Professor of Neuroscience

## Acknowledgements

I would like to start by thanking the people who have helped me on my path toward graduate school. First of all, it is my parents, who found amazing educational opportunities for me as a kid, supported my decision to study in the U.S., and have always made me feel loved despite being an ocean away. Second, it is the teachers at my school back home, most notably Sergey Mendelevich Glagolev, Polina Volkova, and Andrei Nikolayevich Kvashenko, who showed me the beauty of natural sciences and the power of the scientific method. Third, it is my undergraduate research mentor, Lucina Uddin, who introduced me to cognitive neuroscience and functional MRI, creating an environment for me to learn many of the skills I used throughout the PhD. A separate shoutout goes to Paola Odriozola, who taught me the hidden curriculum of grad school applications, having just gone through the process herself.

Here at MIT, my thanks go first and foremost to my advisor Ev Fedorenko. Ev gives me the intellectual freedom to pursue the questions I find fascinating, while also providing a rigorous conceptual and methodological scaffold that allows me to grow as a scientist. She cares deeply about her mentees, both as individuals and as independent thinkers, and fosters a warm welcoming community in her lab, for which I am deeply grateful. Many thanks also to Nancy Kanwisher, who pushes me to think about the big picture, encourages me to use rigorous, interpretable methods, and is always there to share wisdom and provide guidance. Both Ev and Nancy are amazing role models and advocate tirelessly for their mentees. I am incredibly lucky to work with them.

I also thank my committee members, Roger Levy, Josh Tenenbaum, and Marina Bedny. Roger created a space for me to explore computational modeling during my first year of grad school, which has allowed me to relate my work to other areas of cognitive science. Josh has inspired me to think about hard questions ever since my interview day, when he asked me how we know that a chair is a chair. Marina has shown me the value of collaborative science when, after learning that her group and ours worked on similar projects, she offered us to submit our papers to the same journal instead of racing to the finish line.

Many other members of the MIT community contributed greatly to my growth as a scientist. It has been a pleasure to learn from Rebecca Saxe, including her amazing discussion-based class that I TAed and her advocacy efforts in the BCS department and beyond. Ted Gibson's approach to the study of language has substantially influenced my thinking, and his warmth and friendliness as a party host made TEvlab feel like home. Thanks to Ted and Ev, I also got to interact with Ray Jackendoff, who continually reminds me and others about the grand mysteries in the study of meaning. Jacob Andreas teaches an outstanding class on language models and strives to bring together cognitive science and AI, an effort I am excited to contribute to going forward. Noga Zaslavsky encouraged me to start a "generative adversarial collaboration" (GAC) project aimed at examining methods in cognitive neuroscience, which resulted in a workshop with 400+ attendees and a subsequent position paper. Leyla Isik taught me the fundamentals of MEG decoding and, a few years later, bravely agreed to senior author the GAC paper.

My PhD experience would not have been the same without the TEvlab community. I don't have space to describe all the meaningful interactions I've had with this amazing group of people, but I appreciate every one of them. In particular, I thank Idan Blank, who taught me to make effective presentation slides and, by openly sharing his grad school experiences, encouraged me to do the same. Thanks also to Rachel Ryskin, who taught me the fundamentals of mixed effects models and plotting in R – skills that were essential for every empirical chapter of this thesis. Thanks to Hope Kean for her warmth and support, as well as her ability to seamlessly switch between science, philosophy, and girl talk. Thanks to Alvincé Pongos for strengthening my self-confidence as a scientist and for helping me survive the pandemic. Thanks to other current and former lab members who have supported me and made me feel welcome: Greta Tuckute, Saima Malik Moraleda, Carina Kauf, Eghbal Hosseini, Sihan Chen, Thomas Clark, Eric Martinez, Ben Lipkin, Colton Casto, Aalok Sathe, Niharika Jhingan, Yev Diachek, Matt Siegelman, Hannah Small, Josef Affourtit, Cory Shain, Tamar Regev, Sammy Floyd, Yingtong Liu, Giuseppe Ricciardi, Zach Mineroff, Elizabeth Lee, Maya Taliaferro, Leo Zekelman, and Nafisa Syed. Thanks to Levy and Kanwisher labs, the people in which helped shape the way I think about language and the brain, respectively. Thanks to Jon Gauthier, Matthias Hofer, and Tyler Brooke-Wilson for starting a philosophy circle which led to some interesting discussions over the years. Thanks also to my amazing mentees: Yotaro Sueoka, Chandler Cheung, Zawad Chowdhury, Selena She, and David Oluigbo.

Another distinguishing feature of my PhD is the many collaborators I have worked with, both at MIT and externally. Many thanks to Shashank Srikant, Una-May O'Reilly, Marina Bers, Riva Dhamala, Rosemary Varley, Vitor Zimmerer, Yael Benn, Oliver Clark, Kyle Mahowald, Tanya Goldhaber, Zuzanna Balewski, Chloe Seikus, Jack Santos Silva, Alessandro Lenci, Giulia Rambelli, Emmanuele Chersoni, Martin Schrimpf, Stefano Anzellotti, and John Hewitt.

Two non-academic experiences enriched my PhD life. One is the Grad Blog, where I served as an editor for two years. Thanks to Diana Chien and Heather Konar, as well as all my fellow editors, for creating and supporting this initiative. The other is student advocacy and support groups at our department, Gadvocates and REFS. Thanks to Joyce Wang, Lupe Cruz, and Danielle Cosio for creating these groups, as well as all the people (students, faculty, and admins) who have contributed to positive change in the department.

I thank my cohort, with whom I got to share my PhD journey (even if most of it was spent in different labs). Thanks to Alex Ferguson, Peng Qian, Victoria Beja-Glasser, Jungsoo Kim, Heather Kosakowski, Mahdi Ramadan, and others.

Finally, thanks to my non-MIT friends who help me remember about life beyond the PhD: Sasha, Polya, Kolya, Crystal, Julia, Sara, Oscar, Masha, Drew, and Aidar. Thanks to my siblings, Nadia, Tyoma, and Max, and to my extended family for always reminding me that I belong. Spencer, thank you for your unconditional love.

I dedicate this thesis to my home country, with the wish that one day it will be known first and foremost for its science and culture.

# Table of Contents

<b>CHAPTER 1. INTRODUCTION</b>	<b>11</b>
<b>1.1 LANGUAGE AS A FOUNDATION FOR HUMAN COGNITION?</b>	<b>11</b>
<b>1.2 POSSIBLE HYPOTHESES</b>	<b>12</b>
<b>1.3 BACKGROUND: SPECIFICITY OF LANGUAGE PROCESSING IN THE BRAIN</b>	<b>14</b>
<b>1.4 OVERVIEW OF THIS WORK</b>	<b>17</b>
<b>1.5 APPROACH</b>	<b>17</b>
1.5.1 LOCALIZING THE LANGUAGE NETWORK IN INDIVIDUAL BRAINS	17
1.5.1.1 The benefits of functional localization	17
1.5.1.2 Deriving and studying functional regions of interest	19
1.5.1.3 Pre-selecting anatomical regions of interest (parcels)	20
1.5.2 COMBINING NEUROIMAGING EVIDENCE FROM NEUROTYPICALS AND BEHAVIORAL EVIDENCE FROM APHASIA	21
<b>1.6 USING DATA FROM NEUROSCIENCE TO INFORM COGNITIVE THEORIES</b>	<b>23</b>
<b>CHAPTER 2. THE LANGUAGE NETWORK &amp; COMPUTER CODE COMPREHENSION</b>	<b>25</b>
<b>2.1 ABSTRACT</b>	<b>25</b>
<b>2.2 INTRODUCTION</b>	<b>26</b>
<b>2.3 RESULTS</b>	<b>30</b>
2.3.1 MD SYSTEM EXHIBITS ROBUST AND GENERALIZABLE BILATERAL RESPONSES DURING CODE COMPREHENSION	32
2.3.2 NO MD FROIS ARE SELECTIVE FOR CODE COMPREHENSION	35
2.3.3 LANGUAGE SYSTEM RESPONSES DURING CODE COMPREHENSION ARE WEAK AND INCONSISTENT	36
2.3.4 NO CONSISTENT EVIDENCE OF CODE-RESPONSIVE REGIONS OUTSIDE THE MD/LANGUAGE SYSTEMS	39
2.3.5 EFFECT OF PROFICIENCY ON MD AND LANGUAGE RESPONSES	40
<b>2.4 DISCUSSION</b>	<b>40</b>
2.4.1 MD SYSTEM'S ENGAGEMENT REFLECTS THE USE OF DOMAIN-GENERAL RESOURCES	42
2.4.2 THE LANGUAGE SYSTEM IS FUNCTIONALLY CONSERVATIVE	44
2.4.3 LIMITATIONS OF SCOPE	47

<b>2.5 METHOD</b>	<b>48</b>
2.5.1 PARTICIPANTS	48
2.5.2 DESIGN, MATERIALS, AND PROCEDURE	49
2.5.2.1 Program comprehension tasks	49
2.5.2.2 Multiple demand localizer	51
2.5.2.3 Language localizer	51
2.5.3 FMRI DATA ACQUISITION	52
2.5.4 FMRI DATA PREPROCESSING	53
2.5.5 FIRST-LEVEL ANALYSIS	53
2.5.6 DEFINING MD AND LANGUAGE FUNCTIONAL REGIONS OF INTEREST (FROIs)	54
2.5.7 EXAMINING THE FUNCTIONAL RESPONSE PROFILES OF MD AND LANGUAGE FROIs	55
2.5.8 WHOLE-BRAIN ANALYSES	57
<b>2.6 ACKNOWLEDGEMENTS</b>	<b>57</b>
<b>2.7 APPENDIX</b>	<b>58</b>
2.7.1 SUPPLEMENTAL FIGURES	58
2.7.2 SUPPLEMENTAL TABLES	67
<b><u>CHAPTER 3. THE LANGUAGE NETWORK &amp; OBJECT SEMANTICS</u></b>	<b><u>70</u></b>
<b>3.1 ABSTRACT</b>	<b>70</b>
<b>3.2 INTRODUCTION</b>	<b>71</b>
<b>3.3 METHOD</b>	<b>75</b>
3.3.1 PARTICIPANTS	75
3.3.2 DESIGN, MATERIALS, AND PROCEDURE	76
3.3.2.1 Categorization	76
3.3.2.2 Language localizer	77
3.3.2.3 Multiple demand localizer	77
3.3.3 FMRI DATA ACQUISITION	77
3.3.4 FMRI DATA PREPROCESSING	77
3.3.5 FIRST-LEVEL ANALYSIS	77
3.3.6 DEFINING INDIVIDUAL FUNCTIONAL REGIONS OF INTEREST (FROIs)	78

3.3.7 EXAMINING THE FUNCTIONAL RESPONSE PROFILES OF fROIs	78
3.3.8 STATISTICAL ANALYSES	78
<b>3.4 RESULTS</b>	<b>79</b>
3.4.1 THE LANGUAGE NETWORK SHOWS LOW RESPONSES TO BOTH SEMANTIC AND PERCEPTUAL CATEGORIZATION	79
3.4.2 THE MULTIPLE DEMAND NETWORK RESPONDS COMPARABLY TO SEMANTIC AND PERCEPTUAL CATEGORIES	80
3.4.3 WHOLE-BRAIN ANALYSES REVEAL NO REGIONS RESPONSIVE SPECIFICALLY TO SEMANTIC CATEGORIES	81
<b>3.5 DISCUSSION</b>	<b>82</b>
<b>3.6 ACKNOWLEDGEMENTS</b>	<b>83</b>
<b>3.7 APPENDIX</b>	<b>85</b>
3.7.1 MIXED-EFFECTS MODELS CONTRAST SPECIFICATION	85
3.7.2 SUPPLEMENTAL TABLES	86
<b>CHAPTER 4. THE LANGUAGE NETWORK &amp; EVENT SEMANTICS</b>	<b>92</b>
<b>4.1 ABSTRACT</b>	<b>92</b>
<b>4.2 INTRODUCTION</b>	<b>93</b>
<b>4.3 METHOD</b>	<b>97</b>
4.3.1 EXPERIMENT 1: IS THE LANGUAGE NETWORK ACTIVE DURING A NONVERBAL EVENT SEMANTICS TASK?	97
4.3.1.1 Overview	97
4.3.1.2 Participants	97
4.3.1.3 Design, materials, and procedure	98
4.3.1.4 fMRI data acquisition	100
4.3.1.5 fMRI data preprocessing	100
4.3.1.6 First-level analysis	100
4.3.1.7 Defining functional regions of interest (fROIs)	100
4.3.1.8 Examining the functional response profiles of the language fROIs	100
4.3.1.9 Statistical analyses	100
4.3.2 EXPERIMENT 2: IS THE LANGUAGE NETWORK REQUIRED FOR A NONVERBAL EVENT SEMANTICS TASK?	101
4.3.2.1 Overview	101
4.3.2.2 Participants	101
4.3.2.3 Semantic Task: Picture plausibility judgments	102

4.3.2.4	Language Task: Sentence-picture matching	103
4.3.2.5	Data analysis	104
4.3.2.6	Estimating the damage to the language network in patients with aphasia	104
<b>4.4</b>	<b>RESULTS</b>	<b>105</b>
4.4.1	EXPERIMENT 1: IS THE LANGUAGE NETWORK ACTIVE DURING A NONVERBAL EVENT SEMANTICS TASK?	105
4.4.1.1	Behavioral Results	105
4.4.1.2	Neuroimaging Results	106
4.4.2	EXPERIMENT 2: IS THE LANGUAGE NETWORK REQUIRED FOR A NONVERBAL EVENT SEMANTICS TASK?	109
<b>4.5</b>	<b>DISCUSSION</b>	<b>113</b>
4.5.1	THE LANGUAGE NETWORK IS NOT REQUIRED FOR NONVERBAL EVENT SEMANTICS	114
4.5.2	IMPLICATIONS FOR THEORIES OF SEMANTICS IN THE BRAIN	117
4.5.3	IMPLICATIONS FOR NEUROIMAGING STUDIES OF AMODAL SEMANTICS	120
<b>4.6</b>	<b>ACKNOWLEDGMENTS</b>	<b>120</b>
<b>4.7</b>	<b>APPENDIX</b>	<b>122</b>
<b>CHAPTER 5. AMODAL SEMANTICS OUTSIDE THE LANGUAGE NETWORK</b>		<b>124</b>
<b>5.1</b>	<b>ABSTRACT</b>	<b>124</b>
<b>5.2</b>	<b>INTRODUCTION</b>	<b>125</b>
<b>5.3</b>	<b>METHOD</b>	<b>128</b>
5.3.1	PARTICIPANTS	128
5.3.2	CRITICAL EXPERIMENTS	129
5.3.2.1	Experiment 1	130
5.3.2.2	Experiment 2	131
5.3.2.3	Experiment 3	133
5.3.3	LOCALIZER EXPERIMENTS	134
5.3.3.1	Language localizer	135
5.3.3.2	Multiple demand and default mode localizer	135
5.3.4	FMRI DATA ACQUISITION	136
5.3.5	FMRI DATA PREPROCESSING	136
5.3.6	FIRST-LEVEL ANALYSIS	136

5.3.7	DEFINING FUNCTIONAL REGIONS OF INTEREST (FROIs)	136
5.3.8	EXAMINING THE FUNCTIONAL RESPONSE PROFILES OF FROIs	137
5.3.9	STATISTICAL ANALYSES	138
5.3.10	OVERLAP ANALYSIS	138
<b>5.4</b>	<b>RESULTS</b>	<b>139</b>
5.4.1	MULTIPLE BRAIN REGIONS RESPOND TO SEMANTIC TASK ACROSS SENTENCES AND PICTURES	139
5.4.2	THE RESPONSES ARE CONSISTENT ACROSS EXPERIMENTS	141
5.4.3	THE RESPONSE TO SEMANTIC TASKS IN OTHER COGNITIVE NETWORKS	142
5.4.4	SEMANTIC FROIs ARE LARGELY DISTINCT FROM KNOWN FUNCTIONAL NETWORKS	145
<b>5.5</b>	<b>DISCUSSION</b>	<b>146</b>
5.5.1	INFERIOR FRONTAL GYRUS (IFG)	146
5.5.2	POSTERIOR TEMPORAL CORTEX	147
5.5.3	ANGULAR GYRUS (AG) AND INTRAPARIETAL SULCUS (IPS)	148
5.5.4	CEREBELLUM	149
5.5.5	ANTERIOR TEMPORAL LOBE (ATL)	149
5.5.6	POSTERIOR MEDIAL PARIETAL CORTEX	150
<b>5.6</b>	<b>ACKNOWLEDGMENTS</b>	<b>151</b>
<b>5.7</b>	<b>APPENDIX</b>	<b>152</b>
5.7.1	BEHAVIORAL RESULTS	152
5.7.2	THE RESPONSES GENERALIZE TO LEFT-OUT PARTICIPANTS	153
5.7.3	DIFFERENT EVENT SEMANTICS EXPERIMENTS RECRUIT CO-LOCATED BUT DISTINCT CORTICAL REGIONS	154
5.7.4	SUPPLEMENTAL TABLES	156
<b>CHAPTER 6. CONCLUSION</b>		<b>164</b>
<b>6.1</b>	<b>THESIS SUMMARY</b>	<b>164</b>
<b>6.2</b>	<b>LANGUAGE PROCESSING OUTSIDE OF THE LANGUAGE NETWORK</b>	<b>165</b>
6.2.1	MULTIPLE DEMAND NETWORK: HARD LANGUAGE AND EXTERNAL TASKS	165
6.2.2	THEORY OF MIND NETWORK: PRAGMATICS AND SOCIAL INFERENCE	166
6.2.3	THE DEFAULT MODE NETWORK: NARRATIVES AND SITUATION MODELING	167
6.2.4	SEMANTIC DEMAND FROIs AND OTHER BRAIN REGIONS	169

<b>6.3</b>	<b>"LANGUAGE AS A FOUNDATION FOR HUMAN COGNITION" REVISITED</b>	<b>169</b>
<b>6.4</b>	<b>CAVEAT: DEVELOPMENT IS DIFFERENT</b>	<b>171</b>
<b>6.5</b>	<b>IMPLICATIONS FOR INNER THOUGHT</b>	<b>172</b>
<hr/> <b>REFERENCES</b>		<b>176</b>

# Chapter 1

## Introduction

### 1.1 Language as a foundation for human cognition?

Language is a unique cognitive skill that sets humans apart from other animals. Is it a coincidence that a species whose members can flexibly share knowledge amongst each other has achieved such remarkable evolutionary success? Surely not: the ability to share information between individuals is vital to the gradual accumulation of knowledge from generation of generation, enabling each individual to learn not only from their experience, but also from the experience of many others. However, there might also be a more direct link between language and human advancement as a species: language might serve as a foundation for other cognitive abilities that underlie human intelligence.

Many thinkers have posited a strong connection between language and other aspects of human cognition, in fields as diverse as philosophy (Carruthers, 2002; Davidson, 1975; Gauker, 2011; Wittgenstein, 1961), psychology (Sokolov, 1972; Vygotsky, 1934; Watson, 1920), linguistics (Asoulin, 2016; Berwick & Chomsky, 2016; Bickerton, 1990; Chomsky, 2007; Hinzen, 2013; Jackendoff, 1996; Oller Jr., 1981), and artificial intelligence (T. B. Brown et al., 2020; Goldstein & Papert, 1977; Turing, 1950; Winograd, 1976). Edward Sapir was a well-known proponent of this view, stating: “Language and our thought-grooves are inextricably interwoven, are, in a sense, one and the same” (Sapir, 1921). Decades later, Noam Chomsky also advanced the view that language is essential for thought: “The systems of thought … use linguistic expressions for reasoning, interpretation, organizing action, and other mental acts” (Chomsky, 2007). Even Charles Darwin posited a link between language and thought: “The continued use and advancement of [language] would have reacted on the mind by enabling and encouraging it to carry on long trains of thought” (Darwin, 1871).

What does it mean for language to serve as a foundation for non-linguistic cognitive abilities? There is no one answer to this question. In fact, the researchers mentioned above might have very different notions of the language-thought relationship. In this thesis, I evaluate a specific hypothesis about the link between language and thought: that the neural mechanisms used for language processing are also used for other cognitive functions.

In the rest of this chapter, I will first outline different hypotheses that describe this putative relationship more precisely. Then, I will provide motivation for the studies included in this thesis, describe the conceptual and methodological framework underlying my work, and show how it can be used to test the view that language is a foundation for human thought.

## 1.2 Possible hypotheses

My thesis aims to address the following question: how does language processing relate to other cognitive functions in an adult human mind? In this section, I provide a list of key hypotheses addressed in my work.

**H1. Language processing is separate from other cognitive functions.** This hypothesis posits that language processing has a dedicated cognitive module in the brain, distinct from both lower-level processes (such as speech perception) and higher-level processes (such as reasoning). If true, language should be considered one of many cognitive capacities that contribute to human intelligence, but not a “foundation” for them.

**H2. Language shares processing mechanisms with other domains that use symbolic composition.** This hypothesis postulates that the mechanisms that underlie language processing are also used to carry out symbolic composition in other domains, such as music, math, or logic. This view relies on the fact that language use requires composing discrete symbolic units in flexible ways using a system of hierarchical recursive rules.

This ability is sometimes posited as the defining characteristic of human intelligence (e.g., Berwick et al., 2013; Berwick & Chomsky, 2016; Hauser et al., 2002).

**H3. Language shares processing mechanisms with conceptual reasoning.** According to this view, neural circuits that underlie language processing are also used to store, retrieve, and flexibly combine concepts. If so, linguistic processing is recruited during all conceptual reasoning processes, including inner thought. The exact form in which concepts are stored is under debate: a strong view posits that lexical and conceptual units are the same (e.g., Deacon, 1998; Oller Jr., 1981), whereas a weaker view might state that linguistic information constitutes just one part of a distributed conceptual representation (e.g., Barsalou, 2008; Paivio, 1991). Regardless of these differences, the key distinguishing feature of this view is that conceptual processing takes place within the language circuits.

This view that conceptual reasoning relies on linguistic processing mechanisms is often conflated with the postulation of the so-called “language of thought” (Fodor, 1975). The language of thought hypothesis states that inner thought and conceptual reasoning rely on a set of discrete internal symbols (“concepts”), which, just like words in a sentence, combine to yield an infinite number of combinations. In principle, language of thought might be able to operate completely independently from natural language, and thus the language of thought hypothesis is not the subject of this work.

The three hypotheses outlined above are focused on the mechanistic aspect of the language-thought relationship (are the language processing mechanisms also recruited for something else?) and thus leave out many other ways in which language might relate to thought. For instance, the contemporary discussion of the Sapir-Whorf hypothesis (also known as linguistic relativity) centers on the claim that the way a given language partitions the conceptual space into words can affect the way the language users think, even in situations when they are not using the language. For instance, people might be faster to distinguish colors that have different labels than colors with the same label (e.g., Winawer et al., 2007; Regier & Kay, 2009; cf. S. Chen et al., in prep; Martinovic et al., 2020), and users of languages with allocentric spatial reference frames (“the pen is north of the notebook”) might be much more attuned to the cardinal directions than users of language

with egocentric reference frames (“the pen is to the left of the notebook”; e.g., P. Brown & Levinson, 1992, 1993; cf. Li et al., 2011; Li & Gleitman, 2002). However, the Sapir-Whorf accounts typically do not depend on specific, mechanistic hypotheses about the role of language processing circuits. Language might affect conceptual reasoning not because these two functions rely on shared processing mechanisms, but because the language module interacts with other cognitive modules or because linguistic input influenced non-linguistic systems during learning. In all these cases, linguistic relativity accounts would still hold. Thus, the discussion of linguistic relativity is also largely outside the scope of this thesis.

Finally, as mentioned at the beginning of the chapter, language unquestionably affects thought in a more trivial way: it is an important source of information about the world. Language provides us with large amounts of conceptual knowledge from other people, without the need to acquire it from direct observation. This knowledge then serves as a substrate for reasoning and problem solving. In that sense, language is an essential component of human cognition. The question I address is much more specific: do the mechanisms that underlie language processing also contribute to other abilities that make up human intelligence?

### 1.3 Background: specificity of language processing in the brain

Two lines of work provide compelling evidence that language is neurally distinct from other cognitive functions (for a review, see Fedorenko & Varley, 2016). One is *evidence from brain imaging*. Studies investigating brain responses to language found a network of left-lateralized regions in frontal and temporal lobes, which is consistently recruited for language processing. This network responds to written, spoken, and signed linguistic input (e.g., Deniz et al., 2019; Fedorenko et al., 2010; MacSweeney et al., 2002; M. Regev et al., 2013; Scott et al., 2017), is engaged in both comprehension and production (J. Hu et al., 2021; Menenti et al., 2011), and responds during both task-driven paradigms and passive reading/listening (Cheung et al., 2020; Diachek et al., 2020; Fedorenko et al., 2010). So far there appears to be little functional specialization within different language-responsive regions, all of which are sensitive to phonological, syntactic, and semantic

properties of the input (Bautista & Wilson, 2016; Blank, Balewski, et al., 2016; Blank & Fedorenko, 2020; Fedorenko, Behr, et al., 2011; Fedorenko, Duncan, et al., 2012; Fedorenko et al., 2020a; T. I. Regev et al., 2021). Furthermore, these regions show correlated patterns of brain activity, both during language use and during passive rest conditions (Blank et al., 2014a; Blank & Fedorenko, 2017; Braga et al., 2020; Lerner et al., 2011; Silbert et al., 2014; Wilson et al., 2008), suggesting that they form a functionally coherent network (which I henceforth call *the language network*). The language network — if defined using methods that account for inter-individual variability in its location (see **Section 1.5.1**) — is extremely selective, responding to language input but not to mental arithmetic, music perception, executive function tasks, and action/gesture perception (Amalric & Dehaene, 2019; X. Chen et al., 2021; Fedorenko, Behr, et al., 2011; Jouravlev et al., 2019; Liu et al., 2020; Monti et al., 2009, 2012; Pritchett et al., 2018).

The other line of work highlighting the functional specificity of language regions is *evidence from people with global aphasia*, a major impairment of language affecting both production and comprehension. Despite the nearly complete loss of linguistic abilities, some individuals with global aphasia can appreciate music, solve arithmetic problems and logic puzzles, leverage their world knowledge to perform diverse tasks, orient themselves in space, reason about cause and effect, and navigate complex social situations (Bek et al., 2010; Klessinger et al., 2007; Varley, 1998; Varley et al., 2001, 2005; Varley & Siegal, 2000; Willemets et al., 2011). This line of evidence is consistent with earlier reports of intact music skills (e.g., Basso & Capitani, 1985; Luria et al., 1965), conceptual reasoning (e.g., Caramazza et al., 1982), and planning and social skills (Lecours & Joanette, 1980) in the presence of severe language impairment. The rich line of work examining language processing in individuals with aphasia dates back to 19<sup>th</sup> century, when several physicians treating individuals with brain damage (most notably, Broca and Wernicke) noted that certain lesions, localized to the left hemisphere, caused an impairment in language without affecting reasoning (Broca, 1861, 1865; Dax, 1865; Lichtheim, 1885; Wernicke, 1874).

Thus, converging evidence from neuroimaging and individuals with aphasia points to the functional independence of mechanisms underlying language processing, lending

strong support to H1. However, we cannot yet fully discard the alternative hypotheses. H2 states that the mechanisms for syntactic processing (or, more broadly, symbolic combination) are shared between language and other cognitive domains. Prior work has shown that syntactic combination is not fully domain-general: math and music processing do not rely on the language regions (Amalric & Dehaene, 2019; Basso & Capitani, 1985; X. Chen et al., 2021; Fedorenko, Behr, et al., 2011; Varley et al., 2005; cf. A. D. Patel et al., 2008). Yet, it is possible that another domain, more related to language either structurally or functionally, might engage the regions that we currently consider to be language-specific. In this thesis, I probe two such domains: computer programming and visually presented events.

H3 states that language shares processing resources with conceptual reasoning. Although substantial evidence from aphasia indicates that these abilities dissociate in brain-damaged patients (e.g., Antonucci & Reilly, 2008; Chertkow et al., 1997; Dickey & Warren, 2015; Jefferies & Lambon Ralph, 2006; Saygin et al., 2004; cf. Saygin et al., 2003), a number of neuroimaging studies reported overlapping activation to conceptual tasks performed on linguistic and non-linguistic input (e.g., Baldassano et al., 2018; Devereux et al., 2013; Fairhall & Caramazza, 2013; Handjaras et al., 2017; Z. Hu et al., 2019; Jouen et al., 2015; Shinkareva et al., 2011; Thierry & Price, 2006; Vandenberghe et al., 1996; Visser et al., 2012; Wurm & Caramazza, 2019). Oftentimes this activation is observed in left frontal and temporal regions resembling the language network; however, deciding that these regions are indeed the same as the language network requires reverse inference, i.e. inferring function from anatomy (Poldrack, 2006, 2011). Reverse inference can be useful in certain cases (e.g., activity in the anatomic location corresponding to motor cortex reliably indicates motor circuit activity), but its use is complicated for frontal and temporal cortex, where the location of language-specific regions is not only debated among researchers (Tremblay & Dick, 2016), but also varies substantially between participants (Amunts et al., 1999, 2010; Fedorenko & Blank, 2020; Shashidhara, Spronkers, et al., 2019). Thus, to establish the involvement of the language regions in conceptual reasoning (or in any other cognitive process), it is first necessary to localize these regions in individual participants (Fedorenko et al., 2010; Fedorenko & Kanwisher, 2009).

## 1.4 Overview of this work

My thesis work builds upon existing neuroimaging and aphasia studies to probe the limits of language processing specificity. In **Chapter 2**, I test H2 by asking whether language shares processing mechanisms with another domain that uses hierarchically structured symbols — computer programming. If the combinatorial processing within the language regions is not language-specific, we should expect the language regions to engage when processing computer code — a domain that has closely related structure and that, due to its evolutionary recency, must leverage preexisting brain circuits instead of innate domain-specific machinery. In **Chapter 3** and **Chapter 4**, I test H3 by studying whether the language regions are engaged in conceptual reasoning over pictures (in the domains of object and event semantics, respectively). A similar logic holds: if conceptual reasoning leverages the same circuits as language processing, we should observe activity in the language regions during a conceptual task performed even on non-linguistic stimuli. In **Chapter 5**, I use a complementary approach: instead of localizing the language network and measuring its involvement in semantic reasoning, I localize brain regions engaged in conceptual reasoning independent of input domain (linguistic vs. non-linguistic) and then test whether they overlap with the language network.

The main tool I use in my work is functional MRI, which allows me to study brain activity with a relatively high degree of spatial resolution. In Chapter 4, I complement fMRI data with behavioral evidence from individuals with global aphasia, which allows me to test the causal link between language processing and another aspect of human cognition.

## 1.5 Approach

### 1.5.1 Localizing the language network in individual brains

#### 1.5.1.1 The benefits of functional localization

My neuroimaging work uses the approach known as functional localization, in which I define the brain network of interest using a separate task (called a localizer) instead of

pre-selecting regions of interest based solely on anatomy. Functional localization allows the researchers to overcome reliance on anatomical landmarks when establishing whether a given region/network is engaged in a certain task (Poldrack, 2006, 2008) while still providing a stable, reliable way for identifying regions of interest across studies (Saxe, Brett, et al., 2006). It also accounts for individual variability in the functional architecture of the brain (Brett et al., 2002; Nieto-Castañón & Fedorenko, 2012; Saxe, Brett, et al., 2006): instead of using a group activation map, the network is defined in each individual participant.

Localizers are designed to elicit a consistent activation pattern in the brain and are typically robust to changes in their setup. For instance, the language network localizer elicits consistent activation regardless of change in materials, task, language, and spoken vs. written stimuli (Cheung et al., 2020; Fedorenko et al., 2010; Malik-Moraleda, Ayyash et al, 2022; Scott et al., 2017), the multiple demand network localizer elicits a consistent response pattern during a spatial working memory task, a memory probe task on nonwords (harder) vs. sentences (easier), and an arithmetic task (Fedorenko et al., 2013), and the theory of mind network localizer elicits the same response pattern during both verbal and non-verbal narratives that recruit theory of mind reasoning (Jacoby et al., 2016).

The precision afforded by the functional localization method is especially important when testing a hypothesis about functional specificity. The key question of this thesis is whether the language regions are engaged in non-linguistic cognition. Traditional approaches in neuroimaging would address this question using group-level activation maps, which estimate each voxel's responses to conditions of interest across participants. However, if the same voxel belongs to different functional networks in different participants, the group-level approach will fail: it will show the same voxel responding to both conditions (e.g., sentences and computer code) even if, in reality, it responded only to sentences in a subset of participants and only to computer code in the other subset (Fedorenko & Kanwisher, 2009; Nieto-Castañón & Fedorenko, 2012; Shashidhara, Mitchell, et al., 2019). This issue is particularly pressing when studying associative cortex, which shows substantial inter-individual variability in the anatomical location of

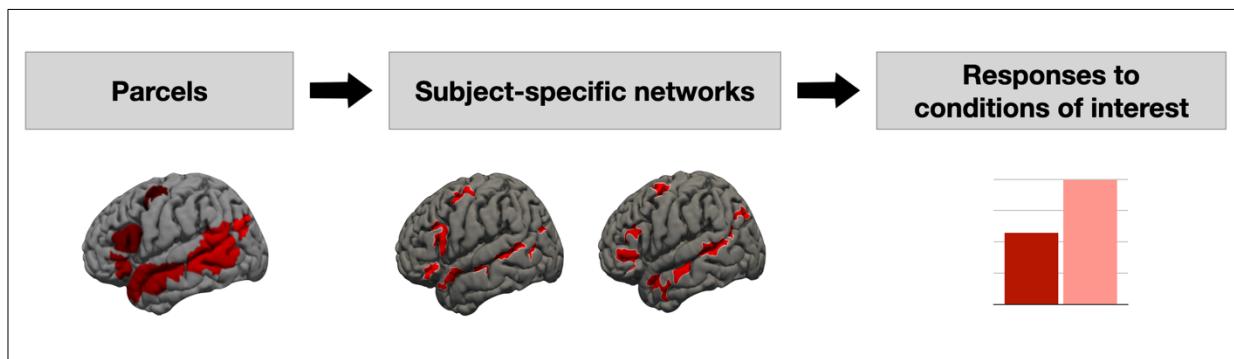
functionally specialized regions (Amunts et al., 1999; Caspers et al., 2006; Tahmasebi et al., 2012; Vázquez-Rodríguez et al., 2019). Thus, to determine true functional selectivity of a given brain region, we have to evaluate it at the level of individual participants rather than averaging responses across many brains.

### 1.5.1.2 Deriving and studying functional regions of interest

The general approach is shown in [Figure 1-1](#) (for details, see Fedorenko et al., 2010). I start with a set of predefined anatomical masks of brain regions (also called “parcels”), which mark the approximate location of the network I aim to examine. Then, in each participant, I select a subset of voxels within each parcel that respond most strongly to the contrast of interest from the localizer task. Thus, the anatomical location of the selected voxels is allowed to vary between participants, provided that it falls within the parcel. The main goal of using anatomical region masks to constrain voxel selection is to have consistent units of comparison across individuals and studies (Julian et al., 2012). The cutoff I use is the top 10% of voxels in each parcel, but the exact cutoff point is not critical as long as it is not overly permissive. The subset of a parcel selected according to this procedure is called a functional region of interest (fROI).

As is common in neuroimaging studies, the metric of interest used for voxel selection is not the voxels’ response to a single condition but rather a contrast value — the difference in responses to two conditions that are matched except for the property of interest. For the language localizer, the contrast I use is sentence reading > nonword reading. Nonwords are pronounceable strings of letters/characters in the same script as the sentences; thus, their visual properties are matched but their content differs (nonwords are syntactically and semantically empty). The exact contrast used to functionally localize the language network can vary: for instance, the spoken sentences > muffled spoken sentences contrast elicits the same activation pattern (Scott et al., 2017), which serves as verification that both contrasts isolate activity related to language, regardless of whether it is spoken or written. The language I use in the studies below is English.

Once the functional network is defined, we can then measure its response to any conditions of interest. We can analyze the network-level response (averaged across all fROIs that comprise the network), fROI-level responses (averaged across voxels within an fROI), or the pattern of voxel responses within each fROI. When analyzing the results, care must be taken that the data used to define the fROIs are different from the data used to derive key measures of interest, such as response magnitude (Kriegeskorte et al., 2009; Vul & Kanwisher, 2010).



**Figure 1-1.** The functional localization approach.

### 1.5.1.3 Pre-selecting anatomical regions of interest (parcels)

How do we decide which brain regions to include in the initial mask used to constrain the voxel search? Several options are possible. First, the regions can be defined anatomically based on previous reports of their involvement in the function of interest (see, e.g., Fedorenko et al., 2013). Second, they can be derived from prior studies quantitatively, e.g., by using a map of activations obtained from many labs and research paradigms (Yarkoni et al., 2011) or from many participants within the same paradigm (Lipkin et al., 2022; Van Essen et al., 2013). Finally, they can be derived from a subset of data collected for the purposes of the study (Fedorenko et al., 2010).

In the work described in this thesis, I use language and multiple demand network parcels derived from the data of many individuals who did the respective network localizer tasks ( $n=220$  for language and  $n=197$  for multiple demand). The procedure for deriving the parcels is part of the group-constrained subject-specific (GcSS) functional localization

approach described in Fedorenko et al (2010); in short, it involves creating a set of binary masks (parcels) from a probabilistic overlap map derived from binarized individual activation maps (thresholded at a prespecified p value). Voxels within each parcel are spatially contiguous.

When conducting whole-brain analyses that require generating new parcels, I use the GcSS procedure to create the parcels based on data from the critical task (e.g., putative computer-code-responsive regions), define fROIs within the parcels, and then test their responses to conditions of interest. In this case, the number of participants used to define the parcels is the same as the number of participants in that study.

### **1.5.2 Combining neuroimaging evidence from neurotypicals and behavioral evidence from aphasia**

Neuroimaging tools, such as functional MRI, provide a unique opportunity to map out the functional response profile of the language network. However, the inferences derived from fMRI are even more powerful when combined with evidence from individuals with language impairments, such as aphasia. Section 1.3 shows that many neuroimaging and aphasia studies provide converging evidence for language specificity in the brain. In Chapter 4, I build on both approaches by reporting data from fMRI studies on neurotypical individuals and from behavioral studies on individuals with aphasia.

Both fMRI and aphasia studies have certain shortcomings that might limit the use of these sources of evidence as tests of the functional specificity of the language regions. However, when used together, they provide complementary sources of evidence that can, in many cases, overcome these shortcomings. Functional MRI, as any neuroimaging method, is a correlational technique that cannot establish the causal link between observed activity and the behavior of interest. Thus, if we observe activity in the language regions in response to a certain task, we cannot establish whether this activity is causally important for that task. In contrast, evidence from aphasia can provide us with causal evidence of a functional dissociation between two functions: if a cognitive ability is preserved in spite

of severe language impairment, we can conclude that this ability does not depend on language processing (as I do in [Chapter 4](#)).

On the flip side, if an individual with aphasia exhibits concomitant deficits in a non-linguistic task, we cannot conclusively state that language shares neural processing mechanisms with that task. The pattern of brain damage after stroke (the most common cause of aphasia) does not respect functional borders and depends primarily on vasculature patterns. Thus, a lesion can easily destroy two or more functionally distinct regions that happen to lie nearby. This situation is common in the inferior frontal lobe, which is known to be functionally heterogeneous (Fedorenko, Duncan, et al., 2012; Fedorenko & Blank, 2020). In contrast, functional localization methods for fMRI allow us to precisely delineate the borders of functionally distinct brain regions and determine which of the functional networks in engaged in a given task. Thus, if evidence from aphasia indicate a correlation between impairment in two functions but neuroimaging shows these functions to be neurally distinct, we can attribute the shared impairment pattern to the fact that the lesion affected two different networks (as I do in [Chapter 3](#)).

Although not done here, correlational and causal approaches can sometimes be combined in a single participant group: individuals with aphasia can undergo functional MRI to have their brain networks localized (Blank et al., 2015, 2017) and neurotypical individuals can undergo both fMRI and a causal intervention that alters their brain activity for a short period of time, such as transcranial magnetic or direct current stimulation (e.g., Groen et al., 2021; Pitcher et al., 2011). Both are promising research directions, although their limitations should also be acknowledged. In case of aphasia studies, brain networks reorganize themselves after damage, and so the functional specialization patterns (and their corresponding cognitive mechanisms) observed in brains after stroke may differ from those observed in neurotypical individuals. For brain stimulation studies, transcranial stimulation techniques are spatially imprecise and, like brain lesions, might affect multiple functionally distinct regions that lie nearby; furthermore, they have limited potential to reach deeper brain structures or the ventral surface of the brain. In general, however, all these approaches reflect the abundant opportunities to investigate

the functional architecture of the human mind and brain using tools from cognitive neuroscience.

## 1.6 Using data from neuroscience to inform cognitive theories

The question addressed in this thesis is, in essence, a cognitive one: does one mental function (language) share processing mechanisms with others? Functionally localizing the language network provides a way to explicitly link this question to neuroscience evidence. A detailed characterization of the functionally defined language network across many studies allows us to infer function (language processing) from a brain activation pattern (voxels responding most strongly to the language localizer contrast). Thus, even though the localizer only covers a specific aspect of language (English sentence reading as contrasted with English-like nonword reading), prior work showing that this activation pattern remains the same regardless of modality, language, or task (Section 1.3) allows me to use localized regions as a functional marker of language processing (Mather et al., 2013).

A long line of research on the neural basis of language (Section 1.3) illustrates the tight link between cognitive theory and neuroscience. Cognitive theory is used to generate hypotheses about the functional organization of the brain (such as a putative shared mechanism for language and music processing; e.g., Baroni et al., 1983; Fay, 1971; Lerdahl & Jackendoff, 1996; Swain, 1995; cf. Jackendoff, 2009), and neuroscience findings, in turn, inform cognitive theory (music and language recruit different brain areas and are therefore processed by different mechanisms; e.g., Basso & Capitani, 1985; X. Chen et al., 2021; Fedorenko, McDermott, et al., 2012; Norman-Haignere et al., 2015; Rogalsky et al., 2011). This mutually enriching connection is observed in many areas of cognitive neuroscience (Mather et al., 2013).

A tight integration of cognitive theory and neuroscience has the potential to lead to evidence-based cognitive ontologies (Lenartowicz et al., 2010; Poldrack et al., 2011; C. J. Price & Friston, 2005), which allow us to generate mechanistic descriptions of cognitive processes as interactions between empirically validated cognitive modules. In the

concluding chapter (Section 6.2), I attempt to do just that. Specifically, I draw upon my work, along with many other cognitive neuroscience studies, to discuss the role of the language network in the broader system of cognitive mechanisms underlying language use. I argue that, not only are the core language processing mechanisms functionally specific to language, but they are also, on their own, insufficient for real-life language use. The language network has to interface with domain-general regions that track discourse-level information, with regions that carry out social reasoning and formal logic (so that the utterances are pragmatically and logically coherent), and possibly with semantic demand regions described in Chapter 5. All these regions are not language-specific — meaning that they respond to both verbal and nonverbal inputs — and can function even in the absence of a functional language system. Thus, instead of serving as a foundation for thought, language processing is only one tool in our rich cognitive toolbox.

# Chapter 2

## The language network & computer code comprehension

Collaborators:

Shashank Srikant, Yotaro Sueoka, Hope H. Kean, Riva Dhamala,  
Una-May O'Reilly, Marina U. Bers, Evelina Fedorenko

This chapter has been published as a journal article in *eLife* (under a CC BY 4.0 license):

Ivanova, A. A., Srikant, S., Sueoka, Y., Kean, H. H., Dhamala, R., O'Reilly, U. M., Bers, M. U. & Fedorenko, E. (2020). Comprehension of computer code relies primarily on domain-general executive brain regions. *eLife*, 9, e58906.

It is reproduced here with slight alterations.

### 2.1 Abstract

Computer programming is a novel cognitive tool that has transformed modern society. What cognitive and neural mechanisms support this skill? Here, we used fMRI to investigate two candidate brain systems: the multiple demand (MD) system, typically recruited during math, logic, problem solving, and executive tasks, and the language system, typically recruited during linguistic processing. We examined MD and language system responses to code written in Python, a text-based programming language (Experiment 1) and in ScratchJr, a graphical programming language (Experiment 2); for both, we contrasted responses to code problems with responses to content-matched sentence problems. We found that the MD system exhibited strong bilateral responses to code in both experiments, whereas the language system responded strongly to sentence problems, but weakly or not at all to code problems. Thus, the MD system supports the use of novel cognitive tools even when the input is structurally similar to natural language.

## 2.2 Introduction

The human mind is endowed with a remarkable ability to support novel cognitive skills, such as reading, writing, map-based navigation, mathematical reasoning, and scientific logic. Recently, humanity has invented another powerful cognitive tool: computer programming. The ability to flexibly instruct programmable machines has led to a rapid technological transformation of communities across the world (Ensmenger, 2012); however, little is known about the cognitive and neural systems that underlie computer programming skills.

Here, we investigate which neural systems support one critical aspect of computer programming: computer code comprehension. By code comprehension, we refer to a set of cognitive processes that allow programmers to interpret individual program tokens (such as keywords, variables, and function names), combine them to extract the meaning of program statements, and, finally, combine the statements into a mental representation of the entire program. It is important to note that code comprehension may be cognitively and neurally separable from cognitive operations required to process program content, i.e., the actual operations described by code. For instance, to predict the output of the program that sums the first three elements of an array, the programmer should identify the relevant elements and then mentally perform the summation. Most of the time, processing program content recruits a range of cognitive processes known as computational thinking (Wing, 2006, 2011), which include algorithm identification, pattern generalization/abstraction, and recursive reasoning (e.g., Kao, 2011). These cognitive operations are notably different from code comprehension per se and may not require programming knowledge at all (Guzdial, 2008). Thus, research studies where people read computer programs should account for the fact that interpreting a computer program involves two separate cognitive phenomena: processing computer code that comprises the program (i.e., code comprehension) and mentally simulating the procedures described in the program (i.e., processing problem content).

Given that code comprehension is a novel cognitive tool, typically acquired in late childhood or in adulthood, we expect it to draw on preexisting cognitive systems.

However, the question of which cognitive processes support code comprehension is non-trivial. Unlike some cognitive inventions that are primarily linked to a single cognitive domain (e.g., reading/writing building on spoken language), code comprehension plausibly bears parallels to multiple distinct cognitive systems. First, it may rely on domain-general executive resources, including working memory and cognitive control (Bergersen & Gustafsson, 2011; Nakagawa et al., 2014; Nakamura et al., 2003). In addition, it may draw on the cognitive systems associated with math and logic (McNamara, 1967; Papert, 1972), in line with the traditional construal of coding as problem-solving (Dalbey & Linn, 1985; Ormerod, 1990; Pea & Kurland, 1984; Pennington & Grabowski, 1990). Finally, code comprehension may rely on the system that supports comprehension of natural languages (Fedorenko et al., 2019; Murnane, 1993; Papert, 1993). Like natural language, computer code makes heavy use of hierarchical structures (e.g., loops, conditionals, and recursive statements), and, like language, it can convey an unlimited amount of meaningful information (e.g., describing objects or action sequences). These similarities could, in principle, make the language circuits well-suited for processing computer code.

Neuroimaging research is well positioned to disentangle the relationship between code comprehension and other cognitive domains. Many cognitive processes are known to evoke activity in specific brain regions/networks: thus, observing activity for the task of interest in a particular region or network with a known function can indicate which cognitive processes are likely engaged in that task (Mather et al., 2013). Prior research (Assem, Glasser, et al., 2020; Duncan, 2010, 2013; Duncan & Owen, 2000) has shown that executive processes — such as attention, working memory, and cognitive control — recruit a set of bilateral frontal and parietal brain regions collectively known as the multiple demand (MD) system. If code comprehension primarily relies on domain-general executive processes, we expect to observe code-evoked responses within the MD system, distributed across both hemispheres. Math and logic also evoke responses within the MD system (Fedorenko et al., 2013), although this activity tends to be left-lateralized (Amalric & Dehaene, 2016, 2019; Goel & Dolan, 2001; Micheloyannis et al., 2005; Monti et al., 2007, 2009; Pinel & Dehaene, 2009; Prabhakaran et al., 1997; Reverberi et al., 2009). If code comprehension draws on the same mechanisms as math and logic, we expect to

observe left-lateralized activity within the MD system. Finally, comprehension of natural language recruits a set of left frontal and temporal brain regions known as the language system (e.g., Fedorenko & Thompson-Schill, 2014). These regions respond robustly to linguistic input, both visual and auditory (Deniz et al., 2019; Fedorenko et al., 2010; Nakai et al., 2020; M. Regev et al., 2013; Scott et al., 2017). However, they show little or no response to tasks in non-linguistic domains, such as executive functions, math, logic, music, action observation, or non-linguistic communicative signals, such as gestures (Fedorenko, Behr, et al., 2011; Jouravlev et al., 2019; Monti et al., 2009, 2012; Pritchett et al., 2018; see Fedorenko & Blank, 2020, for a review). If code comprehension relies on the same circuits that map form to meaning in natural language, we expect to see activity within the language system.

Evidence from prior neuroimaging investigations of code comprehension is inconclusive. Existing studies have provided some evidence for left-lateralized activity in regions that roughly correspond to the language system (Siegmund et al., 2014, 2017), as well as some evidence for the engagement of frontal and parietal regions resembling the MD system (Floyd et al., 2017; Huang et al., 2019; Siegmund et al., 2014, 2017). However, none of these prior studies sought to explicitly distinguish code comprehension from other programming-related processes, and none of them provide quantitative evaluations of putative shared responses to code and other tasks, such as working memory, math, or language (cf. Liu et al., 2020; see Discussion).

Here, we use functional magnetic resonance imaging (fMRI) to evaluate the role of the MD system and the language system in computer code comprehension. Three design features that were lacking in earlier neuroimaging studies of programming allow us to evaluate the relative contributions of these two candidate systems. First, we contrast neural responses evoked by code problems with those evoked by content-matched sentence problems (**Figure 2-1, A**); this comparison allows us to disentangle activity evoked by code comprehension from activity evoked by the underlying program content (which is matched across code and sentence problems).

Second, we use independent ‘localizer’ tasks (Brett et al., 2002; Fedorenko et al., 2010; Saxe, Brett, et al., 2006) to identify our networks of interest: a working memory task to localize the MD system and a passive reading task to localize the language system (**Figure 2-1, B**). The functional localization approach obviates the reliance on the much-criticized ‘reverse inference’ reasoning (Poldrack, 2006, 2011), whereby functions are inferred from coarse macro-anatomical landmarks. Instead, we can directly interpret code-evoked activity within functionally defined regions of interest (Mather et al., 2013). In addition, localization of the MD and language networks is performed in individual participants, which is important given substantial variability in their precise locations across individuals (Fedorenko & Blank, 2020; Shashidhara, Spronkers, et al., 2019) and leads to higher sensitivity and functional resolution (Nieto-Castañón & Fedorenko, 2012).

Third, to draw general conclusions about code comprehension, we investigate two very different programming languages: Python, a popular general-purpose programming language, and ScratchJr, an introductory visual programming language for creating animations designed for young children (Bers & Resnick, 2015). In the Python experiment, we further examine two problem types (math problems and string manipulation) and three basic types of program structure (sequential statements, *for* loops, and *if* statements). Comprehension of both Python and ScratchJr code requires retrieving the meaning of program tokens and combining them into statements, despite the fact that the visual features of the tokens in the two languages are very different (text vs. images). If a brain system is involved in code comprehension, we expect its response to generalize across programming languages and problem types, similar to how distinct natural languages in bilinguals and multilinguals draw on the same language regions (Kroll et al., 2015).

Taken together, these design features of our study allow us to draw precise and generalizable conclusions about the neural basis of code comprehension.

## A Experiment 1 - Python

**code problem**

```
height = 5
weight = 100
bmi = weight/(height*height)
print(bmi)
```

**sentence problem**

Your height is 5 feet and your weight is 100 pounds. The BMI is defined as the ratio between the weight and the square of the height of a person. What is your BMI?

## Experiment 2 - ScratchJr

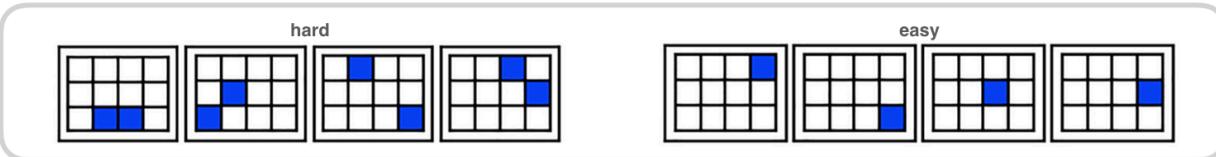
**code problem**

**sentence problem**

Kitten walks right, jumps, and then walks left.

## B

### MD System Localizer



### Language System Localizer

sentence	nonwords
NOBODY COULD HAVE PREDICTED THE EARTHQUAKE	U BIZBY ACWORRILY MIDARAL MAPE LAS POME

**Figure 2-1.** Experimental paradigms. (A) Main task. During code problem trials, participants were presented with snippets of code in Python (Experiment 1) or ScratchJr (Experiment 2); during sentence problem trials, they were presented with text problems that were matched in content with the code stimuli. Each participant saw either the code or the sentence version of any given problem. (B) Localizer tasks. The MD localizer (top) included a hard condition (memorizing positions of 8 squares appearing two at a time) and an easy condition (memorizing positions of 4 squares appearing one at a time). The language localizer (bottom) included a sentence reading and a nonword reading condition, with the words/nonwords appearing one at a time.

## 2.3 Results

Participants performed a program comprehension task inside an MRI scanner. In each trial, participants, all proficient in the target programming language, read either a code problem or a content-matched sentence problem (**Figure 2-1, A**) and were asked to predict the output. In Experiment 1 (24 participants, 15 women), code problems were written in Python, a general-purpose text-based programming language (Sanner, 1999).

In Experiment 2 (19 participants, 12 women), code problems were written in ScratchJr, an introductory graphical programming language developed for children aged 5-7 (Bers, 2018). Both experiments were conducted with adults to facilitate result comparison. Good behavioral performance confirmed that participants were proficient in the relevant programming language and engaged with the task (Python: 99.6% response rate, 85% accuracy on code problems; ScratchJr: 98.6% response rate, 79% accuracy on code problems; see **Supplemental figure 2-2** for detailed behavioral results). Participants additionally performed two functional localizer tasks: a hard vs. easy spatial working memory task, used to define the MD system, and a sentence vs. nonword reading task, used to define the language system (**Figure 2-1, B**; see **Section 2.5** for details).

We then contrasted neural activity in the MD and language systems during code problem comprehension with activity during (a) sentence problem comprehension and (b) the nonword reading condition from the language localizer task. Sentence problem comprehension requires simulating the same operations as code problem comprehension (mathematical operations or string manipulation for Python, video simulation for ScratchJr), so contrasting code problems with sentence problems allows us to isolate neural responses evoked by code comprehension from responses evoked by processing problem content. Nonword reading elicits weak responses in both the language system and the MD system (in the language system, this response likely reflects low-level perceptual and/or phonological processing; in the MD system, it likely reflects the basic task demands associated with maintaining attention or reading pronounceable letter strings). Because the nonword response is much weaker than responses to the localizer conditions of interest (Fedorenko et al., 2010; Mineroff et al., 2018), nonword reading can serve as a control condition for both the MD and language systems, providing a more stringent baseline than simple fixation. Given the abundant evidence that the MD system and the language system are each strongly functionally interconnected (Blank et al., 2014b; Mineroff et al., 2018; Paunov et al., 2019), we perform the key analyses at the system level.

### 2.3.1 MD system exhibits robust and generalizable bilateral responses during code comprehension

We found strong bilateral responses to code problems within the MD system in both Experiments 1 and 2 (**Figure 2-2, Figure 2-3**). These responses were stronger than responses to both the sentence problem condition (Python:  $\beta = 1.03$ ,  $p < 0.001$ , ScratchJr:  $\beta = 1.38$ ,  $p < 0.001$ ) and the control nonword reading condition (Python:  $\beta = 2.17$ ,  $p < 0.001$ ; ScratchJr:  $\beta = 1.23$ ,  $p < 0.001$ ). The fact that code problems drove the MD system more strongly than content-matched sentence problems (despite the fact that sentence problems generally took longer to respond to; see **Supplemental figure 2-2**) demonstrate that the MD system responds to code comprehension specifically rather than simply being activated by the underlying problem content.

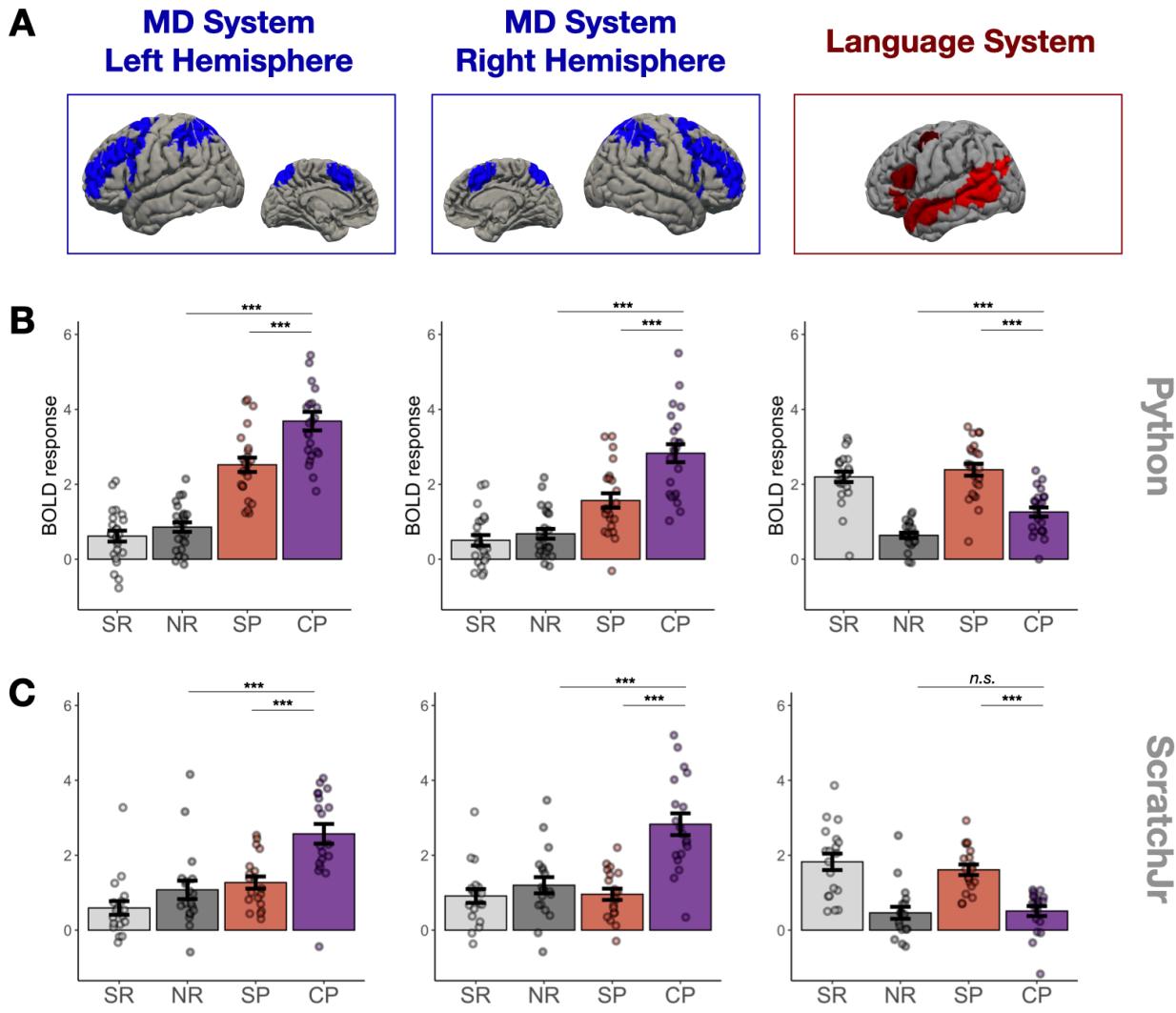
To further test the generalizability of MD responses, we capitalized on the fact that our Python stimuli systematically varied along two dimensions: (1) problem type (math problems vs. string manipulation), and (2) problem structure (sequential statements, *for* loops, *if* statements). Strong responses were observed in the MD system (**Figure 2-4, A, B**) regardless of problem type ( $\beta = 3.02$ ,  $p < 0.001$ ; no difference between problem types) and problem structure ( $\beta = 3.14$ ,  $p < 0.001$ ; sequential problems evoked a slightly weaker response,  $\beta = -0.20$ ,  $p = 0.002$ ). This analysis demonstrates that the responses were not driven by one particular type of problem or by mental operations related to the processing of a particular code structure.

We also tested whether MD responses to code showed a hemispheric bias similar to what is typically seen for math and logic problems (Goel & Dolan, 2001; Micheloyannis et al., 2005; Monti et al., 2007, 2009; Pinel & Dehaene, 2009; Prabhakaran et al., 1997; Reverberi et al., 2009). Neither Python nor ScratchJr problems showed a left-hemisphere bias for code comprehension. For Python, the size of the code problems>sentence problems effect did not interact with hemisphere ( $\beta = 0.11$ ,  $p = 0.46$ ), even though the magnitude of responses to code problems as compared to nonword reading was stronger in the left hemisphere ( $\beta = 0.63$ ,  $p < 0.001$ ). These results show that neural activity evoked by Python code comprehension was bilaterally distributed but that activity evoked by the

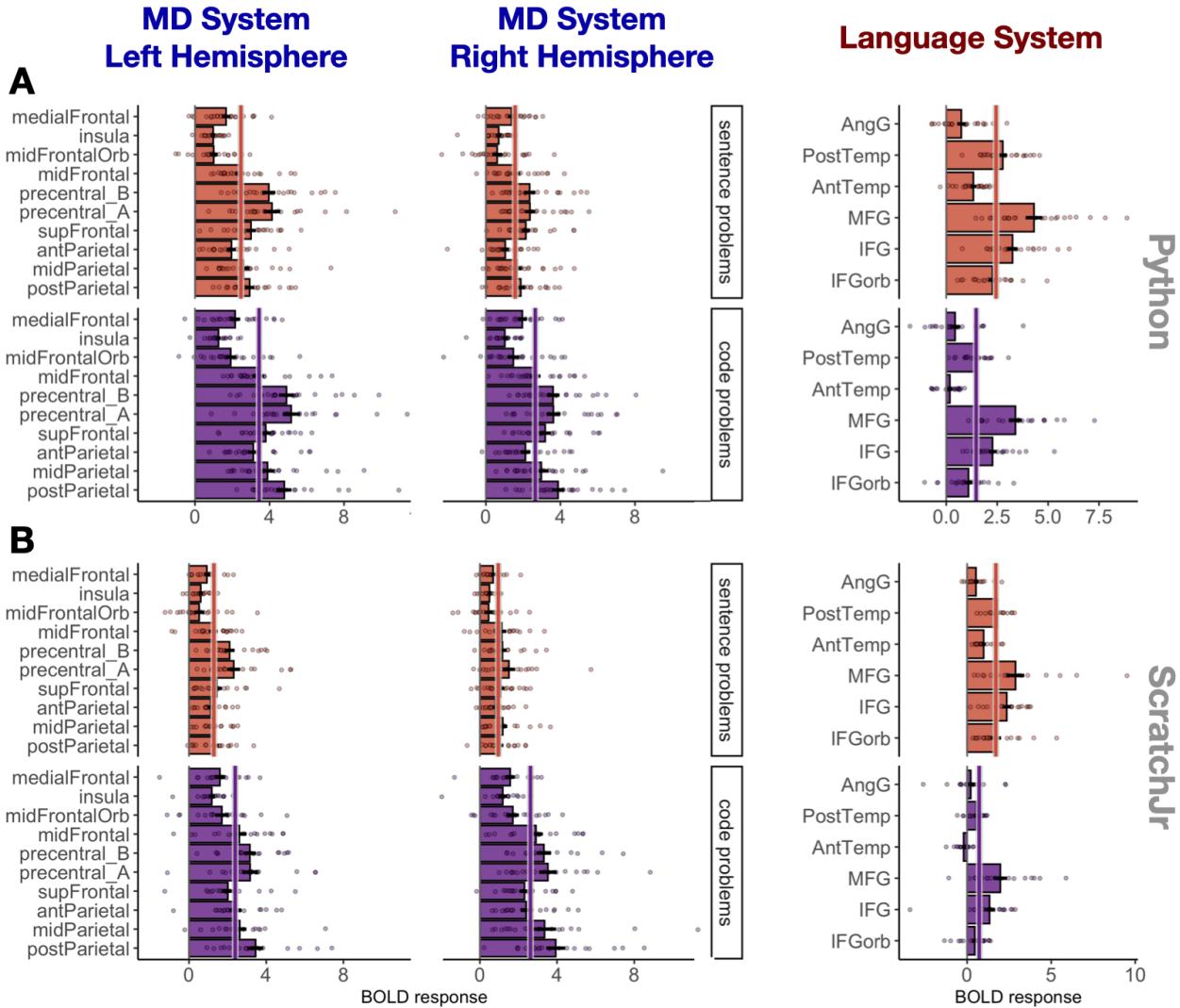
underlying problem content was left-lateralized. For ScratchJr, the size of the code problems>sentence problems effect interacted with hemisphere, with stronger responses in the right hemisphere ( $\beta = 0.57$ ,  $p = 0.001$ ), perhaps reflecting the bias of the right hemisphere toward visuo-spatial processing (Corballis, 2003; Hugdahl, 2011; Sheremata et al., 2010).

Follow-up analyses of activity in individual regions within the MD system demonstrated that 17 of the 20 MD fROIs (all except the fROIs located in left medial frontal cortex and in the left and right insula) responded significantly more strongly to Python code problems than to sentence problems (see **Supplemental table 2-1** for all fROI statistics). Responses to ScratchJr were significantly stronger than responses to sentence problems in 6 of the 10 left hemisphere MD fROIs (the effect was not significant in the fROIs in superior frontal gyrus, the dorsal part of the precentral gyrus, the medial frontal cortex, and the insula) and in 8 of the 10 right hemisphere MD fROIs (the effect was not significant in the fROIs in the medial frontal cortex and the insula; see **Supplemental table 2-2** for all fROI statistics). These analyses demonstrate that code processing is broadly distributed across the MD system rather than being localized to a particular region or to a small subset of regions.

Overall, we show that MD responses to code are strong, do not exclusively reflect responses to problem content, generalize across programming languages and problem types, and are observed across most MD fROIs.



**Figure 2-2.** (A) Candidate brain systems of interest. The areas shown represent the “parcels” used to define the MD and language systems in individual participants (see **Section 2.5** and **Supplemental figure 2-5**). (B, C) Mean responses to the language localizer conditions (SR - sentence reading and NR - nonwords reading) and to the critical task (SP - sentence problems and CP - code problems) in systems of interest across programming languages (**B** – Python, **C** - ScratchJr). In the MD system, we see strong responses to code problems in both hemispheres and to both programming languages; the fact that this response is stronger than the response to content-matched sentence problems suggests that it reflects activity evoked by code comprehension per se rather than just activity evoked by problem content. In the language system, responses to code problems elicit a response that is substantially weaker than that elicited by sentence problems; further, only in Experiment 1 do we observe responses to code problems that are reliably stronger than the responses to the language localizer control condition (nonword reading). Here and elsewhere, error bars show standard error of the mean across participants, and the dots show individual participants’ responses.



**Figure 2-3.** Responses to sentence problems (red) and code problems (purple) during Experiment 1 (Python; A) and Experiment 2 (ScratchJr; B) broken down by region within each system. Abbreviations: mid – middle, ant – anterior, post – posterior, orb – orbital, MFG – middle frontal gyrus, IFG – inferior frontal gyrus, temp – temporal lobe, AngG – angular gyrus, precentral\_A – the dorsal portion of precentral gyrus, precentral\_B – the ventral portion of precentral gyrus. A solid line through the bars in each sub-plot indicates the mean response across the fROIs in that plot.

### 2.3.2 No MD fROIs are selective for code comprehension

To determine whether any fROIs were driven selectively (preferentially) by code problems relative to other cognitively demanding tasks, we contrasted individual fROI responses to code problems with responses to a hard working memory task from the MD

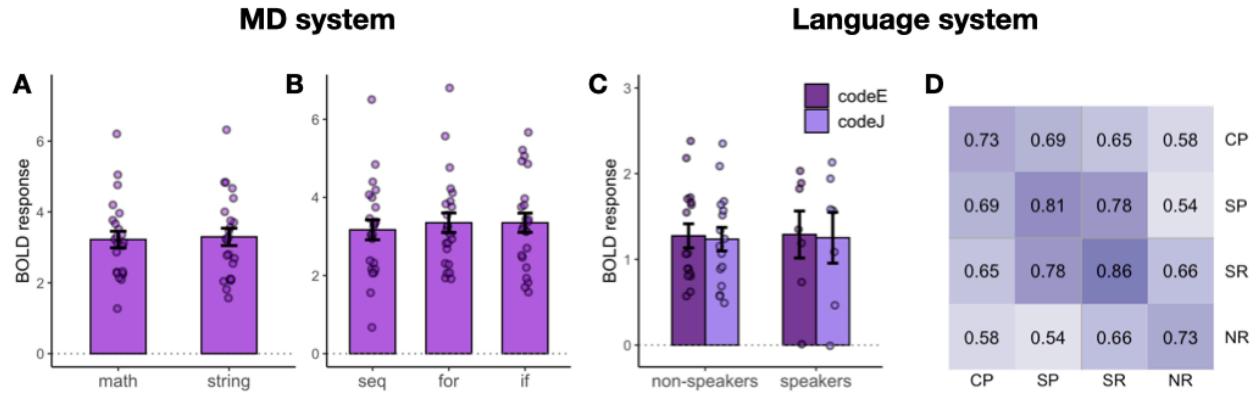
localizer experiment. Three fROIs located in the left frontal lobe (“precentral\_A”, “precentral\_B”, and “midFrontal”) exhibited stronger responses to Python code problems than to the hard working memory task ( $\beta = 1.21$ ,  $p < 0.001$ ;  $\beta = 1.89$ ,  $p < 0.001$ ; and  $\beta = 0.79$ ,  $p = 0.011$ , respectively; **Supplemental figure 2-6**). However, the magnitude of the code problems > sentence problems contrast in these regions ( $\beta = 1.03$ ,  $0.95$ ,  $0.97$ ) was comparable to the average response magnitude across all MD fROIs (average  $\beta = 1.03$ ), suggesting that the high response was caused by processing the underlying problem content rather than by code comprehension per se. Furthermore, neither these nor any other MD fROIs exhibited higher responses to ScratchJr code compared to the hard working memory task (in fact, the “precentral\_A” fROI did not even show a significant code problems>sentence problems effect). We conclude that code comprehension is broadly supported by the MD system (similarly to, e.g., intuitive physical inference; Fischer et al., 2016), but no MD regions are functionally specialized to process computer code.

### **2.3.3 Language system responses during code comprehension are weak and inconsistent**

The responses to code problems within the language system 2 (**Figure 2-2**, **Figure 2-3**) were weaker than responses to sentence problems in both experiments (Python:  $\beta = 0.98$ ,  $p < 0.001$ ; ScratchJr:  $\beta = 0.99$ ,  $p < 0.001$ ). Furthermore, although the responses to code problems were stronger than the responses to nonword reading for Python ( $\beta = 0.78$ ,  $p < 0.001$ ), this was not the case for ScratchJr ( $\beta = 0.15$ ,  $p = 0.29$ ), suggesting that the language system is not consistently engaged during computer code comprehension.

We further tested whether responses to Python code problems within the language system may be driven by the presence of English words. Our stimuli were constructed such that half of the Python problems contained meaningful identifier names, and in the other half, the English identifiers were replaced with their Japanese translations, making them semantically meaningless for non-speakers of Japanese. For this analysis, we divided our participants into two groups — those with no reported knowledge of Japanese (N=18) and those with some knowledge of Japanese (N=6) — and compared

responses within their language regions to code problems with English vs. Japanese identifiers (**Figure 2-4, C**). We found no effect of identifier language ( $\beta = 0.03$ ,  $p = 0.84$ ), knowledge of Japanese ( $\beta = 0.03$ ,  $p = 0.93$ ), or interaction between them ( $\beta = 0.09$ ,  $p = 0.71$ ), indicating that the language system's response to Python code was not driven by the presence of semantically transparent identifiers. This result is somewhat surprising given the language system's strong sensitivity to word meanings (e.g., Anderson et al., 2019; Binder et al., 2009; Fedorenko et al., 2010, 2020; Pereira et al., 2018). One possible explanation is that participants do not deeply engage with the words' meanings in these problems because these meanings are irrelevant to finding the correct solution.



**Figure 2-4.** Follow-up analyses of responses to Python code problems. (A) MD system responses to math problems vs. string manipulation problems. (B) MD system responses to code with different structure (sequential vs. for loops vs. if statements). (C) Language system responses to code problems with English identifiers (codeE) and code problems with Japanese identifiers (codeJ) in participants with no knowledge of Japanese (non-speakers) and some knowledge of Japanese (speakers) (see Section 2.3.3 for details of this manipulation). (D) Spatial correlation analysis of voxel-wise responses within the language system during the main task (SP – sentence problems and CP – code problems) with the language localizer conditions (SR – sentence reading and NR – nonwords reading). Each cell shows a correlation between the activation patterns for each pair of conditions. Within-condition similarity is estimated by correlating activation patterns across independent runs.

Finally, we investigated whether the responses to Python code problems within the language system were driven by code comprehension specifically or rather by the underlying problem content. When examining responses in the MD system, we could easily disentangle the neural correlates of code comprehension vs. the processing of

problem content using univariate analyses: the code problems > sentence problems contrast isolated code comprehension-related processes, and the sentence problems > nonword reading contrast isolated responses to problem content. In the language system, however, the sentence problems > nonword reading response is additionally driven by language comprehension (unlike the MD system, which does not respond to linguistic input in the absence of task demands, as evidenced by its low responses during sentence reading; see also Blank & Fedorenko, 2017; Diacheck et al., 2020). Thus, responses to Python code might be evoked both by problem content and by the language-like features of Python code. To determine the relative contributions of these two factors, we computed voxel-wise spatial correlations within and between the code problem and sentence problem conditions, as well as correlations between these conditions and the sentence/nonword reading conditions from the language localizer task (**Figure 2-4, D**). We reasoned that if a system is driven by problem content, the activation patterns for code and sentence problems should be similar; in contrast, if a system is driven by code comprehension per se, the activation patterns for code and sentence problems should differ. We found that the activation patterns were highly correlated between the code and sentence problems ( $r = 0.69$ ,  $p < 0.001$ ). These correlation values were higher than the correlations between code problems and sentence reading (0.69 vs. 0.65;  $p < 0.001$ ), although lower than the correlations within the code problem condition (0.69 vs. 0.73;  $p < 0.001$ ). The fact that code and sentence problem responses are correlated over and above code problem and sentence reading responses indicates that the language system is sensitive to the content of the stimulus rather than just the stimulus type (code vs. words). Moreover, similarly to the MD system, problem content can account for a substantial portion of the response in the language regions ( $\Delta r = 0.04$ ). Note that a similar spatial correlation analysis in the MD system mirrored the result of univariate analyses (**Supplemental figure 2-9**). Thus, in both MD and language systems, response to Python code is driven both by problem content and by code-specific responses.

Overall, we found that the language system responded to code problems written in Python but not in ScratchJr. Furthermore, Python responses were driven not only by code comprehension, but also by the processing of problem content. We conclude that

successful comprehension of computer code can proceed without engaging the language network.

### **2.3.4 No consistent evidence of code-responsive regions outside the MD/language systems**

To search for code-responsive regions that might fall outside the MD and language systems, we performed a whole-brain GSS analysis (Fedorenko et al., 2010). GSS analysis serves the same goal as the traditional random-effects voxel-wise analysis (A. P. Holmes & Friston, 1998) but accommodates inter-individual variability in the precise locations of functional regions, thus maximizing the likelihood of finding responsive regions (Nieto-Castañón & Fedorenko, 2012). We searched for areas of activation for the code problems>sentence problems contrast (separately for Python and ScratchJr) that were spatially similar across participants. We then examined the response of such regions to code and sentence problems (using an across-runs cross-validation procedure; e.g., Nieto-Castañón & Fedorenko, 2012), as well as to conditions from the two localizer experiments. In both experiments, the discovered regions spatially resembled the MD system (**Supplemental figure 2-7** and **Supplemental figure 2-8**). For Python, any region that responded to code also responded to the spatial working memory task (the MD localizer). In case of ScratchJr, some fROIs responded more strongly to code problems than to the spatial working memory task; these fROIs were located in early visual areas/ventral visual stream and therefore likely responded to low-level visual properties of ScratchJr code (which includes colorful icons, objects, etc.). The traditional random-effects group analyses revealed a similar activation pattern (**Supplemental figure 2-3** and **Supplemental figure 2-4**). These whole-brain analyses demonstrate that the MD system responds robustly and consistently to computer code, recapitulating the results of the fROI-based analyses (**Figures 2.2-2.4**), and show that fROI-based analyses did not miss any non-visual code-responsive or code-selective regions outside the boundaries of the MD system.

### 2.3.5 Effect of proficiency on MD and language responses

We conducted an exploratory analysis to check whether engagement of the MD and/or language system in code comprehension varies with the level of programming expertise. We correlated responses within each system with independently obtained proficiency scores for Experiment 1 participants (see the paper's website for details: <https://github.com/ALFA-group/neural-program-comprehension>) and with in-scanner accuracy scores for Experiment 2 participants. No correlations were significant (see **Supplemental figure 2-10**). However, due to a relatively low number of participants ( $N=24$  and  $N=19$ , respectively), these results should be interpreted with caution.

## 2.4 Discussion

The ability to interpret computer code is a remarkable cognitive skill that bears parallels to diverse cognitive domains, including general executive functions, math, logic, and language. The fact that coding can be learned in adulthood suggests that it may rely on existing cognitive systems. Here, we tested the role of two candidate neural systems in computer code comprehension: the domain-general multiple demand (MD) system (Duncan, 2010) that has been linked to diverse executive demands and implicated in math and logic (e.g., Amalric & Dehaene, 2019; Goel, 2007; Monti et al., 2007, 2009), and the language-selective system (Fedorenko et al., 2011) that has been linked to lexical and combinatorial linguistic processes (e.g., Bautista & Wilson, 2016; Fedorenko et al., 2010, 2020b; Fedorenko, Nieto-Castañon, et al., 2012; Keller et al., 2001; Mollica et al., 2020). We found robust bilateral responses to code problems within the MD system, a pattern that held across two very different programming languages (Python and ScratchJr), types or problems (math and string manipulation), and problem structure (sequential statements, *for* loops, and *if* statements). In contrast, responses in the language system were substantially lower than those elicited by the content-matched sentence problems and exceeded responses to the control condition (nonwords reading) only for one of the two programming languages tested.

Our work uniquely contributes to the study of computer programming in the mind and brain by addressing two core issues that made it difficult to interpret results from prior studies. First, we disentangle responses evoked by code comprehension from responses to problem content (which is often not code-specific) by contrasting code problems with content-matched sentence problems. Our findings suggest that earlier reports of left-lateralized code-evoked activity (Siegmund et al., 2014) may reflect processing program content rather than code comprehension per se. This distinction should also be considered when interpreting results of other studies of programming effects on brain activity, such as debugging (Castelhano et al., 2019), variable tracking (Ikutani & Uwano, 2014; Nakagawa et al., 2014), use of semantic cues or program layout (Fakhoury et al., 2018; Siegmund et al., 2017), program generation (Krueger et al., 2020), and programming expertise (Ikutani et al., 2020).

Second, we analyze responses in brain areas that are functionally localized in individual participants, allowing for straightforward interpretation of the observed responses (Mather et al., 2013; Saxe, Brett, et al., 2006). This approach stands in contrast to the traditional approach, whereby neural responses are averaged across participants on a voxel-by-voxel basis, and the resulting activation clusters are interpreted via ‘reverse inference’ from anatomy (e.g., Poldrack, 2006, 2011). Functional localization is particularly important when analyzing responses in frontal, temporal, and parietal association cortex, which is known to be functionally heterogeneous and variable across individuals (Blank et al., 2017; Braga et al., 2019; Fedorenko & Kanwisher, 2009; Frost & Goebel, 2012; Shashidhara, Spronkers, et al., 2019; Tahmasebi et al., 2012; Vázquez-Rodríguez et al., 2019).

The results of our work align well with the results of another recent study on program comprehension (Liu et al., 2020). Liu et al. investigated the neural correlates of program comprehension by contrasting Python code problems with fake code. The code problem condition was similar to ours, whereas the fake code condition involved viewing scrambled code, followed by a visual recognition task. The code problems > fake code contrast is broader than ours: it includes both code comprehension (interpreting Python code) and the processing of problem content (manipulating characters in a string). Our

results show that the MD system is involved in both processes, but Python code comprehension is bilateral, whereas the processing of problem content is left-lateralized. We would therefore expect the code problems > fake code contrast to activate the MD system, engaging the left hemisphere more strongly than the right due to the demands of problem content processing. This is precisely what Liu et al. find. Further, similar to us, Liu et al. conclude that it is the MD regions, not the language regions, that are primarily involved in program comprehension.

#### **2.4.1 MD system's engagement reflects the use of domain-general resources**

The fact that the MD system responds to code problems over and above content-matched sentence problems underscores the role of domain-general executive processes in code comprehension. Although cognitive processes underlying code interpretation bear parallels to logic and math tasks (Papert, 1972; Pennington & Grabowski, 1990; Perkins & Simmons, 1988) and to natural language comprehension/generation (Fedorenko et al., 2019; Hermans & Aldewereld, 2017), the neural activity we observe primarily resembles activity observed in response to domain-general executive tasks (Assem, Glasser, et al., 2020; Duncan, 2010; Fedorenko et al., 2013). In particular, code comprehension elicits bilateral responses within the MD system, in contrast to math and logic tasks that tend to elicit left-lateralized responses within the MD system, and in contrast to language tasks that elicit responses in the spatially and functionally distinct language system.

We found that responses in the MD system were driven both by the processing of problem content (e.g., summing the contents of an array) and by code comprehension (e.g., identifying variables referring to an array and its elements, interpreting a *for* loop, realizing that the output of the program is the variable being updated inside the *for* loop). Both of these processes plausibly require attention, working memory, inhibitory control, planning, and general flexible relational reasoning—cognitive processes long linked to the MD system (Duncan, 2010, 2013; Duncan & Owen, 2000; Miller & Cohen, 2001) in both humans (Assem, Glasser, et al., 2020; Shashidhara, Mitchell, et al., 2019; Woolgar et al., 2018b) and non-human primates (Freedman et al., 2001; Miller et al., 1996; D. J.

Mitchell et al., 2016). A recent study (Huang et al., 2019) reported neural overlap between operations on programming data structures (which require both code comprehension and the processing of problem content) and a mental rotation task (which requires spatial reasoning). The overlap was observed within brain regions whose topography grossly resembles that of the MD system. In our study, all code-responsive brain regions outside the visual cortex also responded robustly during a spatial memory task (**Supplemental figure 2-7** and **Supplemental figure 2-8**), similarly to the results reported in Huang et al. (2019). However, the MD system is not specifically tuned to spatial reasoning (Duncan, 2010; Fedorenko et al., 2013; Michalka et al., 2015), so the overlap between code comprehension and spatial reasoning likely reflects the engagement of domain-general cognitive processes, like working memory and cognitive control, as opposed to processes specific to spatial reasoning.

Furthermore, given that no regions outside of the MD system showed code-specific responses, it must be the case that code-specific knowledge representations are also *stored* within this system (see Hasson et al., 2015, for a general discussion of the lack of distinction between storage and computing resources in the brain). Such code-specific representations would likely include both knowledge specific to a programming language (e.g., the syntax marking an array in Java vs. Python) and knowledge of programming concepts that are shared across languages (e.g., *for* loops). Much evidence suggests that the MD system can flexibly store task-relevant information in the short term (e.g., Fedorenko et al., 2013; Freedman et al., 2001; Shashidhara, Mitchell, et al., 2019; Wen et al., 2019; Woolgar et al., 2011). However, evidence from studies on processing mathematics (e.g., Amalric & Dehaene, 2019) and physics (e.g., Cetron et al., 2019; Fischer et al., 2016) further suggests that the MD system can store some domain-specific representations in the long term, perhaps for evolutionarily late-emerging and ontogenetically late-acquired domains of knowledge. Our data add to this body of evidence by showing that the MD system stores and uses information required for code comprehension.

We also show that, instead of being concentrated in one region or a subset of the MD system, code-evoked responses are distributed throughout the MD system. This result

seems to violate general metabolic and computational efficiency principles that govern much of the brain's architecture (Chklovskii & Koulakov, 2004; Kanwisher, 2010): if some MD neurons are, at least in part, functionally specialized to process computer code, we would expect them to be located next to each other. Three possibilities are worth considering. First, selectivity for code comprehension in a subset of the MD network may only emerge with years of experience (e.g., in professional programmers). Participants in our experiments were all proficient in the target programming language but most had only a few years of experience with it. Second, code-selective subsets of the MD network may be detectable at higher spatial resolution, using invasive methods like electrocorticography (Parvizi & Kastner, 2018) or single-cell recordings (Mukamel & Fried, 2012). And third, perhaps the need to flexibly solve novel problems throughout one's life prevents the 'crystallization' of specialized subnetworks within the MD cortex. That said, it may also be the case that some subset of the MD network is causally important for code comprehension even though it does not show strong selectivity for it, similar to how damage to some MD areas (mostly, in the left parietal cortex) appears to lead to deficits in numerical cognition (Ardila & Rosselli, 2002; Kahn & Whitaker, 1991; Lemer et al., 2003; Rosselli & Ardila, 1989; Takayama et al., 1994), even though these regions do not show selectivity for numerical tasks in fMRI (Pinel et al., 2004; Shuman & Kanwisher, 2004).

#### 2.4.2 The language system is functionally conservative

We found that the language system does not respond consistently during code comprehension in spite of numerous similarities between code and natural languages (Fedorenko et al., 2019). Perhaps the most salient similarity between these input types is their syntactic/combinatorial structure. Some accounts of language processing claim that syntactic operations that support language processing are highly abstract and insensitive to the nature of the to-be-combined units (e.g., Berwick et al., 2013; Fitch et al., 2005; Fitch & Martins, 2014; Hauser et al., 2002). Such accounts predict that the mechanisms supporting structure processing in language should also get engaged when we process structure in other domains, including computer code. Prior work has already put into question this idea in its broadest form: processing music, whose hierarchical structure

has long been noted to have parallels with linguistic syntax (e.g., Lerdahl & Jackendoff, 1996; cf. Jackendoff, 2009), does not engage the language system (e.g., Fedorenko, Behr, et al., 2011; Rogalsky et al., 2011; Chen et al, in prep.). Our finding builds upon the results from the music domain to show that compositional input (here, variables and keywords combining into statements) and hierarchical structure (here, conditional statements and loops) do not necessarily engage language-specific regions.

Another similarity shared by computer programming and natural language is the use of symbols — units referring to concepts “out in the world”. Studies of math and logic, domains that also make extensive use of symbols, show that those domains do not, in fact, rely on the language system (Amalric & Dehaene, 2019; Cohen et al., 2000; Fedorenko, Behr, et al., 2011; Monti et al., 2009, 2012; Pinel & Dehaene, 2009; Varley et al., 2005), a conclusion consistent with our findings. However, these findings might be explained by the hypothesis that mathematics makes use of a different conceptual space altogether (Cappelletti et al., 2001), in which case the symbol-referent analogy would be weakened. Our work provides an even stronger test of the symbolic reference hypothesis: the computer code problems we designed are not only symbolic, but also refer to the same conceptual representations as the corresponding verbal problems (**Figure 2-1, A**). This parallel is particularly striking in the case of ScratchJr: each code problem refers to a sequence of actions performed by a cartoon character—a clear case of reference to concepts in the physical world. And yet, the language regions do not respond to ScratchJr, showing a clear preference for language over other types of meaningful structured input (see also Ivanova et al., 2019).

The third similarity between code and natural language is the communicative use of those systems (Allamanis et al., 2018). The programming languages we chose are very high-level, meaning that they emphasize human readability (Buse & Weimer, 2010; Klare, 1963) over computational efficiency. ScratchJr is further optimized to be accessible and engaging for young children (Sullivan & Bers, 2019). Thus, code written in these languages is meant to be read and understood by humans, not just executed by machines. In this respect, computer code comprehension is similar to reading in natural language: the goal is to extract a meaningful message produced by another human at some point in

the past. And yet the communicative nature of this activity is not sufficient to recruit the language system, consistent with previous reports showing a neural dissociation between language and other communication-related activities, such as gesture processing (Jouravlev et al., 2019), intentional actions (Pritchett et al., 2018), or theory of mind tasks (Apperly et al., 2006; Dronkers et al., 1998; Jacoby et al., 2016; Paunov et al., 2019; Varley & Siegal, 2000).

Of course, the lack of consistent language system engagement in code comprehension does not mean that the mechanisms underlying language and code processing are completely different. It is possible that both language and MD regions have similarly organized neural circuits that allow them to process combinatorial input or map between a symbol and the concept it refers to. However, the fact that we observed code-evoked activity primarily in the MD regions indicates that code comprehension does not load on the same neural circuits as language and needs to use domain-general MD circuits instead.

More work is required to determine why the language system showed some activity in response to Python code. The most intuitive explanation posits that the language system responds to meaningful words embedded within the code; however, this explanation seems unlikely given the fact that the responses were equally strong when reading problems with semantically meaningful identifiers (English) and semantically meaningless identifiers (Japanese; **Figure 2-4, C**). Another possibility is that participants internally verbalized the symbols they were reading (where “verbalize” means to retrieve the word associated with a certain symbol rather than a simple reading response, since the latter would be shared with nonwords). However, this account does not explain the fact why such verbalization would be observed for Python and not for ScratchJr, where many blocks have easy labels, such as “jump”. It is also inconsistent with observations that even behaviors that ostensibly require subvocal rehearsal (e.g., mathematical operations) do not engage the language system (see e.g., Amalric & Dehaene, 2019; Fedorenko, Behr, et al., 2011). Finally, the account that we consider most likely is that the responses were driven by processing underlying problem content and thus associated with some aspect(s) of computational thinking that were more robustly present in Python compared to ScratchJr problems. Further investigations of the role of the language system

in computational thinking have the potential to shed light on the exact computations supported by these regions.

Finally, it is possible that the language system may play a role in learning to program (Prat et al., 2020), even if it is not required to support code comprehension once the skill is learned. Studies advocating the ‘coding as another language’ approach (Bers, 2018, 2019; Sullivan & Bers, 2019) have found that treating coding as a meaning-making activity rather than merely a problem-solving skill had a positive impact on both teaching and learning to program in the classroom (Hassenfeld et al., 2020; Hassenfeld & Bers, 2020). Such results indicate that the language system and/or the general semantic system might play a role in learning to process computer code, especially in children, when the language system is still developing. This idea remains to be empirically evaluated in future studies.

#### **2.4.3 Limitations of scope**

The stimuli used in our study were short and only included a few basic elements of control flow (such as *for* loops and *if* statements). Furthermore, we focused on code comprehension, which is a necessary but not sufficient component of many other programming activities, such as code generation, editing, and debugging. Future work should investigate changes in brain activity during the processing and generation of more complex code structures, such as functions, objects, and large multi-component programs. Just like narrative processing recruits systems outside the regions that support single sentence processing (Baldassano et al., 2018; Blank & Fedorenko, 2020; Ferstl et al., 2008; Jacoby & Fedorenko, 2020; Lerner et al., 2011; Simony et al., 2016), reading more complex pieces of code might recruit an extended, or a different, set of brain regions. Furthermore, as noted above, investigations of expert programmers may reveal changes in how programming knowledge and use are instantiated in the mind and brain as a function of increasing amount of domain-relevant experience.

Overall, we provide evidence that code comprehension consistently recruits the MD system—which subserves cognitive processing across multiple cognitive domains—but

does not consistently engage the language system, in spite of numerous similarities between natural and programming languages. By isolating neural activity specific to code comprehension, we pave the way for future studies examining the cognitive and neural correlates of programming and contribute to the broader literature on the neural systems that support novel cognitive tools.

## 2.5 Method

### 2.5.1 Participants

For Experiment 1, we recruited 25 participants (15 women, mean age = 23.0 years, SD = 3.0). Average age at which participants started to program was 16 years (SD = 2.6); average number of years spent programming was 6.3 (SD=3.8). In addition to Python, 20 people also reported some knowledge of Java, 18 people reported knowledge of C/C++, 4 of functional languages, and 20 of numerical languages like Matlab and R. Twenty-three participants were right-handed, one was ambidextrous, and one was left-handed (as assessed by Oldfield's (1971) handedness questionnaire); the left-handed participant had a right-lateralized language system and was excluded from the analyses, leaving 24 participants (all of whom had left-lateralized language regions, as evaluated with the language localizer task; see below). Participants also reported their knowledge of foreign languages and completed a one-hour-long Python proficiency test (available on the paper's website, <https://github.com/ALFA-group/neural-program-comprehension>).

For Experiment 2, we recruited 21 participants (13 women, mean age = 22.5 years, SD = 2.8). In addition to ScratchJr, 8 people also reported some knowledge of Python, 6 people reported knowledge of Java, 9 people reported knowledge of C/C++, 1 of functional languages, and 14 of numerical languages like Matlab and R (one participant did not complete the programming questionnaire). Twenty were right-handed and one was ambidextrous; all participants had left-lateralized language regions, as evaluated with the language localizer task (see below). Two participants from Experiment 2 had to be excluded due to excessive motion during the MRI scan, leaving 19 participants.

All participants were recruited from MIT, Tufts University, and the surrounding community and paid for participation. All were native speakers of English, had normal or corrected to normal vision, and reported working knowledge of Python or ScratchJr, respectively. The sample size for both experiments was determined based on previous experiments from our group (e.g., Blank & Fedorenko, 2020; Fedorenko et al., 2020b; Ivanova et al., 2019) and others (e.g., Crittenden et al., 2015; Hugdahl et al., 2015; Shashidhara, Mitchell, et al., 2019). The protocol for the study was approved by MIT's Committee on the Use of Humans as Experimental Subjects (COUHES). All participants gave written informed consent in accordance with protocol requirements.

## 2.5.2 Design, materials, and procedure

All participants completed the main program comprehension task, a spatial working memory localizer task aimed at identifying the multiple demand (MD) brain regions (Fedorenko, Behr, et al., 2011), and a language localizer task aimed at identifying language-responsive brain regions (Fedorenko et al., 2010).

### 2.5.2.1 Program comprehension tasks

The **program comprehension task in Experiment 1** included three conditions: programs in Python with English identifiers, programs in Python with Japanese identifiers, and sentence versions of those programs (visually presented). The full list of problems can be found on the paper's website, <https://github.com/ALFA-group/neural-program-comprehension>. Each participant saw 72 problems, and any given participant saw only one version of a problem. Half of the problems required performing mathematical operations, and the other half required string manipulations. In addition, both math and string-manipulation problems varied in program structure: 1/3 of the problems of each type included only sequential statements, 1/3 included a *for* loop, and 1/3 included an *if* statement.

During each trial, participants were instructed to read the problem statement and press a button when they were ready to respond (the minimum processing time was restricted to 5 s and the maximum to 50 s; mean reading time was 19 s). Once they pressed the button, four response options were revealed, and participants had to indicate their response by pressing one of four buttons on a button box. The response screen was presented for 5 s (see **Supplemental figure 2-1, A**, for a schematic of trial structure). Each run consisted of 6 trials (2 per condition), and 3 fixation blocks (at the beginning and end of the run, and after the third trial), each lasting 10 s. A run lasted, on average, 176 s (SD = 34 s), and each participant completed 12 runs. Condition order was counterbalanced across runs and participants.

The **program comprehension task in Experiment 2** included two conditions: short programs in ScratchJr and the sentence versions of those programs (visually presented). ScratchJr is a language designed to teach programming concepts to young children (Bers, 2018): users can create events and sequences of events (stories) with a set of characters and actions. The full list of problems used in the study can be found on the paper's website. Each participant saw 24 problems, and any given participant saw only one version of a problem. Furthermore, problems varied in the complexity of the code snippet (3 levels of difficulty; 8 problems at each level).

During each trial, participants were presented with a fixation cross for 4 s, followed by a description (either a code snippet or a sentence) to read for 8 s. The presentation of the description was followed by 5-9 s of fixation, and then by a video (average duration: 4.13 s, SD: 1.70 s) that either did or did not match the description. Participants had to indicate whether the video matched the description by pressing one of two buttons on a button box in the scanner. The response window started with the onset of the video and included a 4 s period after the video offset. A trial lasted, on average, 27.46 s (SD = 2.54 s; see **Supplemental figure 2-1, B**, for a schematic of trial structure). Each run consisted of 6 trials (3 per condition), and a 10 s fixation at the beginning and end of the run. A run lasted, on average, 184.75 s (SD = 3.86 s); each participant completed 4 runs. Condition order was counterbalanced across runs and participants.

### 2.5.2.2 Multiple demand localizer

To localize MD system within individual participants, we conducted a **spatial working memory task**. Participants had to keep track of four (easy condition) or eight (hard condition) sequentially presented locations in a  $3 \times 4$  grid (**Figure 2-1**, B; Fedorenko et al., 2011). In both conditions, they performed a two-alternative forced-choice task at the end of each trial to indicate the set of locations they just saw. The hard > easy contrast has been previously shown to reliably activate bilateral frontal and parietal MD regions (Assem, Glasser, et al., 2020; Blank et al., 2014b; Fedorenko et al., 2013). Numerous studies have shown that the same brain regions are activated by diverse executively-demanding tasks (Duncan & Owen, 2000; Fedorenko et al., 2013; Hugdahl et al., 2015; Shashidhara, Mitchell, et al., 2019; Woolgar et al., 2011). Stimuli were presented in the center of the screen across four steps. Each step lasted 1 s and revealed one location on the grid in the easy condition, and two locations in the hard condition. Each stimulus was followed by a choice-selection step, which showed two grids side by side. One grid contained the locations shown across the previous four steps, while the other contained an incorrect set of locations. Participants were asked to press one of two buttons to choose the grid that showed the correct locations. Condition order was counterbalanced across runs. Experimental blocks lasted 32 s (with 4 trials per block), and fixation blocks lasted 16 s. Each run (consisting of 4 fixation blocks and 12 experimental blocks) lasted 448 s. Each participant completed 2 runs.

### 2.5.2.3 Language localizer

The **language localizer task** was conducted in order to identify the language system within individual participants. Participants read sentences (e.g., NOBODY COULD HAVE PREDICTED THE EARTHQUAKE IN THIS PART OF THE COUNTRY) and lists of unconnected, pronounceable nonwords (e.g., U BIZBY ACWORRILY MIDARAL MAPE LAS POME U TRINT WEPS WIBRON PUZ) in a blocked design. Each stimulus consisted of twelve words/nonwords. For details of how the language materials were constructed, see Fedorenko et al. (2010). The materials are available at [http://web.mit.edu/evelina9/www/funcloc/funcloc\\_localizers.html](http://web.mit.edu/evelina9/www/funcloc/funcloc_localizers.html). The sentences >

nonword-lists contrast isolates processes related to language comprehension (responses evoked by, e.g., visual perception and reading are subtracted out) and has been previously shown to reliably activate left-lateralized fronto-temporal language processing regions, be robust to changes in task and materials, and activate the same regions regardless of whether the materials were presented visually or auditorily (Fedorenko et al., 2010; Mahowald & Fedorenko, 2016; Scott et al., 2017). Further, a similar network emerges from task-free resting-state data (Braga et al., 2020). Stimuli were presented in the center of the screen, one word/nonword at a time, at the rate of 450 ms per word/nonword. Each stimulus was preceded by a 100 ms blank screen and followed by a 400 ms screen showing a picture of a finger pressing a button, and a blank screen for another 100 ms, for a total trial duration of 6 s. Participants were asked to press a button whenever they saw the picture of a finger pressing a button. This task was included to help participants stay alert. Condition order was counterbalanced across runs. Experimental blocks lasted 18 s (with 3 trials per block), and fixation blocks lasted 14 s. Each run (consisting of 5 fixation blocks and 16 experimental blocks) lasted 358 s. Each participant completed 2 runs.

### **2.5.3 fMRI data acquisition**

Structural and functional data were collected on the whole-body, 3 Tesla, Siemens Trio scanner with a 32-channel head coil, at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT. T1-weighted structural images were collected in 176 sagittal slices with 1mm isotropic voxels (TR = 2,530 ms, TE = 3.48 ms). Functional, blood oxygenation level dependent (BOLD), data were acquired using an EPI sequence (with a 90° flip angle and using GRAPPA with an acceleration factor of 2), with the following acquisition parameters: thirty-one 4mm thick near-axial slices acquired in the interleaved order (with 10% distance factor), 2.1 mm × 2.1 mm in-plane resolution, FoV in the phase encoding (A>>P) direction 200mm and matrix size 96 mm × 96 mm, TR = 2,000 ms and TE = 30 ms. The first 10 s of each run were excluded to allow for steady state magnetization.

#### **2.5.4 fMRI data preprocessing**

fMRI data were analyzed using SPM12 (release 7487), CONN EvLab module (release 19b), and other custom MATLAB scripts. Each participant's functional and structural data were converted from DICOM to NIFTI format. All functional scans were coregistered and resampled using B-spline interpolation to the first scan of the first session (Friston et al., 1995). Potential outlier scans were identified from the resulting subject-motion estimates, as well as from BOLD signal indicators, using default thresholds in CONN preprocessing pipeline (5 standard deviations above the mean in global BOLD signal change, or framewise displacement values above 0.9 mm; Nieto-Castañon, 2020). Functional and structural data were independently normalized into a common space (the Montreal Neurological Institute [MNI] template; IXI549Space) using SPM12 unified segmentation and normalization procedure (Ashburner & Friston, 2005) with a reference functional image computed as the mean functional data after realignment across all timepoints omitting outlier scans. The output data were resampled to a common bounding box between MNI-space coordinates (-90, -126, -72) and (90, 90, 108), using 2mm isotropic voxels and 4th order spline interpolation for the functional data, and 1mm isotropic voxels and trilinear interpolation for the structural data. Last, the functional data were smoothed spatially using spatial convolution with a 4 mm FWHM Gaussian kernel.

#### **2.5.5 First-level analysis**

Responses in individual voxels were estimated using a General Linear Model (GLM) in which each experimental condition was modeled with a boxcar function convolved with the canonical hemodynamic response function (HRF) (fixation was modeled implicitly, such that all timepoints that did not correspond to one of the conditions were assumed to correspond to a fixation period). Temporal autocorrelations in the BOLD signal timeseries were accounted for by a combination of high-pass filtering with a 128 seconds cutoff and whitening using an AR(0.2) model (first-order autoregressive model linearized around the coefficient  $a=0.2$ ) to approximate the observed covariance of the functional data in the context of Restricted Maximum Likelihood estimation (ReML). In addition to

experimental condition effects, the GLM design included first-order temporal derivatives for each condition (included to model variability in the HRF delays), as well as nuisance regressors to control for the effect of slow linear drifts, subject-motion parameters, and potential outlier scans on the BOLD signal.

For the localizer experiments, we modeled the entire blocks. For the Python program comprehension experiment, we modeled the period from the onset of the code/sentence problem and until the button press (the responses were modeled as a separate condition; see **Supplemental figure 2-1, A**); for the ScratchJr program comprehension experiment, we modeled the period of the code/sentence presentation (the video and the response were modeled as a separate condition; see **Supplemental figure 2-1, B**).

### 2.5.6 Defining MD and language functional regions of interest (fROIs)

The fROI analyses examined responses in individually defined MD and language fROIs (functional regions of interest). These fROIs were defined using the Group-constrained Subject-Specific (GSS) approach (Fedorenko et al. 2010; Julian et al. 2012) where a set of spatial masks, or parcels, is combined with each individual subject's localizer activation map, to constrain the definition of individual fROIs. The parcels delineate the expected gross locations of activations for a given contrast based on prior work and large numbers of participants and are sufficiently large to encompass the variability in the locations of individual activations. For the MD system, we used a set of 20 parcels (10 in each hemisphere) derived from a group-level probabilistic activation overlap map for the hard > easy spatial working memory contrast in 197 participants. The parcels included regions in frontal and parietal lobes, as well as a region in the anterior cingulate cortex. For the language system, we used a set of six parcels derived from a group-level probabilistic activation overlap map for the sentences > nonwords contrast in 220 participants. The parcels included two regions in the left inferior frontal gyrus (LIFG, LIFGorb), one in the left middle frontal gyrus (LMFG), two in the left temporal lobe (LAntTemp and LPostTemp), and one extending into the angular gyrus (LAngG). Both sets of parcels are available on the paper's website; see **Supplemental figure 2-5** for labeled images of MD and language parcels.

To create each set of parcels, individual activation maps for the relevant localizer contrast were binarized (by turning all voxels significant at the  $p < .001$  whole-brain threshold (uncorrected) into 1s, and the rest into 0s) and overlaid in the MNI space to create a probabilistic overlap map. For the multiple demand network, the individual activation maps were averaged across the two hemispheres prior to binarizing. The map was then smoothed ( $\text{FWHM} = 6\text{mm}$ ), and voxels with fewer than 10% of participants overlapping were excluded. The resulting map was divided into regions using a watershed algorithm. Finally, we excluded parcels that did not show significant effects for the relevant localizer contrast in a left-out run or did not contain supra-threshold voxels in at least 60% of the participants. For the multiple demand network, we also a) excluded parcels in the visual cortex (the hard condition includes more visual information than the easy condition and thus yields more activation in the visual cortex), and b) divided a parcel that encompassed parts of both the precentral gyrus and the opercular portion of the inferior frontal gyrus according to the macroanatomical boundary.

For each participant, each set of parcels was intersected with the participant's activation map for the relevant contrast (sentences>nonwords for the language network, hard>easy spatial WM for the multiple demand network). Within each parcel, the voxels were sorted based on their  $t$ -values for the relevant contrast, and the top 10% of voxels were selected as that participant's fROI (see Figure 1-1 for sample language fROIs). This top n% approach ensures that the fROIs can be defined in every participant, thus enabling us to have stable units of comparison across individuals (Nieto-Castañón & Fedorenko, 2012).

### 2.5.7 Examining the functional response profiles of MD and language fROIs

**Univariate analyses.** We evaluated MD and language system responses by estimating their response magnitudes to the conditions of interest using individually defined fROIs (see above). For each fROI in each participant, we averaged the responses across voxels to get a single value for each of the conditions (the responses to the localizer conditions were estimated using an across-runs cross-validation procedure, where one run was used to define the fROI and the other to estimate the response magnitudes, then the procedure

was repeated switching which run was used for fROI definition vs. response estimation, and finally the estimates were averaged to derive a single value per condition per fROI per participant). We then ran a linear mixed-effect regression model to compare the responses to the critical code problem condition with (a) the responses to the sentence problem condition from the critical task, and (b) the responses to the nonword reading condition from the language localizer task. We included *condition* as a fixed effect and *participant* and *fROI* as random intercepts. For the MD system, we additionally tested the main (fixed) effect of *hemisphere* and the interaction between *hemisphere* and *condition*. We used dummy coding for *condition*, with code problems as the reference category, and sum coding for *hemisphere*. For follow-up analyses, we used the variable of interest (*problem type/structure/identifier language*) as a fixed effect and *participant* and *fROI* as random intercepts; dummy coding was used for all variables of interest. For fROI analyses, we used *condition* as a fixed effect and *participant* as a random intercept. The analyses were run using the *lmer* function from the *lme4* R package (Bates et al., 2015); statistical significance of the effects was evaluated using the *lmerTest* package (Kuznetsova et al., 2017).

**Spatial correlation analyses.** To further examine the similarity of the fine-grained patterns of activation between conditions in the language system, we calculated voxel-wise spatial correlations in activation magnitudes within the code problem condition (between odd and even runs), within the sentence problem condition (between odd and even runs), between these two conditions (we used odd and even run splits here, too, to match the amount of data for the within- vs. between-condition comparisons, and averaged the correlation values across the different splits), and between these two critical conditions and each of the sentence and nonword reading conditions from the language localizer. The correlation values were calculated for voxels in each participant's language fROIs, and then averaged across participants and fROIs for plotting (the values were weighted by fROI size). We also used the *lme4* R package to calculate statistical differences between spatial correlation values for code vs. other conditions (with *participant* and *fROI* as random intercepts); for this analysis, the correlation values were Fischer-transformed.

### **2.5.8 Whole-brain analyses**

For each of the critical experiments (Python and ScratchJr), we conducted (a) the Group-constrained Subject-Specific (GSS) analysis (Fedorenko et al., 2010; Julian et al., 2012), and (b) the traditional random effects group analysis (A. P. Holmes & Friston, 1998) using the code problems > sentence problems contrast. The analyses were performed using the spm\_ss toolbox ([http://www.nitrc.org/projects/spm\\_ss](http://www.nitrc.org/projects/spm_ss)), which interfaces with SPM and the CONN toolbox (<https://www.nitrc.org/projects/conn>).

## **2.6 Acknowledgements**

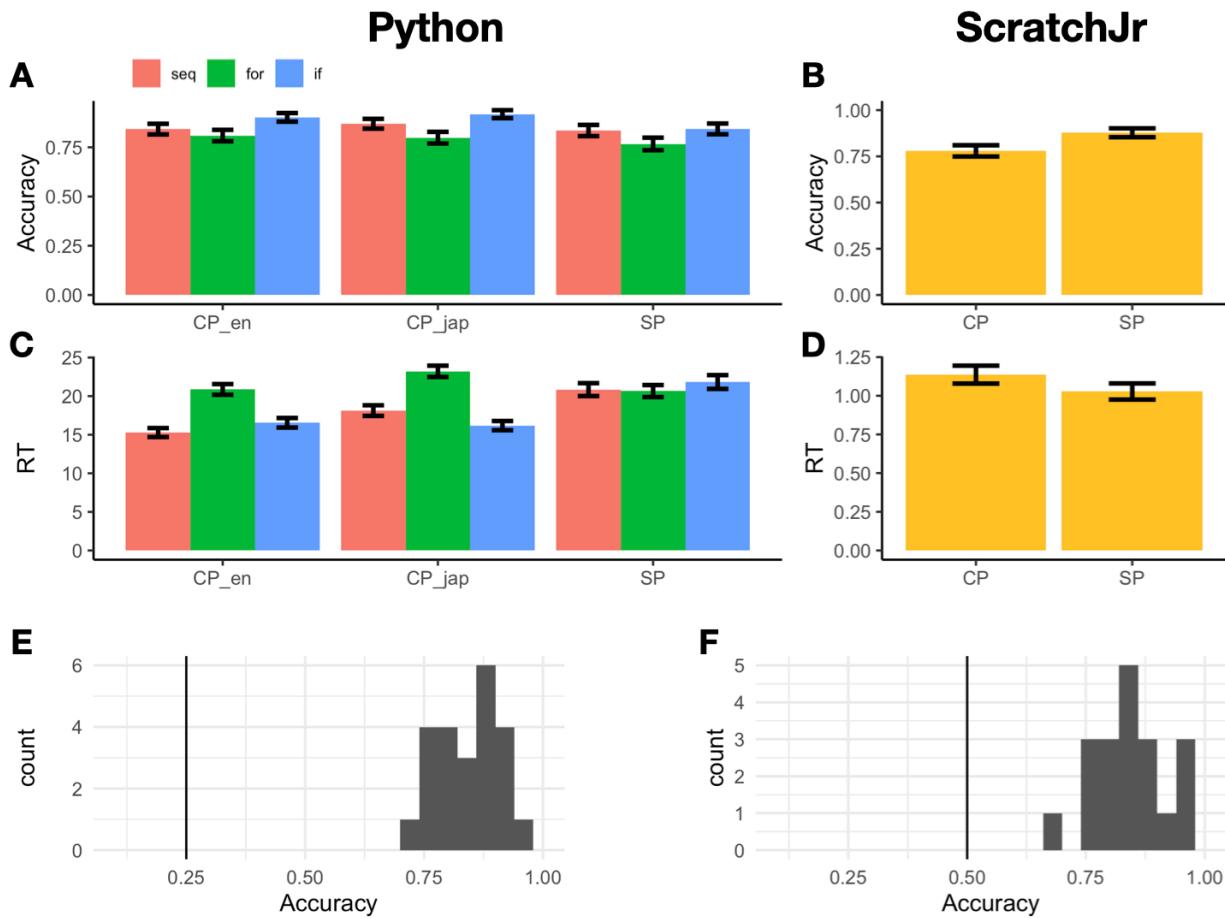
We would like to acknowledge the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT and its support team (Steve Shannon, Atsushi Takahashi, and Dima Ayyash), Rachel Ryskin for advice on statistics, Alfonso Nieto-Castañón for help with analyses, ALFA group at CSAIL for helpful discussions on the experiment design, and Josef Affourtit, Yev Diachek, and Matt Siegelman (EvLab), and Ruthi Aladjem, Claudia Mihm, and Kaitlyn Leidl (DevTech research group at Tufts University) for technical support during experiment design and data collection.

## 2.7 Appendix

### 2.7.1 Supplemental figures

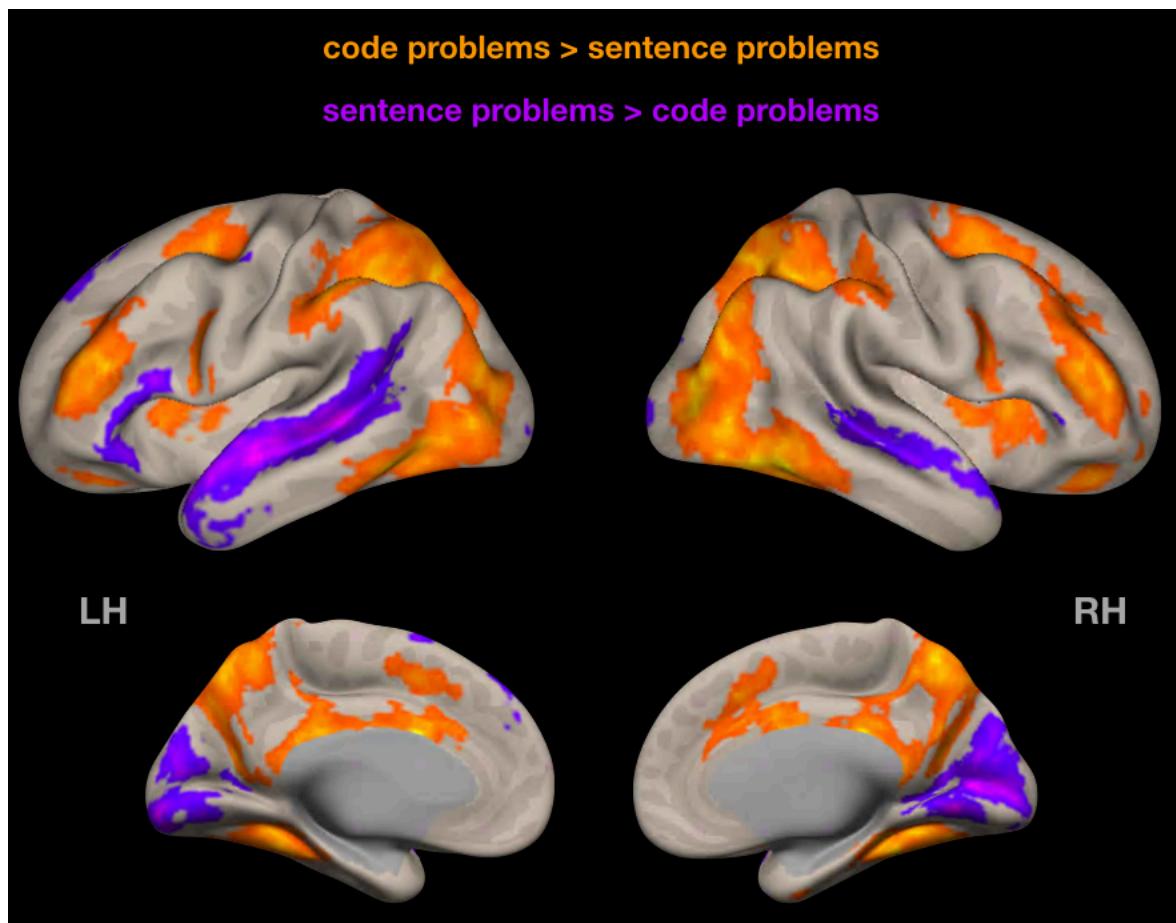
A	Problem	Response
	<pre>height = 5 weight = 100 bmi = weight/(height*height) print(bmi)</pre>	A. 5 B. 6 C. 8 D. Other
	Your height is 5 feet and your weight is 100 pounds. The BMI is defined as the ratio between the weight and the square of the height of a person. What is your BMI?	A. 5 B. 6 C. 8 D. Other
self-paced (5 s min, 50 s max)		5 s
B	Problem	Video
		
	Kitten walks right, jumps, and then walks left.	
8 s		avg 4.13 s
Response		Did the instructions match the video? A. yes      B. no
		Did the instructions match the video? A. yes      B. no
		4 s

**Supplemental figure 2-1.** Trial structure of the critical task. (A) Experiment 1 - Python (B) Experiment 2 - ScratchJr. All analyses use fMRI responses to the “problem” step.

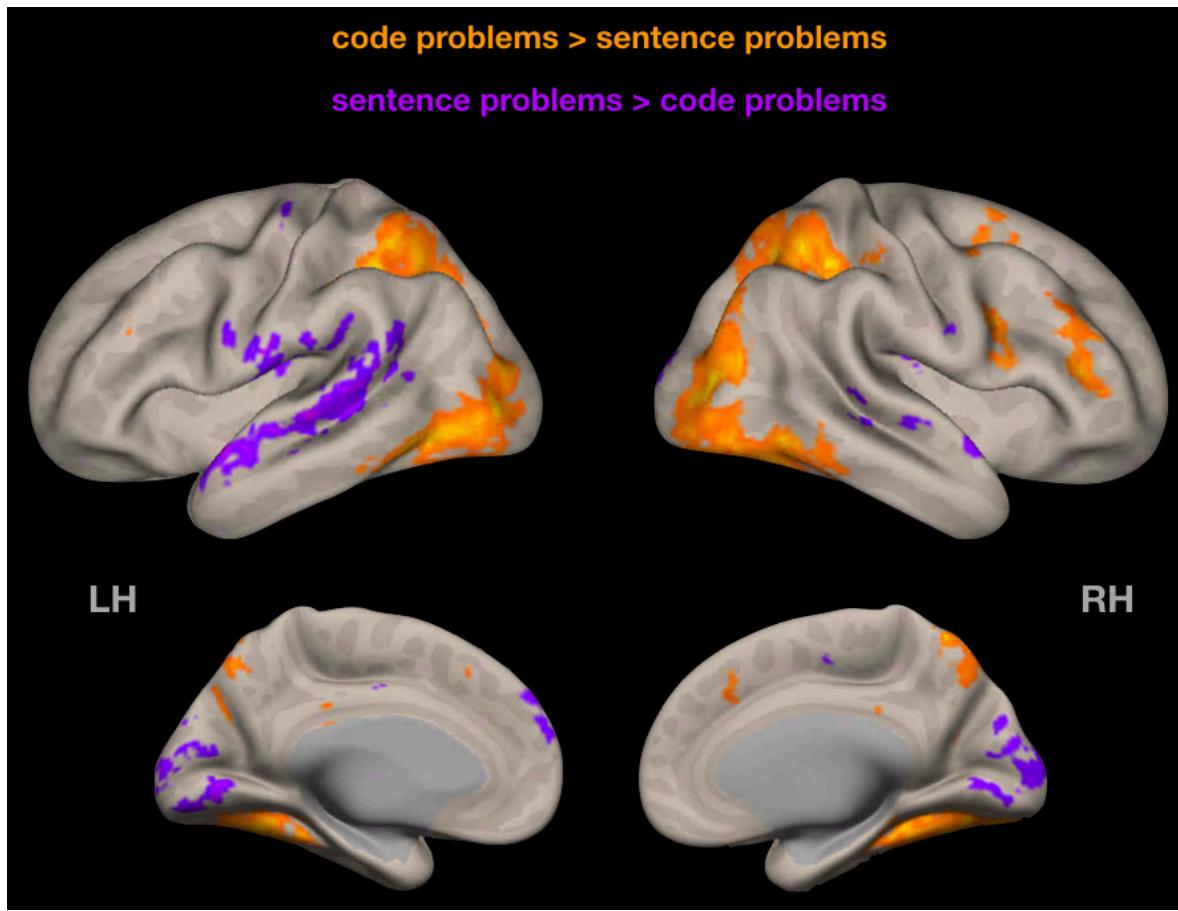


**Supplemental figure 2-2.** Behavioral results. (A) Python code problems had mean accuracies of 85.1% and 86.2% for the English-identifier (CP\_en) and Japanese-identifier (CP\_jap) conditions, respectively, and sentence problems (SP) had a mean accuracy of 81.5%. There was no main effect of condition (CP\_en, CP\_jap, SP), problem structure (seq – sequential, for – for loops, if – if statements), or problem content (math vs. string); however, there was a three-way interaction among Condition (sentence problems > code with English identifiers), Problem Type (string > math), and Problem Structure (for loop > sequential;  $p = 0.02$ ). Accuracy data from one participant had to be excluded due to a bug in the script. (B) ScratchJr code problems had a mean accuracy of 78.0%, and sentence problems had a mean accuracy of 87.8% (the difference was significant:  $p = 0.006$ ). (C) Python problems with English identifiers had a mean response time (RT) of 17.56 s ( $SD = 9.05$ ), Python problems with Japanese identifiers had a mean RT of 19.39 s ( $SD = 10.1$ ), and sentence problems had a mean RT of 21.32 s ( $SD = 11.6$ ). Problems with Japanese identifiers took longer to answer than problems with English identifiers ( $\beta = 3.10$ ,  $p = 0.002$ ), and so did sentence problems ( $\beta = 6.12$ ,  $p < 0.001$ ). There was also an interaction between Condition (sentence problems > code with English identifiers) and Program Structure (for > seq;  $\beta = -5.25$ ,  $p < 0.001$ ), as well as between Condition (CP\_jap > CP\_en) and Program Structure (if > seq;  $\beta = -2.83$ ,  $p = 0.04$ ). There was no significant difference in response times between math and string manipulation problems. (D) ScratchJr code problems had a mean RT of 1.14 s ( $SD = 0.86$ ), and sentence problems had a mean RT of 1.03 s ( $SD = 0.78$ ; the difference was not significant). The RTs are reported with respect to

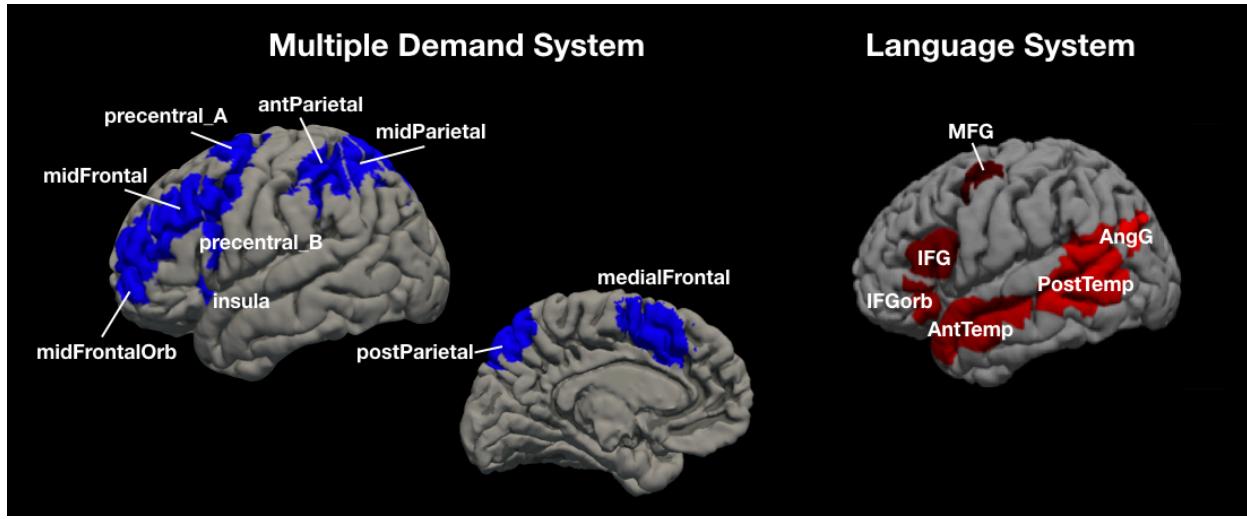
video offset. Items where >50% participants chose the incorrect answer for the (easy) verbal condition were excluded from accuracy calculations. (E) Mean accuracies for all Python participants were above chance. (F) Mean accuracies for all ScratchJr participants were above chance.



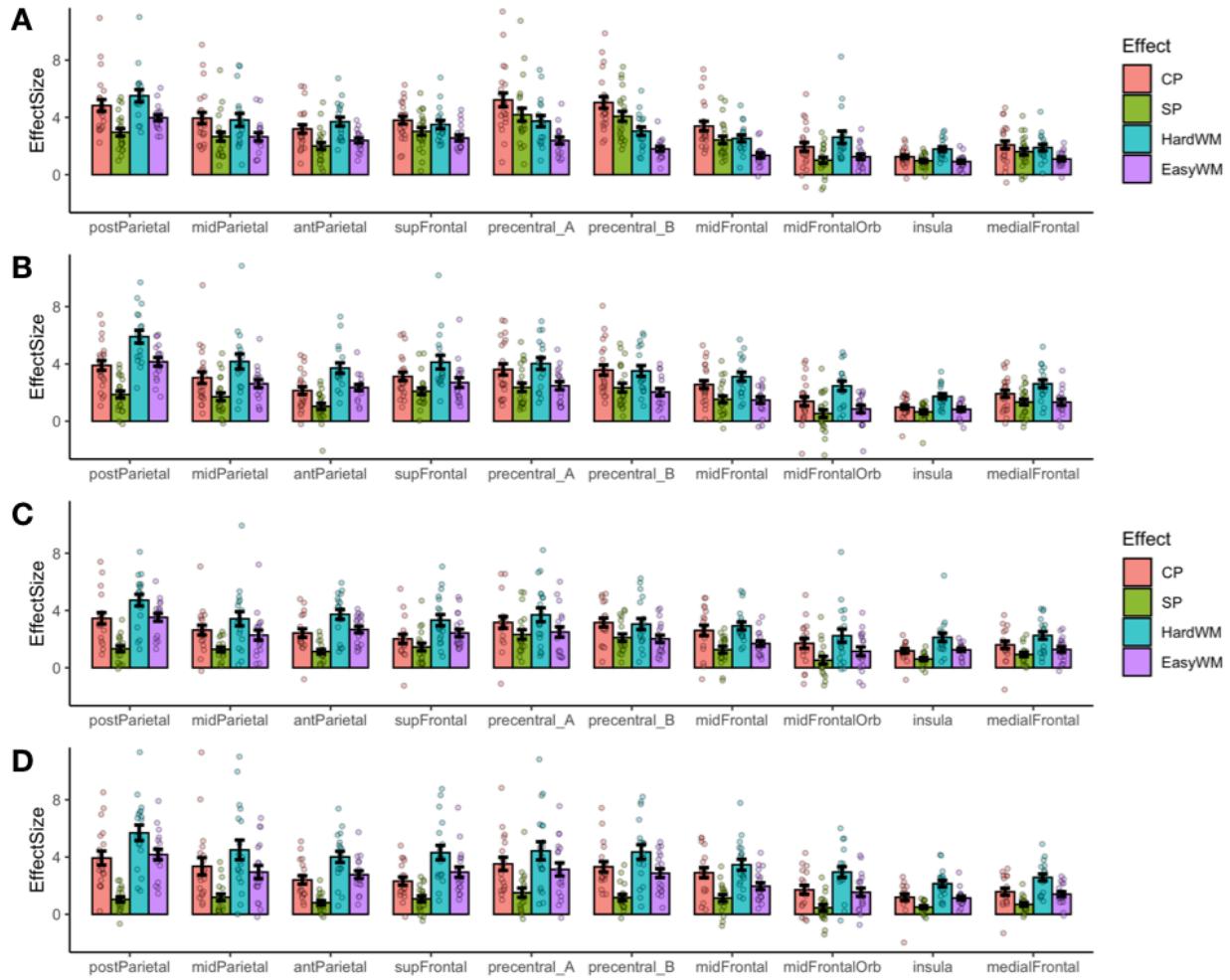
**Supplemental figure 2-3.** Random-effects group-level analysis of Experiment 1 data (Python, code problems > sentence problems contrast). Similarly to analyses reported in the main text, code-evoked activity is bilateral and recruits fronto-parietal but not temporal regions. Cluster threshold  $p < 0.05$ , cluster-size FDR-corrected; voxel threshold:  $p < 0.001$ , uncorrected.



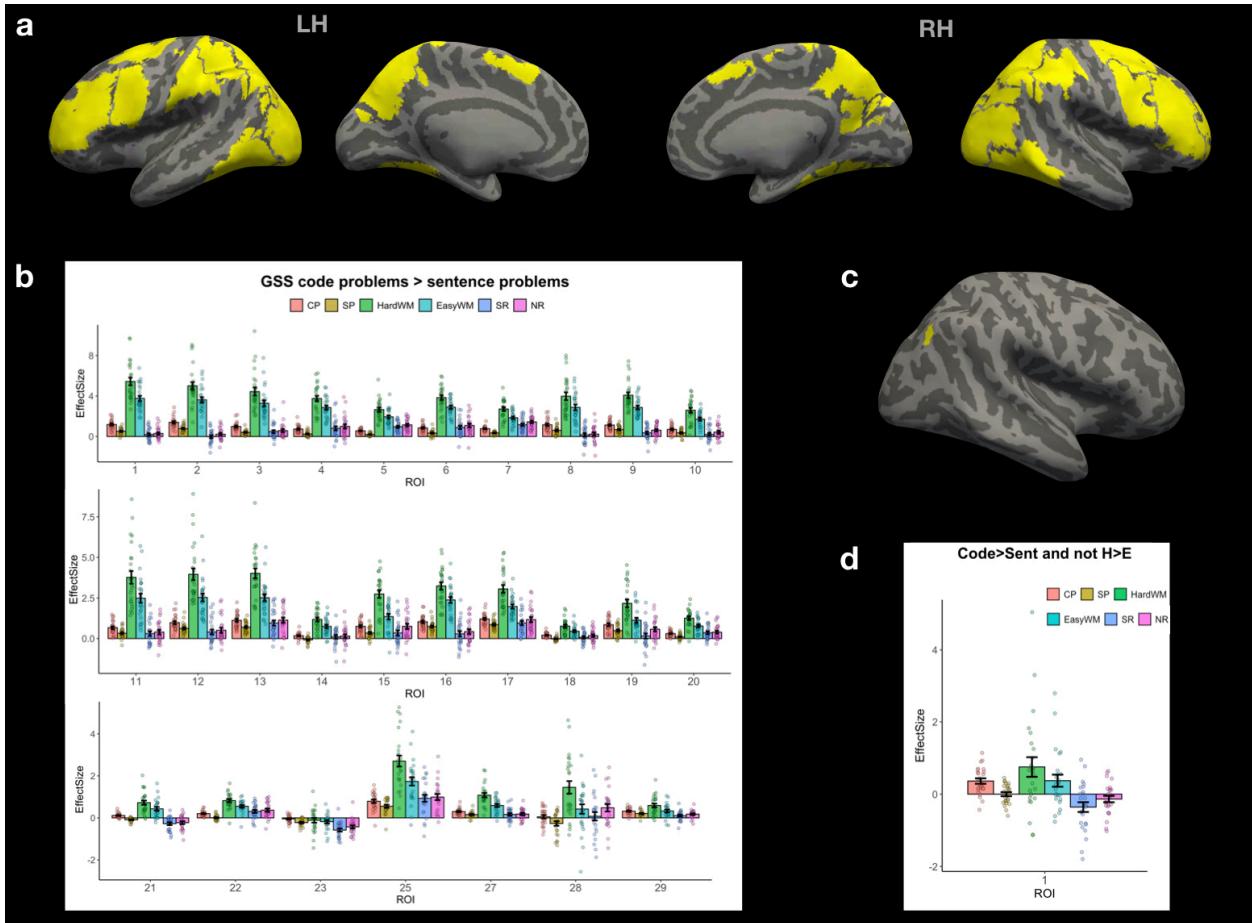
**Supplemental figure 2-4.** Random-effects group-level analysis of Experiment 2 data (ScratchJr, code problems > sentence problems contrast). Similarly to analyses reported in the main text, ScratchJr-evoked activity has a small right hemisphere bias. Cluster threshold  $p < 0.05$ , cluster-size FDR-corrected; voxel threshold:  $p < 0.001$ , uncorrected.



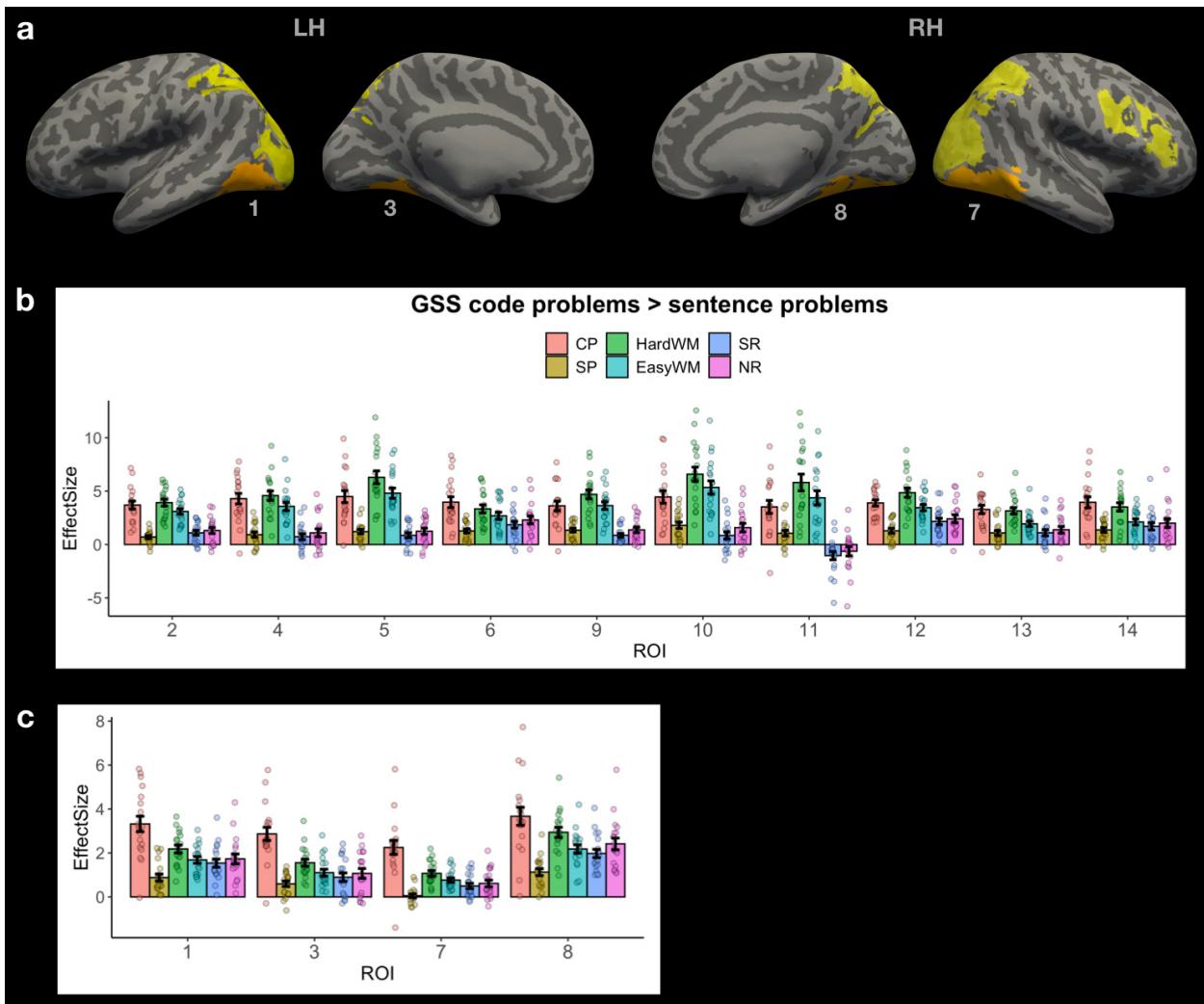
**Supplemental figure 2-5.** The parcels in the two candidate brain systems of interest, multiple demand (MD) and language. The parcels are derived from group-level representations of MD and language activity and are used to define the functional regions of interest (fROIs) in individual participants (NB: we show the left hemisphere parcels for the MD system, but the system is bilateral). For each participant, the network of interest is comprised of the top 10% of voxels within each parcel with the highest t-value for the relevant contrast (MD - hard vs. easy spatial working memory task; language – sentence reading vs. nonword reading; see **Method**). Abbreviations: mid – middle, ant – anterior, post – posterior, orb – orbital, MFG – middle frontal gyrus, IFG – inferior frontal gyrus, temp – temporal lobe, AngG – angular gyrus.



**Supplemental figure 2-6.** ROI-level responses in the multiple demand system to the critical task (CP – code problems, SP – sentence problems) and the spatial working memory task (HardWM – hard working memory task, EasyWM – easy working memory task). (A) Experiment 1, Python; left hemisphere fROIs; (B) Experiment 1, Python; right hemisphere fROIs; (C) Experiment 2, ScratchJr; left hemisphere fROIs; (D) Experiment 2, ScratchJr; right hemisphere fROIs. No fROIs prefer both Python and ScratchJr code problems over the spatial working memory task.



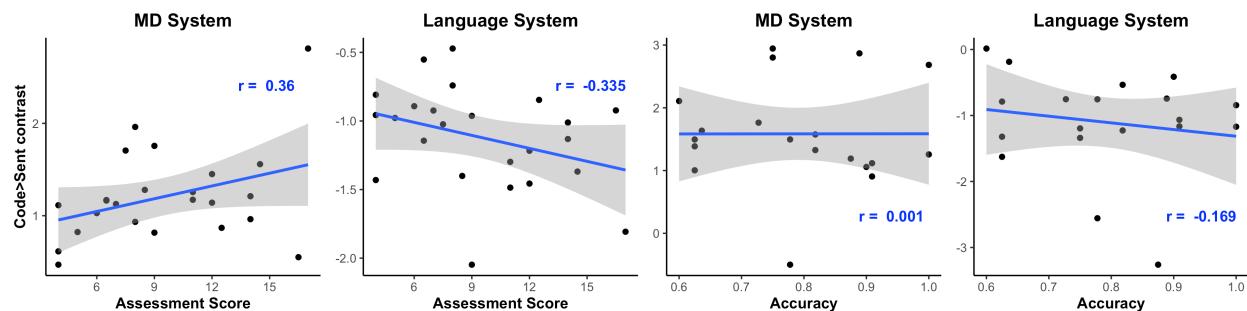
**Supplemental figure 2-7.** Whole-brain group-constrained subject-specific analysis (GSS; Fedorenko et al, 2010) based on data from Experiment 1 shows the absence of code-only brain regions. (a) Parcels defined on the group level using the code problems > sentence problems contrast,  $p$  threshold 0.001, inter-subject overlap  $\geq 70\%$ . (b) Activation profile for the top 10% of voxels within each parcel in (a) across conditions. All code-sensitive regions exhibit high activity during the spatial working-memory task, suggesting that they belong to the MD system. (c) Parcels defined using the contrast above plus the “not hard working-memory task > easy working-memory task” contrast,  $p=0.5$ . Only one parcel was significant (right hemisphere). (d) Even that parcel’s response profile shows high activity in response to the working-memory task, modulated by difficulty, rather than a code-specific response. Abbreviations: CP – code problems; SP – sentence problems; HardWM – hard working memory task; EasyWM – easy working memory task; SR – sentence reading; NR – nonword reading.



**Supplemental figure 2-8.** Whole-brain group-constrained subject-specific analysis (GSS; Fedorenko et al, 2010) based on data from Experiment 2. (a) Parcels defined on the group level using the code problems > sentence problems contrast,  $p$  threshold 0.001, inter-subject overlap  $\geq 70\%$ . Parcels where the responses to ScratchJr code were stronger than responses to all other tasks are labeled and marked in orange; they include parts of early visual cortex and parts of the ventral visual stream. (b) Activation profile for the top 10% of voxels within each parcel in (a) marked in yellow. All regions exhibit high activity during the spatial working-memory task, suggesting that they belong to the MD system. (c) Activation profile for the top 10% of voxels within each parcel in (a) marked in orange. These fROIs exhibit higher responses to ScratchJr problems compared to a working memory task; given that they are located in the visual cortex, we can infer that they respond to low-level visual properties of ScratchJr code. A follow-up conjunction analysis using the contrast in (a) plus the “not hard working-memory task > easy working-memory task” contrast,  $p=0.5$ , revealed no significant parcels, indicating the lack of code-selective response. Abbreviations: CP – code problems; SP – sentence problems; HardWM – hard working memory task; EasyWM – easy working memory task; SR – sentence reading; NR – nonword reading.

MD System Left Hemisphere				MD System Right Hemisphere					
CP	SP	SR	NR	CP	SP	SR	NR		
0.87	0.82	0.51	0.55	CP	0.83	0.75	0.5	0.51	CP
0.82	0.82	0.57	0.57	SP	0.75	0.75	0.55	0.52	SP
0.51	0.57	0.69	0.59	SR	0.5	0.55	0.62	0.55	SR
0.55	0.57	0.59	0.61	NR	0.51	0.52	0.55	0.59	NR

**Supplemental figure 2-9.** Spatial correlation analysis of voxel responses within the MD system during the Python experiment (CP – code problems and SP – sentence problems) with the language localizer conditions for the same participants (SR – sentence reading and NR – nonword reading). Each cell shows a correlation between voxel-level activation patterns for each condition. Within-condition similarity is estimated by correlating activation patterns across independent runs. Code problems correlate with sentence problems much more strongly than with sentence reading ( $\beta = -0.59$ ,  $p < 0.001$ ) and with nonword reading ( $\beta = -0.55$ ,  $p < 0.001$ ), but substantially weaker than with other code problems ( $\beta = 0.11$ ,  $p < 0.001$ ). There was no main effect of hemisphere, but there was an interaction between code/sentence problems and reading conditions (sentence reading:  $\beta = 0.17$ ,  $p < 0.001$ , nonword reading:  $\beta = 0.13$ ,  $p = 0.002$ ), indicating that the correlation patterns of code/sentence problems were somewhat less robust in the right hemisphere.



**Supplemental figure 2-10.** The effect of programming expertise on code-specific response strength within the MD and language system in Experiment 1, Python (A, B) and Experiment 2, ScratchJr (C, D). Python expertise was evaluated with a separate one-hour-long Python assessment (see the paper's website <https://github.com/ALFA-group/neural-program-comprehension>); ScratchJr expertise was estimated with in-scanner response accuracies. No correlations were significant.

## 2.7.2 Supplemental tables

**Supplemental table 2-1.** Responses to Python code problems (CP) vs. sentence problems (SP) and nonword reading (NR) in individual fROIs within the multiple demand system. P values are Bonferroni-corrected for the number of regions. Non-significant values are italicized and marked in gray.

Hemisphere	fROI	Regression Term	Beta	p value
L	postParietal	Intercept	4.79	2.00E-20
		CP>SP	1.86	3.00E-07
		CP>NR	4.17	1.05E-18
L	midFrontal	Intercept	3.39	1.56E-16
		CP>SP	0.97	0.001
		CP>NR	2.5	9.62E-14
L	precentral_B	Intercept	4.91	1.23E-14
		CP>SP	0.95	0.002
		CP>NR	3.38	1.68E-18
L	antParietal	Intercept	3.13	4.40E-15
		CP>SP	1.17	1.07E-06
		CP>NR	2.54	3.11E-17
L	midFrontalOrb	Intercept	1.92	4.73E-09
		CP>SP	0.91	0.004
		CP>NR	1.01	8.90E-04
L	medialFrontal	Intercept	2.15	2.48E-11
		CP>SP	0.49	0.16
		CP>NR	1.41	4.49E-09
L	midParietal	Intercept	3.89	2.84E-14
		CP>SP	1.28	2.52E-05
		CP>NR	2.95	1.11E-15
L	precentral_A	Intercept	5.16	1.27E-12
		CP>SP	1.03	9.55E-04
		CP>NR	3.32	1.13E-17
L	supFrontal	Intercept	3.8	5.83E-18
		CP>SP	0.79	0.002
		CP>NR	2.96	1.14E-19
L	insula	Intercept	1.26	6.46E-16
		CP>SP	0.28	0.141
		CP>NR	0.66	6.33E-07
R	postParietal	Intercept	3.87	4.58E-18
		CP>SP	2	1.09E-10
		CP>NR	3.51	1.68E-19
R	midFrontal	Intercept	2.65	2.66E-14
		CP>SP	1.06	2.62E-04
		CP>NR	1.81	1.90E-09
R	precentral_B	Intercept	3.61	6.29E-14
		CP>SP	1.26	1.61E-05
		CP>NR	2.43	2.34E-13
R	antParietal	Intercept	2.12	5.07E-13
		CP>SP	1.09	5.91E-05
		CP>NR	1.61	6.99E-09
R	midFrontalOrb	Intercept	1.48	1.71E-05
		CP>SP	0.87	0.006

R	medialFrontal	<i>CP&gt;NR</i>	0.8	0.016
		Intercept	1.96	5.46E-11
		<i>CP&gt;SP</i>	0.58	0.004
		<i>CP&gt;NR</i>	1.09	2.04E-08
R	midParietal	Intercept	2.98	1.53E-11
		<i>CP&gt;SP</i>	1.31	1.14E-04
		<i>CP&gt;NR</i>	2.1	2.62E-09
R	precentral_A	Intercept	3.62	7.18E-13
		<i>CP&gt;SP</i>	1.26	1.38E-05
		<i>CP&gt;NR</i>	2.14	1.34E-11
R	supFrontal	Intercept	3.19	1.00E-16
		<i>CP&gt;SP</i>	1.04	6.87E-04
		<i>CP&gt;NR</i>	2.57	8.34E-14
R	insula	Intercept	1.02	1.12E-09
		<i>CP&gt;SP</i>	0.33	0.135
		<i>CP&gt;NR</i>	0.54	4.98E-04

**Supplemental table 2-2.** Responses to ScratchJr code problems (CP) vs. sentence problems (SP) and nonword reading (NR) in individual fROIs within the multiple demand system. P values are Bonferroni-corrected for the number of regions. Non-significant values are italicized and marked in gray.

Hemisphere	fROI	Regression Term	Beta	p value
L	postParietal	Intercept	3.45	2.76E-14
		<i>CP&gt;SP</i>	2.13	1.88E-06
		<i>CP&gt;NR</i>	2.65	1.33E-08
L	midParietal	Intercept	2.63	9.39E-13
		<i>CP&gt;SP</i>	1.34	1.23E-04
		<i>CP&gt;NR</i>	1.65	2.75E-06
L	antParietal	Intercept	2.42	2.42E-14
		<i>CP&gt;SP</i>	1.3	0.001
		<i>CP&gt;NR</i>	1.42	2.75E-04
L	supFrontal	Intercept	2.02	4.13E-07
		<i>CP&gt;SP</i>	0.57	1
		<i>CP&gt;NR</i>	1.14	0.007
L	precentral_A	Intercept	3.17	1.43E-07
		<i>CP&gt;SP</i>	0.85	1
		<i>CP&gt;NR</i>	0.8	1
L	precentral_B	Intercept	3.15	2.98E-10
		<i>CP&gt;SP</i>	1.05	0.031
		<i>CP&gt;NR</i>	1.33	0.002
L	midFrontal	Intercept	2.61	2.63E-09
		<i>CP&gt;SP</i>	1.34	0.01
		<i>CP&gt;NR</i>	1.5	0.003
L	midFrontalOrb	Intercept	1.7	1.23E-05
		<i>CP&gt;SP</i>	1.18	0.017
		<i>CP&gt;NR</i>	0.89	0.182
L	insula	Intercept	1.18	3.61E-07
		<i>CP&gt;SP</i>	0.58	0.161
		<i>CP&gt;NR</i>	0.12	1

L	medialFrontal	Intercept	1.6	9.49E-07
		<i>CP&gt;SP</i>	0.67	0.416
		<i>CP&gt;NR</i>	0.47	1
R	postParietal	Intercept	3.93	2.11E-16
		<i>CP&gt;SP</i>	2.89	6.90E-07
		<i>CP&gt;NR</i>	3.22	6.16E-08
R	midParietal	Intercept	3.35	1.75E-09
		<i>CP&gt;SP</i>	2.17	2.12E-04
		<i>CP&gt;NR</i>	1.83	0.002
R	antParietal	Intercept	2.4	1.43E-15
		<i>CP&gt;SP</i>	1.59	1.67E-05
		<i>CP&gt;NR</i>	1.53	3.58E-05
R	supFrontal	Intercept	2.32	1.95E-10
		<i>CP&gt;SP</i>	1.24	0.007
		<i>CP&gt;NR</i>	1.21	0.009
R	precentral_A	Intercept	3.52	8.22E-09
		<i>CP&gt;SP</i>	2.02	2.26E-06
		<i>CP&gt;NR</i>	1.27	0.004
R	precentral_B	Intercept	3.32	1.99E-10
		<i>CP&gt;SP</i>	2.15	1.67E-07
		<i>CP&gt;NR</i>	1.48	2.21E-04
R	midFrontal	Intercept	2.91	2.76E-11
		<i>CP&gt;SP</i>	1.77	4.57E-05
		<i>CP&gt;NR</i>	1.26	0.007
R	midFrontalOrb	Intercept	1.71	3.49E-06
		<i>CP&gt;SP</i>	1.26	0.006
		<i>CP&gt;NR</i>	0.34	1
R	insula	Intercept	1.19	1.02E-07
		<i>CP&gt;SP</i>	0.69	0.088
		<i>CP&gt;NR</i>	0.2	1
R	medialFrontal	Intercept	1.57	4.39E-08
		<i>CP&gt;SP</i>	0.89	0.064
		<i>CP&gt;NR</i>	0.31	1

# Chapter 3

## The language network & object semantics

### Collaborators:

Yael Benn, Oliver Clark, Zachary Mineroff, Chloe Seikus, Jack Santos Silva,  
Rosemary Varley, Evelina Fedorenko

A related study (including a different set of analyses on the same data) is published as a preprint:

Benn\*, Y., Ivanova\*, A. A., Clark, O., Mineroff, Z., Seikus, C., Silva, J. S., Varley, R. & Fedorenko, E. (2021). No evidence for a special role of language in feature-based categorization. *bioRxiv*. <https://doi.org/10.1101/2021.03.18.436075v2>

### 3.1 Abstract

The relationship between language and thought is the subject of long-standing debate. Here, we test a specific aspect of this relationship: engagement of language vs. domain-general resources in object categorization. We test two hypotheses: 1. If the target category is specified with a verbal label, the label gets activated every time a categorization decision is made, leading to activity within the language network; 2. Categories that involve accessing semantic information about the object (e.g., “dangerous animals”) will recruit the language network to a greater extent than categories based on a perceptual attribute (e.g., “things that are blue”). In an fMRI study, participants read a category label and then viewed a sequence of images; for each image, they had to indicate whether it belonged to a given category; the process was repeated for 32 categories. We found low responses in the language network during both semantic and perceptual categorization, thus disconfirming both hypotheses. The domain-general multiple demand network was recruited equally strongly during semantic and perceptual categorization, suggesting that the tasks are comparable in difficulty. We conclude that

the language network is not recruited for object categorization in general or for semantic categorization in particular.

## 3.2 Introduction

The role of language in mediating or augmenting thought has been subject to much discussion (see [Chapter 1](#)). According to one view, language processing mechanisms are recruited for many cognitive functions, such as math, logic, and thought (e.g., Baldo et al., 2010, 2015; Bermúdez, 2007; Bickerton, 1995; Carruthers, 2002; Darwin, 1871; Dennett, 1994). However, a large body of evidence supports a different view: that language is cognitively and neurally independent from the rest of human cognition. This evidence includes the lack of activity in the language brain regions during non-linguistic tasks that allegedly require language (e.g., Amalric & Dehaene, 2016, 2019; Fedorenko, Behr, et al., 2011; Ivanova et al., 2021; Monti et al., 2009, 2012), the retained ability of some individuals with aphasia to perform such tasks (e.g., Bek et al., 2013; Benn et al., 2013; Siegal & Varley, 2006; Varley et al., 2005), and variability across cultures in the use of language resources during thought (Kim, 2002). Here, we investigate a specific case of putative language-thought interaction: categorizing images based on a verbal label.

Like other animals, humans can convert rich, multi-dimensional perceptual inputs into a latent lower-dimensional structured representation of the world (e.g., Mareschal & Quinn, 2001; Mervis & Rosch, 1981; Murphy, 2002; Pearce, 1994; E. E. Smith & Medin, 1981; L. B. Smith & Heise, 1992; Wasserman et al., 1988). In contrast to other animals, humans additionally label individual categories with words—the core building blocks of a powerful communication system that allows us to share complex thoughts with one another. Even though categorization is a basic cognitive capacity that evolved long before language, there is evidence that word learning affects category learning in development (e.g., Ferguson & Waxman, 2017; Gershkoff-Stowe et al., 1997; Plunkett et al., 2008; Sloutsky & Fisher, 2004; Waxman & Gelman, 2009) and, to some extent, in adulthood (Brojde et al., 2011; Lupyan et al., 2007; Lupyan & Casasanto, 2015). Here, we focus not on category learning, but on the process of categorization itself. Specifically, we ask: do

people recruit language processing mechanisms when grouping objects into categories that have been specified via a verbal label?

One account (Luo et al., 2021; Lupyan, 2012; Regier & Kay, 2009) posits that linguistic labels get activated during categorization and are used to highlight the features relevant for categorization. Evidence in favor of this account comes from studies that show impaired categorization performance when access to language resources is reduced, such as when performing a verbal interference task (Lupyan, 2009), during transcranial direct current stimulation (Lupyan et al., 2012), or in individuals with aphasia (Lupyan & Mirman, 2013). Indirect evidence of the involvement of linguistic resources also comes from left-lateralized EEG responses during categorization of exemplars across linguistic category bounds (Gilbert et al., 2006; cf. K. J. Holmes & Wolff, 2012).

Interpreting some of this evidence is complicated for two reasons. First, an overall performance measure cannot disentangle the relative contributions of object categorization vs. instruction processing. Given that instructions in categorization tasks are often presented via a verbal category label, impaired performance might result from participants failing to understand task instructions rather than failing to perform the categorization task itself. Second, the language processing regions in the human brain are closely adjacent to other, non-language specific regions, meaning that damage resulting from a stroke might affect both linguistic and non-linguistic mechanisms. In particular, the language network in the left hemisphere, especially in the left frontal cortex, lies adjacent to the domain-general multiple demand network, which supports executive functions, like working memory and inhibitory control (Assem, Glasser, et al., 2020; Duncan, 2010, 2013; Fedorenko, Duncan, et al., 2012; Fedorenko et al., 2013). As a result, left hemisphere damage can lead to joint linguistic and domain-general executive deficits (Baldo et al., 2010; Gainotti et al., 1986) and left-lateralized activity might, in principle, arise from non-linguistic neural mechanisms (K. J. Holmes & Wolff, 2012).

In fact, existing studies provide mixed results with respect to the role of language in categorization. Various studies of individuals with aphasia describe general categorization impairments (Koemeda-Lutz et al., 1987), deficits for specific category

types (e.g., Burger & Muma, 1980; Lupyán & Mirman, 2013), or no deficits at all (Hough, 1993). Further, variations in the task (such as showing the category label to the participant during the entire trial vs. just at the beginning of the trial) significantly affected categorization performance in participants with aphasia (Koameda-Lutz et al., 1987), suggesting that task demands may contribute to the observed results. Executive task demands might also underlie the effects of transcranial stimulation and verbal interference described above (given the close adjacency of language and multiple demand regions and the fact that verbal and non-verbal interference tasks are hard to match for executive demand). Finally, some aphasia researchers have argued for a relationship between categorization difficulties and conceptual-semantic rather than purely linguistic impairments (Caramazza et al., 1982; Whitehouse et al., 1978; cf. Le Dorze & Nesporous, 1989), necessitating a clearer distinction between linguistic and conceptual processing.

Here, we use functional MRI to test the recruitment of linguistic vs. domain-general resources in label-based categorization. Our goal was to test (a) whether linguistic resources are active during label-based categorization in general, and (b) whether linguistic resources are recruited specifically during semantic categorization, in line with accounts implicating language in semantic processing (e.g., Binder et al., 2009; Binder & Desai, 2011). To do so, we asked participants to complete the critical categorization task along with two localizer experiments: language localizer and multiple demand localizer.

The language localizer is designed to identify brain regions that respond more strongly to sentences than to pronounceable but meaningless sequences of letters (“nonwords”). A large number of studies has shown that this sentences>nonwords contrast picks out a set of regions that are strongly and selectively recruited for language processing, including spoken, written, and signed language comprehension, spoken and written language production, and inner speech (Amit et al., 2017; Braga et al., 2020; Fedorenko et al., 2010, 2011; Giglio et al., 2021; Hu, Small, et al., 2021; Menenti et al., 2011; Scott et al., 2017; Silbert et al., 2014). These regions also respond to linguistic units at different levels of the processing hierarchy, including both phrases and single words (albeit no single region is sensitive just to word-level or sentence-level meaning; Blank, Balewski, et al.,

2016; Blank & Fedorenko, 2020). Therefore, if a certain task requires activating linguistic representations, we expect to observe activity in regions identified with the language localizer (henceforth called the language network).

The multiple demand localizer identifies a set of regions that respond to a wide range of cognitively demanding tasks. Specifically, these regions are sensitive to general cognitive effort, exhibiting higher activity when the task is more difficult (Assem, Glasser, et al., 2020; Duncan, 2010; Fedorenko et al., 2013; Hugdahl et al., 2015). The hard>easy response signature in the multiple demand network holds across many diverse tasks, including spatial working memory, logic, math, relational reasoning, and cognitive control (Assem, Glasser, et al., 2020; Coetzee & Monti, 2018; Fedorenko et al., 2013; Shashidhara, Mitchell, et al., 2019). Thus, if a task is cognitively challenging, we expect it to elicit activity in the multiple demand network.

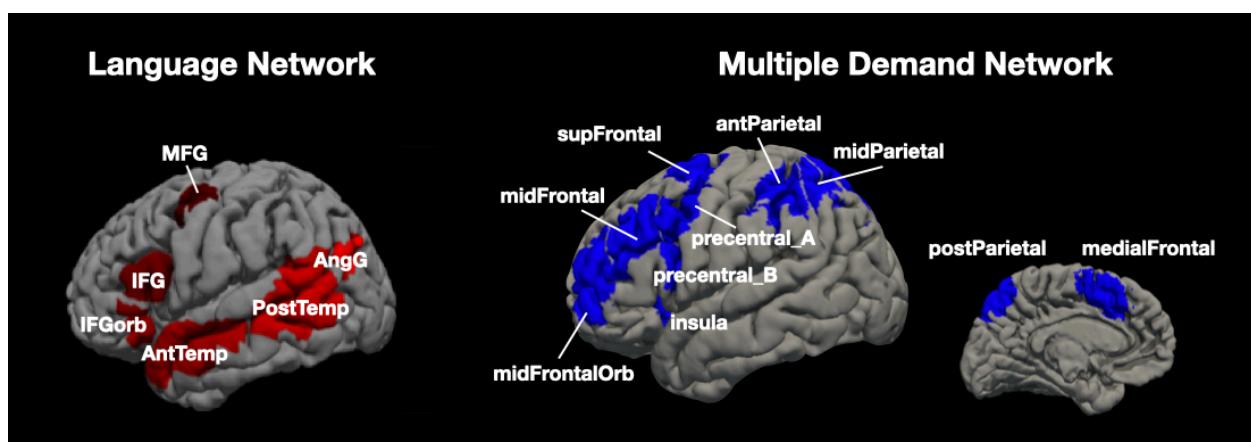
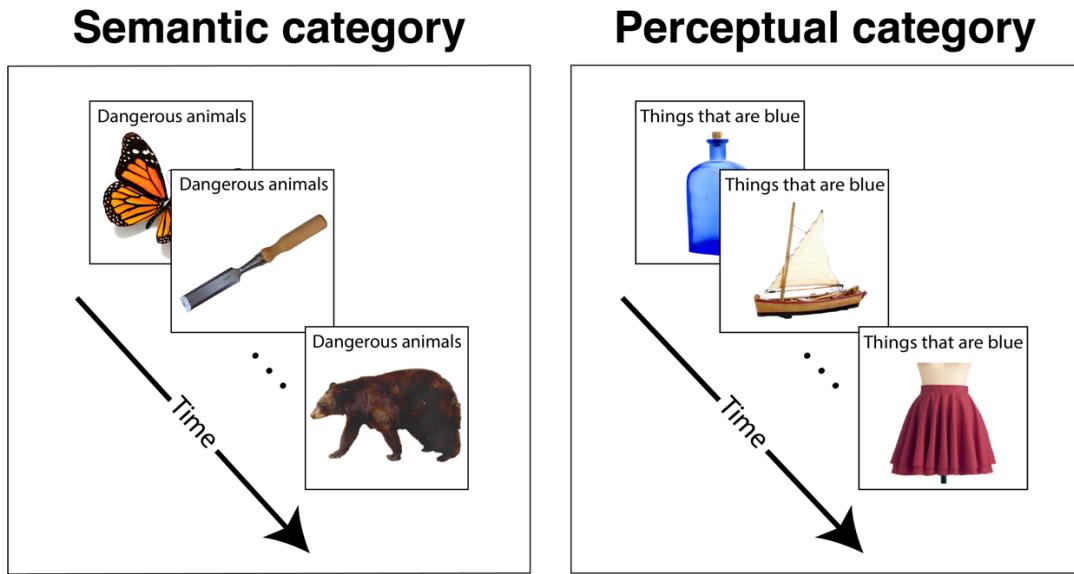


Figure 3-1. Parcels used to define the language and multiple demand networks.

Examining activation patterns in both the language and the multiple demand networks (**Figure 3-1**) allows us to examine the relative contributions of linguistic and cognitive control resources to semantic and perceptual categorization. If linguistic resources are engaged during categorization, we would expect an overall high response of the language network to categorization conditions. Further, if these regions are specifically engaged during semantic processing, the response to semantic categories should exceed the response to perceptual categories. If either category type loads specifically onto the

domain-general resources, we would expect to see greater responses to that category type in the domain-general multiple demand regions. Finally, if a brain network does not respond to either semantic or perceptual categorization, we can conclude that this network is not recruited for this task.



**Figure 3-2.** Sample semantic and perceptual categorization trials. The category label was presented at the center of the screen at the beginning of each block (not shown) and was additionally shown at the top of the screen during each trial.

### 3.3 Method

#### 3.3.1 Participants

Fourteen neurotypical participants (7 F, age  $M=22.31$ ,  $SD=3.51$ ) were recruited from MIT and the surrounding community and paid \$60 for their participation. All were native speakers of English. One participant was left-handed (see Willems et al., 2014, for motivation to include left-handers in cognitive neuroscience research) but showed typical left-lateralized language activation as determined by the language localizer task (described below). All participants gave informed consent in accordance with the

requirements of MIT's Committee On the Use of Humans as Experimental Subjects (COUHES).

### 3.3.2 Design, materials, and procedure

Each participant completed a language localizer task aimed at identifying language-responsive brain regions (Fedorenko et al., 2010), a spatial working memory (WM) task aimed at identifying the multiple demand network (Fedorenko et al., 2013), and the critical categorization task. Some participants completed one or more additional tasks for unrelated studies. The entire scanning session lasted two hours.

#### 3.3.2.1 Categorization

The critical categorization task was modeled on the study by Lupyan & Mirman (2013). We used 32 of the 34 categories used by L&M (dropping 'BODY PARTS' and 'FACIAL FEATURES'). Unlike L&M, who used normed color drawings (Rossion & Pourtois, 2004), we used high-quality color photographs selected from the Hemera Photo Objects 5000 and Google Images. For each category, we selected 8-15 targets and 25-27 distractors. Distractors included some items which were related to the target category (for example, for the category 'DANGEROUS ANIMALS', 13 of the 26 distractors were animals that were not dangerous, and the rest were not animals, and the category 'ANIMALS WITH STRIPES' included distractors that were animals without stripes, and inanimate objects with stripes). A total of 1,087 unique images were used (any given image appeared as a target in 0-2 categories and as a distractor in 0-2 categories). All photographs depicted objects on a white background.

During the critical experiment, participants viewed a sequence of images and were asked to indicate whether the object depicted in the image belonged to a given category (**Figure 3-2**). Each block started with a category label presented for 2s, followed by 12 images presented sequentially at the fixed speed of 2s per image. The category label remained on the screen to minimize memory demands. Participants were asked to press a button if the picture belonged to the target category and not to press anything if it did not. The images for each category block were randomly selected from the general set of pictures for that

category; the order of images within category blocks was therefore randomized for each participant. The number of targets varied across blocks (minimum: 4, maximum: 6) so as to minimize the implicit learning of a fixed number of targets, which could have incentivized participants to keep track of the total number of targets. Category blocks lasted 26s (2s category label presentation + 2s \* 12 images), and fixation blocks lasted 14s. Each run, consisting of 12 category blocks and 4 fixation blocks, lasted 368s. Each participant completed 3 runs. Across the 3 runs, any given participant saw a random subset of the 32 categories, with some categories repeating (but never repeating within a run; see **Supplemental table 3-3** for details). Category order was counterbalanced across runs and participants.

### **3.3.2.2 Language localizer**

The language localizer task was the same as that described in **Section 2.5.2.3**.

### **3.3.2.3 Multiple demand localizer**

The multiple demand localizer task was the same as that described in **Section 2.5.2.2**. Twelve participants completed two runs and two participants completed one run.

## **3.3.3 fMRI data acquisition**

The data acquisition procedure was the same as that described in **Section 2.5.3**.

## **3.3.4 fMRI data preprocessing**

The preprocessing procedure was the same as that described in **Section 2.5.4**.

## **3.3.5 First-level analysis**

First-level analysis procedure was the same as that described in **Section 2.5.5**. For the localizer tasks, we modeled entire blocks. For the picture plausibility task, we modeled entire blocks (labeling them as semantic categorization vs. perceptual categorization)

with the exception of the instruction reading phase, which was modeled as a separate condition.

### **3.3.6 Defining individual functional regions of interest (fROIs)**

The procedure for defining language and multiple demand parcels and fROIs is described in Section 2.5.6. For the putative SEM>PERC categorization regions, we used the same procedure to generate the parcels based on the data collected in this study.

### **3.3.7 Examining the functional response profiles of fROIs**

After defining fROIs in individual participants, we evaluated their responses to the conditions of interest by averaging the responses across voxels to get a single value per condition per fROI. Responses to conditions used to define the fROI were estimated using an across-runs cross-validation procedure, where one run was used to define the fROI and the other to estimate the response magnitudes, then the procedure was repeated switching which run was used for fROI definition vs. response estimation, and finally the estimates were averaged to derive a single value per condition per fROI per participant. This cross-validation procedure allows one to use all of the data for defining the fROIs as well as for estimating their responses (see Nieto-Castañón & Fedorenko, 2012, for discussion), while ensuring the independence of the data used for fROI definition and response estimation (Kriegeskorte et al., 2009).

Two participants completed only one run of the multiple demand localizer task; therefore, we did not estimate the strength of their responses to the hard and easy multiple demand localizer conditions but ensured that the whole-brain activation maps for the hard>easy contrast looked as expected.

### **3.3.8 Statistical analyses**

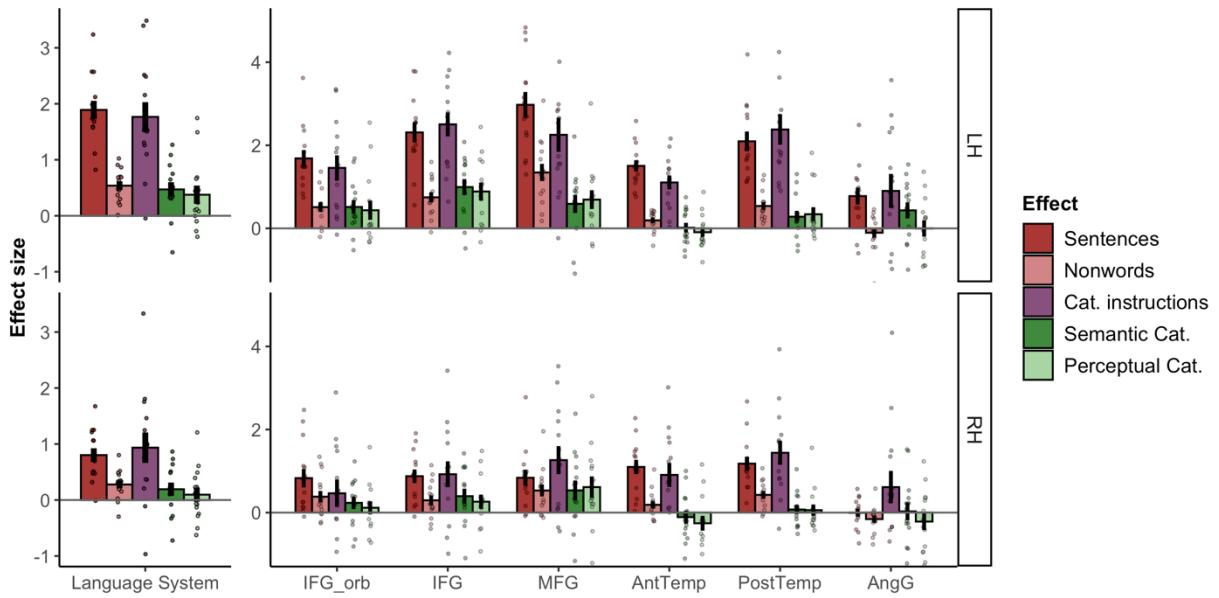
We analyzed our data using mixed effect regression models (Baayen et al., 2008). In all models, condition was a fixed effect and participant was a random intercept. The model

for the multiple demand network included hemisphere as an additional fixed effect. For language and multiple demand network analyses, we also included fROI as a random intercept and then ran follow-up analyses on individual fROIs using false discovery rate (FDR) correction (Benjamini & Hochberg, 1995) for the number of fROIs in each network. Neuroimaging analyses used custom contrasts (see [Section 3.7.1 \(Appendix\)](#) for detailed contrast specification). The mixed effect analyses were run using the *lmer* function from the *lme4* R package (Bates et al., 2015); statistical significance of the effects was evaluated using the *lmerTest* package (Kuznetsova et al., 2017). The hypotheses-specific contrasts were defined using the *hypr* package (Rabe et al., 2020).

## 3.4 Results

### 3.4.1 The language network shows low responses to both semantic and perceptual categorization

In the language network, response to categorization was not significantly different from 0 ( $\beta=0.28$ ,  $SE=0.18$ ,  $p=.128$ ; see [Figure 3-3](#)). Response to semantic categories was not significantly different from response to perceptual categories ( $\beta=0.10$ ,  $SE=0.09$ ,  $p=.272$ ) or nonword reading ( $\beta=0.08$ ,  $SE=0.09$ ,  $p=.385$ ), and significantly weaker than responses to sentences ( $\beta=1.01$ ,  $SE=.07$ ,  $p<.001$ ) and instruction reading ( $\beta=1.02$ ,  $SE=.09$ ,  $p<.001$ ). Follow-up analyses in individual language fROIs ([Supplemental table 3-1](#)) showed that none of them had significantly different responses to semantic and perceptual categories. Four fROIs (lMFG, lIFG, lIFGorb, and rIFG) had above-zero responses to semantic categorization, but these responses were not significantly different from responses during the control nonword reading task. Thus, our results suggest that the language network is not involved in either semantic or perceptual categorization in neurotypical participants.

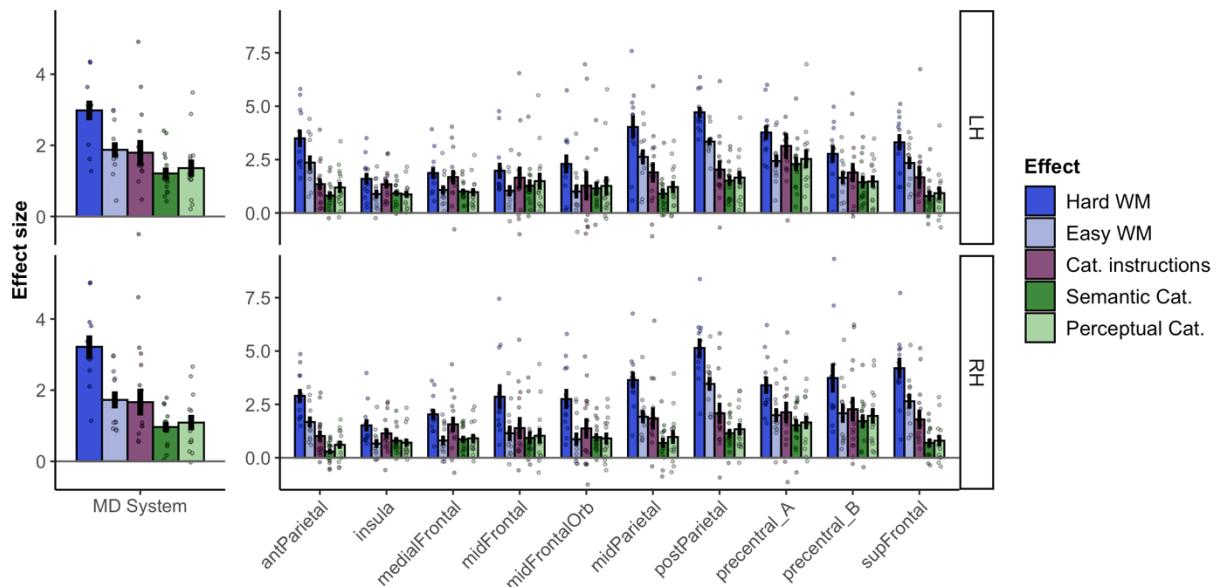


**Figure 3-3.** Categorization responses within the language brain network. Top: left hemisphere (LH); bottom: right hemisphere (RH). Left: responses averaged across fROIs; right: responses in individual fROIs. Abbreviations: Cat. – categorization, IFG – inferior frontal gyrus, orb – pars orbitalis, MFG – middle frontal gyrus, AntTemp – anterior temporal lobe, PostTemp – posterior temporal lobe, AngG – angular gyrus. Here and elsewhere, dots show values for individual participants, and error bars show standard error of the mean.

### 3.4.2 The multiple demand network responds comparably to semantic and perceptual categories

The multiple demand network had equally strong responses to semantic and perceptual categories ( $\beta=0.14$ ,  $SE=0.09$ ,  $p=.139$ ; see **Figure 3-4**), which were significantly above 0 ( $\beta=1.16$ ,  $SE=0.22$ ,  $p<.001$ ). Responses to semantic categorization were stronger than responses to control conditions from the language localizer task (semantic categorization  $>$  sentences:  $\beta=0.73$ ,  $SE=0.09$ ,  $p<.001$ ; semantic categorization  $>$  nonwords:  $\beta=0.41$ ,  $SE=0.09$ ,  $p<.001$ ). However, they were weaker than responses to the spatial working memory task ( $\beta=-1.41$ ,  $SE=.08$ ,  $p<.001$ ), indicating that the working memory task was more effortful. They were also weaker than responses to the instruction stage of the categorization task ( $\beta=-0.64$ ,  $SE=.09$ ,  $p<.001$ ), indicating that the most cognitively demanding stage of the task is the initial category label processing. The responses to categorization were stronger in the left hemisphere ( $\beta=0.26$ ,  $SE=0.09$ ,  $p=.005$ ), but there

was no interaction between hemisphere and category type ( $\beta=0.02$ ,  $SE=0.18$ ,  $p=.895$ ). Follow-up analyses in individual fROIs (**Supplemental table 3-2**) showed that no fROIs had a significant difference between semantic and perceptual categorization. Responses to categorization overall were significantly above 0 in all fROIs, but weaker than the overall responses to the working memory task in almost all fROIs (except left middle frontal and middle frontal orbital fROIs). Thus, the MD network was engaged in categorization but did not show a preference for either semantic or perceptual categories.



**Figure 3-4.** Categorization responses within the multiple demand brain network. Top: left hemisphere (LH); bottom: right hemisphere (RH). Left: responses averaged across fROIs; right: responses in individual fROIs. Abbreviations: WM – working memory task, Cat. – categorization, ant – anterior, mid – middle, post – posterior, precentral\_A – dorsal precentral gyrus, precentral\_B – ventral precentral gyrus.

### 3.4.3 Whole-brain analyses reveal no regions responsive specifically to semantic categories

We conducted a whole-brain analysis to identify fROIs that might respond more strongly to semantic or perceptual categorization but lie outside the language and multiple demand fROIs described above. The GSS analysis (see **Section 3.3.6**) revealed no regions that passed the threshold of having significant voxels in at least 60% of participants,

indicating that no brain region exhibits a consistent preference for either semantic or perceptual categorization (as defined in our study).

### 3.5 Discussion

In this experiment, we used fMRI to examine responses to label-based object categorization in the human brain. The language network exhibited low responses during categorization, indicating that linguistic processing mechanisms are not involved in object categorization, even when the category label is presented verbally beforehand. Moreover, language network responses were low even when the category was semantic, i.e., when participants needed to access conceptual information about the objects. This result goes against the claim that semantic processing relies on language processing mechanisms.

Other results from psycho- and neurolinguistics also support the view that linguistic resources do not typically mediate categorization in humans. If access to linguistic representations were necessary for categorization, categorizing images would take longer than categorizing words; instead, they take approximately the same amount of time (Potter & Faulconer, 1975). When asked to match a picture with a label, participants do not explicitly generate/rehearse verbal labels in advance unless there is an additional memory demand (e.g., if images disappear from the screen) (Pontillo et al., 2015). Our work therefore adds to the growing body of evidence for a separation between linguistic and visual semantic processing.

That said, many studies have shown that linguistic labels influence categorization behavior in infants (e.g., Ferguson & Waxman, 2017; Gershkoff-Stowe et al., 1997; Plunkett et al., 2008; Sloutsky & Fisher, 2004; Waxman & Gelman, 2009) and adults (e.g., Brojde et al., 2011; Lupyan, 2009; Lupyan et al., 2007; Zettersten & Lupyan, 2020), so the relationship between words and categories is clearly an important one. What we are showing here is that the mechanisms responsible for language processing are not engaged during object categorization, nor are they specifically recruited for semantic categorization. It is possible that linguistic labels, once acquired, may influence

categorization via other brain systems, e.g., semantic, domain-general, or perceptual. The cognitive and neural mechanisms underlying the influence of labels on categorization thus remain to be determined (see, e.g., Gliozzi et al., 2009; Ivanova & Hofer, 2020; Luo et al., 2021; Lupyan, 2012, for some modeling proposals). And, of course, linguistic processing is required to process the category label before performing categorization itself, which is reflected in high language network responses during the instruction presentation stage.

The multiple demand network responded during categorization, although it showed no preference toward either semantic or conceptual categories. The response to categorization within the multiple demand network was stronger in the left hemisphere, consistent with the view that label-based categorization recruits the left hemisphere more strongly (e.g., A. Franklin et al., 2008; Gilbert et al., 2006). This makes the categorization task similar to logic and math, which also evoke left-lateralized responses within the multiple demand network (Amalric & Dehaene, 2016; Monti et al., 2009, 2012; Pinel & Dehaene, 2009). Importantly, our result demonstrates that, just because the function is left-lateralized, it is not necessarily related to language, at least not in fully formed brains (contra, e.g., Gilbert et al., 2006; see also K. J. Holmes & Wolff, 2012).

Overall, our study shows that object categorization is not a language-dependent task in the adult brain, regardless of whether the categorization is made on the basis of semantic or perceptual features. Instead, this task relies on the domain-general multiple demand system, which supports diverse goal-directed behaviors. Our work provides evidence against the view of language as an aid for categorization and suggests that, even when people receive information verbally, they convert and store it in a nonverbal format prior to performing other cognitive tasks.

### 3.6 Acknowledgements

We would like to acknowledge the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT, and its support team (Steve Shannon and Atsushi Takahashi). EF was supported by NIH awards R00-HD057522, R01-DC016607,

R01-DC016950, a grant from the Simons Foundation to the Simons Center for the Social Brain at MIT, and by support from the Brain and Cognitive Sciences Department and the McGovern Institute for Brain Research at MIT. The authors thank Naveen Hanif and Anis Adila Khairil Anuar, who have helped with development and piloting of the paradigms, and Alvincé Pongos for help with data analysis.

## 3.7 Appendix

### 3.7.1 Mixed-effects models contrast specification

Condition contrasts were designed to test the following null hypotheses.

#### Language network:

- (1)  $\frac{SEM + PERC}{2} = 0$
- (2)  $SEM = PERC$  (main)
- (3)  $SEM = Sentences$
- (4)  $SEM = Nonwords$
- (5)  $SEM = Instruction\ reading\ (categorization)$

#### Multiple demand network:

- (1)  $\frac{SEM + PERC}{2} = 0$
- (2)  $SEM = PERC$  (main)
- (3)  $HardWM = EasyWM$
- (4)  $SEM = \frac{HardWM + EasyWM}{2}$
- (5)  $SEM = Sentences$
- (6)  $SEM = Nonwords$
- (7)  $SEM = Instruction\ reading\ (categorization)$

SEM – semantic categories; PERC – perceptual categories, HardWM – hard working memory task, EasyWM – easy working memory task.

### 3.7.2 Supplemental tables

**Supplemental table 3-1.** Mixed-effect linear regression results for individual fROIs within the language network. P values are FDR-corrected for the number of regions (n=12). SEM – semantic categorization, PERC – perceptual categorization, Instruct – instruction/category label processing, S – sentence reading (language localizer), N – nonword reading (language localizer).

fROI	Hemisphere	Regression Term	Beta	p value
IFG_orb	LH	Intercept	0.48	0.031 *
		PERC-SEM	-0.08	0.866
		S-SEM	1.17	<.001 ***
		N-SEM	0	0.99
		Instruct-SEM	0.94	<.001 ***
		Intercept	0.94	<.001 ***
	RH	PERC-SEM	-0.11	0.866
		S-SEM	1.32	<.001 ***
		N-SEM	-0.25	0.505
		Instruct-SEM	1.51	<.001 ***
		Intercept	0.64	0.031 *
		PERC-SEM	0.1	0.866
MFG	LH	S-SEM	2.38	<.001 ***
		N-SEM	0.76	0.248
		Instruct-SEM	1.66	<.001 ***
		Intercept	-0.04	0.68
		PERC-SEM	-0.11	0.866
		S-SEM	1.49	<.001 ***
	RH	N-SEM	0.18	0.505
		Instruct-SEM	1.09	<.001 ***
		Intercept	0.31	0.204
		PERC-SEM	0.06	0.866
		S-SEM	1.82	<.001 ***
		N-SEM	0.26	0.505
AntTemp	LH	Instruct-SEM	2.1	<.001 ***
		Intercept	0.21	0.416
		PERC-SEM	-0.44	0.866
		S-SEM	0.34	0.207
		N-SEM	-0.54	0.248
		Instruct-SEM	0.47	0.083
	RH	Intercept	0.18	0.416
		PERC-SEM	-0.12	0.866
		S-SEM	0.59	0.016 *
		N-SEM	0.15	0.686
		Instruct-SEM	0.23	0.313
		Intercept	0.33	0.126
PostTemp	LH	PERC-SEM	-0.13	0.866
		S-SEM	0.48	0.034 *
		N-SEM	-0.1	0.765
		Instruct-SEM	0.53	0.019 *
		Intercept	0.57	0.039 *
		PERC-SEM	0.08	0.866
	RH	S-SEM	0.3	0.177
		Intercept	0.18	0.416
		PERC-SEM	-0.12	0.866
		S-SEM	0.59	0.016 *
		N-SEM	0.15	0.686
		Instruct-SEM	0.23	0.313
AngG	LH	Intercept	0.21	0.416
		PERC-SEM	-0.44	0.866
		S-SEM	0.34	0.207
		N-SEM	-0.54	0.248
		Instruct-SEM	0.47	0.083
		Intercept	0.21	0.416
	RH	PERC-SEM	-0.44	0.866
		S-SEM	0.34	0.207
		N-SEM	-0.54	0.248
		Instruct-SEM	0.47	0.083
		Intercept	0.21	0.416
		PERC-SEM	-0.44	0.866

		N-SEM	0	0.99	
AntTemp	RH	Instruct-SEM	0.73	0.001	**
		Intercept	-0.18	0.392	
		PERC-SEM	-0.15	0.866	
		S-SEM	1.21	<.001	***
		N-SEM	0.3	0.482	
PostTemp	RH	Instruct-SEM	1.02	<.001	***
		Intercept	0.07	0.68	
		PERC-SEM	-0.02	0.927	
		S-SEM	1.1	<.001	***
		N-SEM	0.35	0.248	
AngG	RH	Instruct-SEM	1.37	<.001	***
		Intercept	-0.09	0.68	
		PERC-SEM	-0.25	0.866	
		S-SEM	-0.04	0.886	
		N-SEM	-0.19	0.686	
		Instruct-SEM	0.58	0.027	*

**Supplemental table 3-2.** Mixed-effect linear regression results for individual fROIs within the multiple demand system. P values are FDR-corrected for the number of regions (n=20). SEM – semantic categorization, PERC – perceptual categorization, Instruct – instruction/category label processing, S – sentence reading (language localizer), N – nonword reading (language localizer), H – hard working memory task (multiple demand localizer), E – easy working memory task (multiple demand localizer).

fROI	Hemisphere	Regression Term	Beta	p value
postParietal	LH	Intercept	1.59	<.001 ***
		PERC-SEM	0.14	0.901
		H-E	1.37	<.001 ***
		HE-SEM	2.62	<.001 ***
		N-SEM	-0.98	0.002 **
		S-SEM	-1.43	<.001 ***
		Instruct-SEM	0.52	0.056
		Intercept	1.06	0.003 **
		PERC-SEM	0.31	0.901
		H-E	1.4	0.001 **
midParietal	LH	HE-SEM	2.38	<.001 ***
		N-SEM	0.02	0.96
		S-SEM	-0.57	0.169
		Instruct-SEM	1	0.025 *
		Intercept	1	<.001 ***
		PERC-SEM	0.38	0.901
		H-E	1.14	<.001 ***
		HE-SEM	2.17	<.001 ***
		N-SEM	-0.1	0.873 *
		S-SEM	-0.64	0.023 *
antParietal	LH	Instruct-SEM	0.54	0.064
		Intercept	0.86	0.009 **
		PERC-SEM	0.13	0.901
		H-E	0.97	0.007 **
		HE-SEM	2.11	<.001 ***
		N-SEM	-0.28	0.55
		S-SEM	-0.4	0.256
		Instruct-SEM	0.88	0.025 *
		Intercept	2.41	<.001 ***
		PERC-SEM	0.24	0.901
supFrontal	LH	H-E	1.34	<.001 ***
		HE-SEM	0.94	0.003 **
		N-SEM	-0.62	0.155
		S-SEM	-1.03	0.008 **
		Instruct-SEM	0.85	0.025 *
		Intercept	1.46	<.001 ***
		PERC-SEM	0.04	0.901
		H-E	1.12	0.001 **
		HE-SEM	0.67	0.017 *
		N-SEM	-0.68	0.091
precentral_A	LH	S-SEM	-0.79	0.021 *
		Instruct-SEM	0.45	0.18
		Intercept	1.38	<.001 ***
		PERC-SEM	0.21	0.901
		H-E	0.94	0.003 **
precentral_B	LH			
midFrontal	LH			

		HE-SEM	0.27	0.299	
		N-SEM	-0.83	0.032	*
		S-SEM	-1.21	<.001	***
		Instruct-SEM	0.38	0.216	
midFrontalOrb	LH	Intercept	1.21	0.009	**
		PERC-SEM	0.12	0.901	
		H-E	1.29	0.002	**
		HE-SEM	0.67	0.052	
		N-SEM	-0.58	0.225	
		S-SEM	-1.04	0.016	*
		Instruct-SEM	0.13	0.727	
insula	LH	Intercept	0.89	<.001	***
		PERC-SEM	-0.04	0.901	
		H-E	0.72	<.001	***
		HE-SEM	0.38	0.006	**
		N-SEM	-0.45	0.021	*
		S-SEM	-0.57	<.001	***
medialFrontal	LH	Instruct-SEM	0.43	0.016	*
		Intercept	0.99	<.001	***
		PERC-SEM	-0.03	0.901	
		H-E	0.79	0.001	**
		HE-SEM	0.51	0.012	*
		N-SEM	-0.43	0.124	
		S-SEM	-0.6	0.016	*
postParietal	RH	Instruct-SEM	0.67	0.016	*
		Intercept	1.24	<.001	***
		PERC-SEM	0.19	0.901	
		H-E	1.68	<.001	***
		HE-SEM	3.18	<.001	***
		N-SEM	-0.84	0.068	
		S-SEM	-1.14	0.006	**
midParietal	RH	Instruct-SEM	0.94	0.025	*
		Intercept	0.83	0.006	**
		PERC-SEM	0.28	0.901	
		H-E	1.72	<.001	***
		HE-SEM	2.14	<.001	***
		N-SEM	0.12	0.873	
		S-SEM	-0.36	0.375	
antParietal	RH	Instruct-SEM	1.16	0.016	*
		Intercept	0.45	0.009	**
		PERC-SEM	0.32	0.901	
		H-E	1.23	<.001	***
		HE-SEM	2.02	<.001	***
		N-SEM	0.06	0.928	
		S-SEM	-0.2	0.515	
supFrontal	RH	Instruct-SEM	0.73	0.029	*
		Intercept	0.75	0.015	*
		PERC-SEM	0.11	0.901	
		H-E	1.55	<.001	***
		HE-SEM	2.76	<.001	***
		N-SEM	-0.05	0.948	
		S-SEM	-0.2	0.594	
precentral_A	RH	Instruct-SEM	1.1	0.019	*
		Intercept	1.59	<.001	***

		PERC-SEM	0.13	0.901	
		H-E	1.4	0.001	**
		HE-SEM	1.18	0.002	**
		N-SEM	-0.4	0.438	
		S-SEM	-0.74	0.074	
		Instruct-SEM	0.6	0.157	
precentral_B	RH	Intercept	1.84	<.001	***
		PERC-SEM	0.24	0.901	
		H-E	1.65	<.001	***
		HE-SEM	1.23	0.003	**
		N-SEM	-0.68	0.225	
		S-SEM	-1.06	0.023	*
		Instruct-SEM	0.56	0.216	
midFrontal	RH	Intercept	0.98	0.009	**
		PERC-SEM	0.1	0.901	
		H-E	1.71	<.001	***
		HE-SEM	1.03	0.003	**
		N-SEM	-0.24	0.662	
		S-SEM	-0.63	0.105	
midFrontalOrb	RH	Instruct-SEM	0.46	0.216	
		Intercept	0.94	0.004	**
		PERC-SEM	-0.06	0.901	
		H-E	1.89	<.001	***
		HE-SEM	0.87	0.012	*
		N-SEM	-0.42	0.404	
		S-SEM	-0.91	0.023	*
		Instruct-SEM	0.41	0.272	
insula	RH	Intercept	0.75	0.001	**
		PERC-SEM	-0.1	0.901	
		H-E	0.85	<.001	***
		HE-SEM	0.32	0.019	*
		N-SEM	-0.39	0.048	*
		S-SEM	-0.52	0.003	**
		Instruct-SEM	0.34	0.04	*
medialFrontal	RH	Intercept	0.89	<.001	***
		PERC-SEM	0.04	0.901	
		H-E	1.23	<.001	***
		HE-SEM	0.6	0.006	**
		N-SEM	-0.36	0.225	
		S-SEM	-0.61	0.02	*
		Instruct-SEM	0.7	0.016	*

**Supplemental table 3-3.** Distribution of category use across participants (i.e., the number of times each participant saw each category during the categorization experiment, summed across runs).

Condition	Category	Subject ID													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
SEM	animals that live in water	1	1	1	0	0	0	1	1	2	1	1	2	2	2
SEM	birds	2	0	1	1	1	3	1	2	1	2	3	1	3	1
SEM	clothes	1	1	1	1	1	0	1	0	0	2	1	0	1	1
SEM	dangerous animals	2	1	2	1	1	0	1	1	1	1	1	0	0	1
SEM	farm animals	1	1	3	2	0	0	2	2	2	1	3	1	2	1
SEM	fruit	1	2	1	1	2	2	0	0	1	1	1	2	1	0
SEM	home appliances	2	2	0	1	1	2	2	1	0	1	0	2	0	1
SEM	insects	2	2	1	3	1	2	1	2	1	2	1	2	1	2
SEM	musical instruments	0	1	1	0	3	0	2	1	2	2	2	1	1	1
SEM	non food things found in the kitchen	2	1	0	0	2	2	0	2	2	1	1	2	0	1
SEM	objects found in the laundry room	1	1	0	2	1	1	1	2	1	1	1	0	1	1
SEM	objects that hold water	0	1	2	2	0	1	2	0	0	1	1	1	2	2
SEM	objects used for transportation	1	2	1	1	1	1	0	1	3	1	0	0	1	1
SEM	things that fly	1	2	3	1	0	1	2	2	1	0	1	2	0	2
SEM	tools	1	0	0	1	2	2	2	0	0	0	0	1	0	0
SEM	vegetables	0	0	1	0	2	1	0	1	1	1	1	1	3	1
PERC	animals with stripes	2	1	2	2	0	1	2	0	1	2	0	1	0	2
PERC	long thin objects	0	0	2	2	1	1	1	1	0	2	2	1	2	2
SEM	small objects	1	2	0	1	2	1	0	1	1	1	1	0	1	1
PERC	things made of wood	1	1	1	3	2	1	2	1	0	1	2	2	0	0
PERC	things that are blue	2	1	1	1	1	1	1	1	0	0	1	0	1	1
PERC	things that are brown	1	2	1	0	3	1	2	2	2	1	2	0	1	1
PERC	things that are green	1	0	1	1	0	2	1	1	2	2	1	1	0	1
PERC	things that are orange	2	1	2	2	1	2	0	1	1	1	0	2	2	1
PERC	things that are red	1	1	1	1	0	1	1	1	2	0	2	1	2	1
PERC	things that are round	1	1	1	1	1	1	0	1	3	1	2	0	1	1
SEM	things that are soft	0	0	2	0	2	1	1	2	1	0	2	2	1	2
SEM	things that are very large	1	1	1	2	1	2	1	1	1	1	1	2	3	0
PERC	things that are white	3	2	1	1	1	0	1	0	1	1	0	3	0	2
PERC	things that are yellow	2	3	0	0	0	1	2	2	1	3	1	1	2	0
SEM	things with doors	0	1	1	0	1	1	2	2	1	1	0	1	2	3
SEM	things with handles	0	1	1	1	2	1	1	1	1	1	1	1	0	0

# Chapter 4

## The language network & event semantics

### Collaborators:

Zachary Mineroff, Vitor Zimmerer, Nancy Kanwisher, Rosemary Varley,  
Evelina Fedorenko

This chapter has been published as a journal article in *Neurobiology of Language* (under a CC BY 4.0 license):

Ivanova, A. A., Mineroff, Z., Zimmerer, V., Kanwisher, N., Varley, R., & Fedorenko, E. (2021). The language network is recruited but not required for nonverbal event semantics. *Neurobiology of Language*, 2(2), 176-201. [https://doi.org/10.1162/nol\\_a\\_00030](https://doi.org/10.1162/nol_a_00030)

It is reproduced here with slight alterations.

### 4.1 Abstract

The ability to combine individual concepts of objects, properties, and actions into complex representations of the world is often associated with language. Yet combinatorial event-level representations can also be constructed from nonverbal input, such as visual scenes. Here, we test whether the language network in the human brain is involved in and necessary for semantic processing of events presented nonverbally. In Experiment 1, we scanned participants with fMRI while they performed a semantic plausibility judgment task vs. a difficult perceptual control task on sentences and line drawings that describe/depict simple agent-patient interactions. We found that the language network responded robustly during the semantic task performed on both sentences and pictures (although its response to sentences was stronger). Thus, language regions in healthy adults are engaged during a semantic task performed on pictorial depictions of events. But is this engagement necessary? In Experiment 2, we tested two individuals with global aphasia, who have sustained massive damage to perisylvian

language areas and display severe language difficulties, against a group of age-matched control participants. Individuals with aphasia were severely impaired on the task of matching sentences to pictures. However, they performed close to controls in assessing the plausibility of pictorial depictions of agent-patient interactions. Overall, our results indicate that the left fronto-temporal language network is recruited but not necessary for semantic processing of nonverbally presented events.

## 4.2 Introduction

Many thinkers have argued for an intimate relationship between language and thought, in fields as diverse as philosophy (Carruthers, 2002; Davidson, 1975; Wittgenstein, 1961), psychology (Sokolov, 1972; Vygotsky, 1934; Watson, 1920), linguistics (Berwick & Chomsky, 2016; Bickerton, 1990; Chomsky, 2007; Hinzen, 2013; Jackendoff, 1996), and artificial intelligence (T. B. Brown et al., 2020; Goldstein & Papert, 1977; Turing, 1950; Winograd, 1976). According to such accounts, language enables us to access our vast knowledge of objects, properties, and actions — often referred to as semantic knowledge — and flexibly combine individual semantic units to produce complex situation-specific representations called ‘thoughts’. The hypothesis that language is critical for thought crucially depends on whether or not language is essential for combinatorial semantic processing: if we can access and combine individual concepts in the absence of language, that would constitute evidence against the necessity of language in forming novel thoughts. Here, we test the link between language and thought by examining the role of the language network in a nonverbal combinatorial semantic task.

Recent evidence from neuroscience suggests that language processing is largely distinct from other aspects of cognition (Fedorenko & Blank, 2020; Fedorenko & Varley, 2016). A network of left-lateralized frontal and temporal brain regions (here referred to as ‘the language network’) has been found to respond to written/spoken/signed words and sentences, but not to mental arithmetic, music perception, executive function tasks, action/gesture perception, or computer programming (Amalric & Dehaene, 2019; X. Chen et al., 2021; Fedorenko, Bers, et al., 2011; Ivanova et al., 2020; Jouravlev et al., 2019; Liu et al., 2020; MacSweeney et al., 2002; Monti et al., 2009, 2012; Pritchett et al., 2018).

Similarly, investigations of patients with profound disruption of language capacity (global aphasia) have shown that some of these individuals can solve arithmetic and logic problems, appreciate and create music, and think about others' thoughts in spite of their language impairment (Basso & Capitani, 1985; Luria et al., 1965; Varley et al., 2005; Varley & Siegal, 2000), providing converging evidence that language is subserved by domain-specific cognitive mechanisms.

Despite this significant progress in dissociating linguistic and non-linguistic processing, the role of the language network in nonverbal semantics remains unclear. Semantics is often considered to be an integral part of linguistic processing (Altshuler et al., 2019; Binder et al., 2009; Fillmore, 2006; Milberg & Blumstein, 1981; Pinker & Levin, 1991; Talmy, 2000): each content word is linked to an underlying semantic representation ("lexical semantics"), which then combine to form phrase- and sentence-level meanings ("combinatorial semantics"). This tight integration between language and semantics suggests that the frontal and temporal language regions may play an important role in storing and processing semantic information (see Hasson et al., 2015 for general arguments against the separation of storage and processing/computation in the brain). However, many semantic representations can also be activated by nonverbal input (e.g., the concept CAT can be evoked not only by the word "cat", but also by a picture / sight of a cat), suggesting that language does not necessarily have a privileged role in semantic processing. In this work, we ask whether the fronto-temporal language network supports semantic processing for both verbal and nonverbal stimuli or whether it is only engaged in the semantic processing of *verbal* input.

A large body of work has aimed to address the role of the language network in nonverbal semantics; however, different sources of evidence have produced conflicting results. Neuroimaging studies that explicitly compared verbal and nonverbal semantic processing of objects (e.g., Devereux et al., 2013; Fairhall & Caramazza, 2013; Handjaras et al., 2017; Shinkareva et al., 2011; Vandenberghe et al., 1996; Visser et al., 2012), actions (e.g., Wurm & Caramazza, 2019), and events (Baldassano et al., 2018; Z. Hu et al., 2019; Jouen et al., 2015; Thierry & Price, 2006) often reported overlapping activation in left-lateralized frontal and temporal areas, which may reflect the engagement of the language

network. In contrast, neuropsychology studies have often reported dissociations between linguistic and semantic deficits in patients with aphasia (e.g., Antonucci & Reilly, 2008; Chertkow et al., 1997; Dickey & Warren, 2015; Jefferies & Lambon Ralph, 2006; Saygin et al., 2004; cf. Saygin et al., 2003), suggesting that verbal and nonverbal semantic processes rely on distinct neural circuits. Both groups of studies have limitations that make it difficult to reconcile their findings. The neuroimaging studies have typically relied on group analyses — an approach known to overestimate overlap in cases of nearby functionally distinct areas (Nieto-Castañón & Fedorenko, 2012) — and/or do not report effect sizes, which are critical for interpreting the functional profiles of the regions in question (a region that responds similarly strongly to verbal and nonverbal semantic tasks plausibly supports computations that are different from a region that responds to both, but shows a 2-3 times stronger response to verbal semantics; see, e.g., G. Chen et al. (2017) for discussion). Meanwhile, the aphasia studies have typically investigated cases where only some of the language regions were damaged, leaving open the possibility that the intact portions of the language network were still contributing to nonverbal semantic processing. Further, neuroimaging and aphasia studies typically rely on different experimental paradigms, making it challenging to directly compare their results.

It should also be noted that few neuropsychological studies (with the exception of Dresang et al., 2019; Marshall et al., 1993) have investigated the processing of verbal and nonverbal events (as opposed to individual objects or actions). Constructing event-level mental representations requires object and action processing but is not reducible to them (Dresang et al., 2019) and therefore may engage additional cognitive operations. In particular, to understand an event, we must identify *relations* between participating entities and assign them *thematic roles* (Estes et al., 2011). This process of identifying who did what to whom has traditionally been considered a hallmark of the language system (Fillmore, 1968; Gruber, 1965). Thus, if any aspect of semantic processing requires language, event understanding would seem to be one of the strongest candidates.

Event processing has perhaps been most extensively investigated in EEG research, where a number of studies have reported that semantic violations in visually presented scenes/events evoke the N400 response, a marker of semantic processing (Coco et al.,

2020; Cohn, 2020; Jouen et al., 2019; Proverbio & Riva, 2009; Sitnikova et al., 2008; Võ & Wolfe, 2013; West & Holcomb, 2002; see Kutas & Federmeier, 2011, for a review), similarly to semantic violations in sentences, where the N400 component was originally discovered (Kutas & Hillyard, 1980). The EEG results have been taken to suggest that linguistic and visual semantic processing rely on a shared mechanism. However, because the neural generators of the N400 remain debated (Lau et al., 2008, 2016; Matsumoto et al., 2005; Zhu et al., 2019), this evidence does not definitively demonstrate the involvement of the language network in visual event processing.

Here, we synergistically combine neuroimaging and neuropsychological evidence to ask whether the language network is engaged during and/or necessary for nonverbal event semantics. We focus on the understanding of agent-patient relations (“who did what to whom”) in visually presented scenes. Identification of thematic relations is critical to understanding and generating sentences (Carlson & Tanenhaus, 1988; Fillmore, 2002; Jackendoff, 1987), but “agent” and “patient” are not exclusively linguistic notions: they likely constitute part of humans’ ‘core knowledge’ (Rissman & Majid, 2019; Spelke & Kinzler, 2007; Strickland, 2017; L. Wagner & Lakusta, 2009) and are integral to visual event processing (Cohn & Paczynski, 2013; Hafri et al., 2018). Investigating the role of the language network in processing agent-patient relations therefore constitutes an important test of the relationship between language and combinatorial event semantics.

We used two kinds of evidence in our study: (1) fMRI in neurotypical participants, and (2) behavioral data from two individuals with global aphasia and a group of age-matched healthy controls. All participants were asked to evaluate the plausibility of events, presented either as sentences (neurotypicals only) or pictures. To ensure that participants could not rely on low-level visual cues when evaluating picture plausibility, we used line drawings rather than photographs. The line drawings were highly controlled: each picture pair depicted two animate participants engaged in a certain interaction, but the participants’ roles in this interaction were either plausible (e.g., a cop arresting a criminal) or implausible (e.g., a criminal arresting a cop). This manipulation allowed us to ensure that participants could not infer picture plausibility based solely on the attributes of a single participant; rather, they had to evaluate the event as a whole.

To foreshadow our results, we find that language-responsive brain areas in neurotypical participants respond during the plausibility task for both sentences and pictures (although the responses are lower for pictures). However, individuals with global aphasia, who sustained severe damage to language areas, perform well on the picture plausibility task, suggesting that the language network is not required for constructing combinatorial representations of visually depicted events.

## 4.3 Method

### 4.3.1 Experiment 1: Is the language network active during a nonverbal event semantics task?

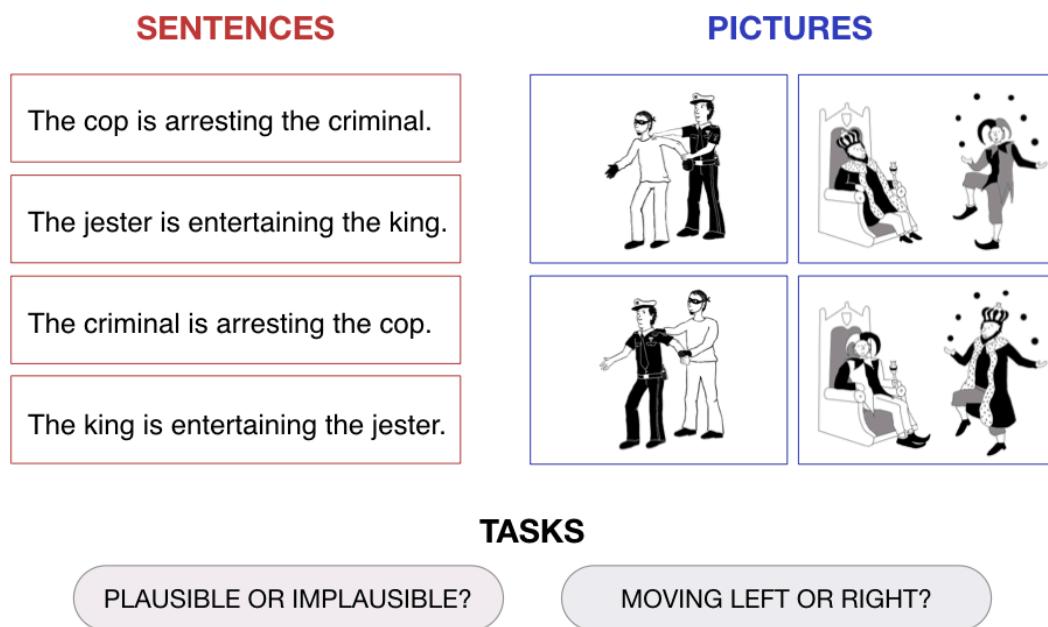
#### 4.3.1.1 Overview

In the first experiment, we presented neurotypical participants with sentences and pictures describing/depicting agent-patient interactions that were either plausible or implausible (**Figure 4-1**), while the participants were undergoing an fMRI scan. Participants performed a semantic judgment task on the sentences and pictures, as well as a difficulty-matched low-level perceptual control task on the same stimuli, in a 2x2 blocked design. In separate blocks, participants were instructed to indicate either i) whether the stimulus was plausible or implausible (the semantic task) or ii) whether the stimulus was moving to the left or right (the perceptual task). The language regions in each participant was identified using a separate functional language localizer task (sentences > nonwords contrast; Fedorenko et al., 2010). We then measured the response of those regions to sentences and pictures during the semantic and perceptual tasks.

#### 4.3.1.2 Participants

Twenty-four participants took part in the fMRI experiment (11 female, mean age = 25 years, SD = 5.2). The participants were recruited from MIT and the surrounding Cambridge/Boston, MA, community and paid for their participation. All were native

speakers of English, had normal hearing and vision, and no history of language impairment. All were right-handed (as assessed by Oldfield's (1971) handedness questionnaire, or self-report). Two participants had low behavioral accuracy scores (<60%), and one had right-lateralized language regions (as evaluated by the language localizer task; see below); they were excluded from the analyses, which were therefore based on data from 21 participants. The protocol for the study was approved by MIT's Committee on the Use of Humans as Experimental Subjects (COUHES). All participants gave written informed consent in accordance with protocol requirements.



**Figure 4-1.** Sample stimuli used in the experiment. For both sentences and pictures, participants were required to perform either a semantic plausibility judgment task (“Plausible or implausible?”) or a control perceptual task (“Moving left or right?”). The full set of materials is available at <https://osf.io/gsudr/>.

#### 4.3.1.3 Design, materials, and procedure

All participants completed a language localizer task aimed at identifying language-responsive brain regions (Fedorenko et al., 2010) and the critical picture/sentence plausibility task.

The language localizer task is described in Section 2.5.2.3.

The picture plausibility task included two types of stimuli: (1) black-and-white line drawings depicting plausible and implausible agent-patient interactions (created by an artist for this study), and (2) simple sentences describing the same interactions. Sample stimuli are shown in **Figure 4-1**, and a full list of materials is available on the project's website (<https://osf.io/gsudr/>). Forty plausible-implausible pairs of pictures, and forty plausible-implausible pairs of corresponding sentences were used. The full set of materials was divided into two lists, such that List 1 used plausible pictures and implausible sentences for odd-numbered items, and implausible pictures and plausible sentences for even-numbered items, and List 2 did the opposite. Thus, each list contained either a picture or a sentence version of any given event. Stimuli were presented in a blocked design (each block included either pictures or sentences) and were moving either to the right or to the left for the duration of stimulus presentation. At the beginning of each block, participants were told which task they would have to perform next: semantic or perceptual. The semantic task required them to indicate whether the depicted / described event is plausible or implausible by pressing one of two buttons. The perceptual task required them to indicate the direction of stimulus movement (right or left). To ensure that participants always perform the right task, a reminder about the task and the response buttons ("plausible=1/implausible=2", or "moving right=1/left=2") was visible in the lower right-hand corner of the screen for the duration of the block. Each stimulus (a picture or a sentence) was presented for 1.5 s, with 0.5 s intervals between stimuli. Each block began with a 2-second instruction screen to indicate the task, and consisted of 10 trials, for a total duration of 22 s. Trials were presented with a constraint that the same response (plausible/implausible in the semantic condition, or right/left in the perceptual condition) did not occur more than 3 times in a row. Each run consisted of 3 fixation blocks and 8 experimental blocks (2 per condition: semantic task – pictures, semantic task – sentences, perceptual task – pictures, perceptual task - sentences) and lasted 242 s (4 min 2 s). The order of conditions was palindromic and varied across runs and participants. Each participant completed 2 runs.

#### **4.3.1.4 fMRI data acquisition**

The data acquisition procedure was the same as that described in Section 2.5.3.

#### **4.3.1.5 fMRI data preprocessing**

The preprocessing procedure was the same as that described in Section 2.5.4.

#### **4.3.1.6 First-level analysis**

First-level analysis procedure was the same as that described in Section 2.5.5. For the language localizer, we modeled entire blocks (sentence reading and nonword reading). For the picture plausibility task, we modeled entire blocks (sentences, semantic task; pictures, semantic task; sentences, perceptual task; pictures, perceptual task) with the exception of the instruction reading phase, which was modeled as a separate condition.

#### **4.3.1.7 Defining functional regions of interest (fROIs)**

The procedure for defining language parcels and fROIs is described in Section 2.5.6. Note that the multiple demand network was not considered in this study.

#### **4.3.1.8 Examining the functional response profiles of the language fROIs**

fROI response estimation procedure was the same as that described in Section 3.3.7.

#### **4.3.1.9 Statistical analyses**

To analyze fROI responses, we ran a linear mixed-effect regression model with two fixed effects (stimulus type and task) and two random effects (participant and fROI). We used sum coding for both stimulus type and task. Planned follow-up comparisons examined response to sentences and pictures during the semantic task within each fROI; the results were FDR-corrected (Benjamini & Hochberg, 1995) for the number of regions. The formula used for the main mixed linear effects model was  $EffectSize \sim StimType * Task +$

$(1 | fROI) + (1 | Participant)$ . The formula used for the follow-up comparisons was  $EffectSize \sim StimType * Task + (1 | Participant)$ . The analysis was run using the *lmer* function from the *lme4* R package (Bates et al., 2015); statistical significance of the effects was evaluated using the *lmerTest* package (Kuznetsova et al., 2017).

To analyze differences in response times (RT) and accuracy across conditions, we ran a linear (for RT) and logistic (for accuracy) mixed effect regression models that aimed to mirror the structure of the mixed effect models in the neuroimaging analyses. Specifically, the behavioral models used task and stimulus type as fixed effects (with sum contrast coding) and participant and item as random intercepts. The formulae were  $Accuracy/RT \sim StimType * Task + (1 | fROI) + (1 | Participant)$ .

### **4.3.2 Experiment 2: Is the language network required for a nonverbal event semantics task?**

#### **4.3.2.1 Overview**

In the second experiment, we examined two individuals with global aphasia, a disorder characterized by severe linguistic impairments, together with a group of age-matched controls. The participants performed two critical tasks: the picture plausibility judgment task (identical to the “picture, semantic” condition from Experiment 1) and the sentence-picture matching task based on the same set of pictures.

#### **4.3.2.2 Participants**

Two participants with global aphasia, S.A. and P.R., took part in the study. Both had large lesions that had damaged the left inferior frontal gyrus, the inferior parietal lobe (supramarginal and angular gyri) and the superior temporal lobe. At the time of testing, they were 68 and 70 years old respectively. S.A. was 22 years 5 months post-onset of his neurological condition, and P.R. was 14 years 7 months post-onset. S.A. had a subdural empyema in the left sylvian fissure, with associated meningitis that led to a secondary vascular lesion in left middle cerebral artery territory. P.R. also had a vascular lesion in left middle cerebral artery territory.

Both participants were male, native English speakers, and did not present with visual impairments. S.A. was pre-morbidly right-handed; P.R. was pre-morbidly left-handed, but a left hemisphere lesion that resulted in profound aphasia indicated that he, like most left-handers, was left-hemisphere dominant for language (Pujol et al., 1999). Both individuals were classified as severely agrammatic (**Table 4-2**), but their non-linguistic cognitive skills were mostly spared (**Table 4-3**). They performed the semantic task and the sentence-picture matching task with a 7-months period between the two.

We also tested two sets of neurotypical control participants, one for the semantic task and one for the language task. The semantic task control participants were 12 healthy participants (7 females) ranging in age from 58 to 78 years (mean age 65.5 years). The language task control participants were 12 healthy participants (5 females) ranging in age from 58 to 78 years (mean age 64.7 years). None of the healthy participants had a history of speech or language disorders, neurological diseases or reading impairments. All were native English speakers, and had normal, or corrected-to-normal, vision.

Participants undertook the experiments individually, in a quiet room. An experimenter was present throughout the testing session. The stimuli were presented on an Acer Extensa 5630G laptop, with the experiment built using DMDX (Forster & Forster, 2003). Ethics approval was granted by the UCL Research Ethics Committee (LC/2013/05). All participants provided informed consent prior to taking part in the study.

#### 4.3.2.3 Semantic Task: Picture plausibility judgments

The same picture stimuli were used as those in Experiment 1 (see **Figure 4-1**), plus one additional plausible-implausible pair of pictures (which was omitted from the fMRI experiment to have a total number of stimuli be divisible by four, for the purposes of grouping materials into blocks and runs), for a total of 82 pictures (41 plausible-implausible pairs). Four of the 82 pictures were used as training items (see below).

The stimuli were divided into 2 sets, with an equal number of plausible and implausible pictures; each plausible-implausible pair was split across the 2 sets, to minimize repetition of the same event participants within a set. The order of the trials was randomized within each set, so that each participant saw the pictures in a different sequence. A self-timed break was placed between the two sets.

Prior to the experiment, participants were shown two pairs of pictures, which acted as training items. The pairs consisted of one plausible and one implausible event. They were given clear instructions to focus on the relationship between the two characters and assess whether they thought the interaction was plausible, in adherence with normal expectations, or implausible, at odds with expectations. They were asked to press a green tick (the left button on the mouse) if they thought the picture depicted a plausible event, and a red cross (the right button on the mouse) if they thought the picture depicted an implausible event. They were asked to do so as quickly and accurately as possible. The pictures appeared for a maximum of 8 seconds, with the inter-stimulus interval of 2 seconds. Accuracies and reaction times were recorded. Participants had the opportunity to ask any questions, and the instructions for participants with aphasia were supplemented by gestures to aid comprehension of the task. Participants had to indicate that they understood the task prior to starting.

#### **4.3.2.4 Language Task: Sentence-picture matching**

The same 82 pictures were used as in the plausibility judgment experiment. In this task, a sentence was presented below each picture that either described the picture correctly (e.g., “the cop is arresting a criminal” for the first sample picture in **Figure 4-1**) or had the agent and patient switched (“the criminal is arresting the cop”). Simple active subject-verb-object sentences were used. Combining each picture with a matching and a mismatching sentence resulted in 164 trials in total.

For the control participants, the trials were split into two sets of 82, with an equal number of plausible and implausible pictures, as well as an equal number of matches and mismatches in each set. In order to avoid tiring the participants with aphasia, the

experiment was administered across two testing sessions each consisting of two sets of 41 stimuli and occurring within the same week. For both groups, the order of the trials was randomized separately for each participant, and no pictures from the same pair (e.g., an event involving a cop and a criminal) appeared in a row. A self-timed break was placed between the two sets.

Prior to the experiment, participants were told that they would see a series of pictures with accompanying sentences, and their task was to decide whether the sentence matched the depicted event. They were asked to press a green tick (the left button on the mouse) if they thought the sentence matched the picture, and a red cross (the right button on the mouse) if they thought the sentence did not match the picture. They were asked to do so as quickly and accurately as possible. The picture / sentence combinations appeared for a maximum of 25 seconds, with the inter-stimulus interval of 2 seconds. Accuracies and reaction times were recorded. As in the critical task, participants had the opportunity to ask any questions, and the instructions for participants with aphasia were supplemented by gestures.

#### **4.3.2.5 Data analysis**

We used exact binomial test to test whether patients' performance on either task was significantly above chance, as well as the Crawford-Howell (1998) test for dissociation to compare patient performance relative to controls across the two tasks. We excluded all items with reaction times and/or accuracies outside 3 standard deviations of the control group mean (4 items for the semantic task and 11 items for the sentence-picture matching task).

#### **4.3.2.6 Estimating the damage to the language network in patients with aphasia**

To visualize the extent of the damage to the language network, we combined available structural MRI of one patient with aphasia (P.R.) with a probabilistic activation overlap map of the language network. The map was created by overlaying thresholded individual activation maps for the language localizer contrast (sentences > nonwords, as

described in Experiment 1) in 220 healthy participants. The maps were thresholded at the  $p < 0.001$  whole-brain uncorrected level, binarized, and overlaid in the common space, so that each voxel contains information on the proportion of participants showing a significant language localizer effect (see Woolgar, Duncan, Manes & Fedorenko (2018b) for more details). The map can be downloaded from the project's website (<https://osf.io/gsudr/>).

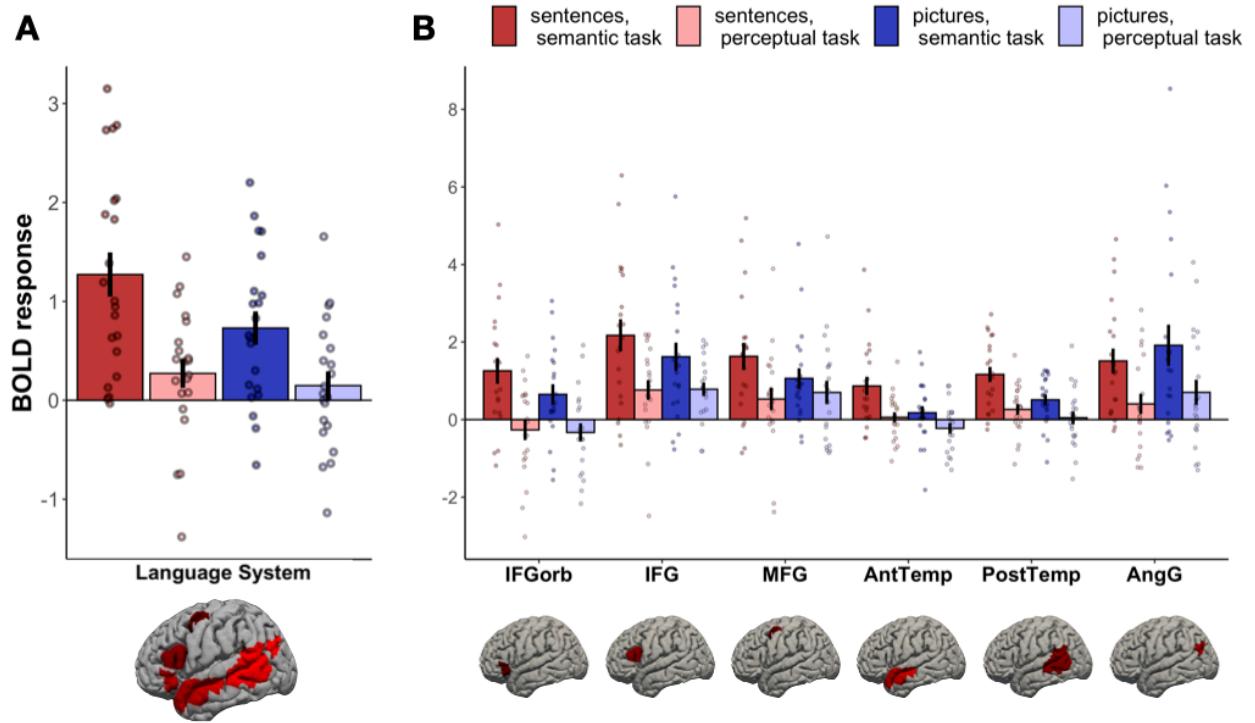
## 4.4 Results

### 4.4.1 Experiment 1: Is the language network active during a nonverbal event semantics task?

#### 4.4.1.1 Behavioral Results

All participants were engaged during the task: the overall response rate was 91.7% (sentence semantic - 89.9%, sentence perceptual - 91.6%, picture semantic - 93.6%, picture perceptual - 91.9%). Average response times were 1.27 s ( $SD = 0.46$ ) for the semantic sentence task, 1.16 s ( $SD = 0.38$ ) for the perceptual sentence task, 1.22 s ( $SD = 0.35$ ) for the semantic picture task, and 1.19 ( $SD = 0.36$ ) for the perceptual picture task. A linear mixed effect model with *task* and *stimulus type* as fixed effects and *participant* and *item number* as random intercepts showed a small main effect of task (semantic > perceptual;  $\beta = .06$ ,  $p < .001$ ), no main effect of stimulus type ( $\beta = 0.02$ ,  $p = 0.287$ ), and no interaction between task and stimulus type ( $\beta = .03$ ,  $p = 0.359$ ).

Average accuracies were 0.81 for the semantic sentence task, 0.79 for the perceptual sentence task, 0.75 for the semantic picture task, and 0.75 for the perceptual picture task. A logistic mixed effect model with the same structure as the linear RT model above showed no significant effects of either task ( $\beta = .09$ ,  $p = 0.198$ ) or stimulus type ( $\beta = .12$ ,  $p = 0.101$ ), and no interaction between them ( $\beta = .04$ ,  $p = 0.759$ ). Due to a technical error, accuracy data for 14 participants were only recorded for one of the two runs.



**Figure 4-2.** BOLD response during the four experimental conditions within (A) the language network as a whole and (B) each of the six language fROIs. The fROI labels correspond to approximate anatomical locations: IFGorb – the orbital portion of the left inferior frontal gyrus; IFG – left inferior frontal gyrus; MFG – left middle frontal gyrus; AntTemp – left anterior temporal cortex; PostTemp – left posterior temporal cortex; AngG – left angular gyrus. Within each parcel, the responses to the critical experiment conditions are extracted from the top 10% most language-responsive voxels (selected in each of the 21 individuals separately). Error bars indicate standard error of the mean across participants; dots indicate individual participants' responses.

#### 4.4.1.2 Neuroimaging Results

Although diverse non-linguistic tasks have been previously shown not to engage the language network (Fedorenko & Varley, 2016), we here found that the language regions responded more strongly during the semantic task on both sentences and pictures compared to the perceptual control task (**Figure 4-2**). A linear mixed effect model with *task* and *stimulus type* as fixed effects and *participant* and *fROI* as random effect intercepts showed a significant effect of task (semantic > perceptual,  $\beta = 0.93$ ,  $p < .001$ ), stimulus type (sentences > pictures,  $\beta = 0.23$ ,  $p = 0.018$ ), and an interaction between them ( $\beta = .43$ ,  $p = 0.025$ ). These results demonstrate that the language network responds to the semantic

task performed on both sentences and pictures, although this task effect is stronger for sentences.

**Table 4-1.** Regression model terms for fROI-based statistical analyses. The p-values are FDR-corrected for the number of regions (n=6). Significant terms are highlighted in bold. The fROI labels correspond to the approximate anatomical locations: IFGorb – the orbital portion of the left inferior frontal gyrus; IFG – left inferior frontal gyrus; MFG – left middle frontal gyrus; AntTemp – left anterior temporal cortex; PostTemp – left posterior temporal cortex; AngG – left angular gyrus.

ROI	Regression Term	Beta	p-value
IFGorb	Intercept	0.33	0.104
	Stimulus (Sent>Pic)	0.34	0.215
	Task (Sem>Perc)	1.25	<0.001
	Stimulus:Task	0.54	0.283
IFG	Intercept	1.33	<0.001
	Stimulus (Sent>Pic)	0.27	0.259
	Task (Sem>Perc)	1.12	<0.001
	Stimulus:Task	0.57	0.28
MFG	Intercept	0.98	0.002
	Stimulus (Sent>Pic)	0.2	0.259
	Task (Sem>Perc)	0.73	<0.001
	Stimulus:Task	0.74	0.231
AntTemp	Intercept	0.22	0.104
	Stimulus (Sent>Pic)	0.49	0.002
	Task (Sem>Perc)	0.6	<0.001
	Stimulus:Task	0.41	0.24
PostTemp	Intercept	0.5	<0.001
	Stimulus (Sent>Pic)	0.43	0.006
	Task (Sem>Perc)	0.68	<0.001
	Stimulus:Task	0.44	0.24
AngG	Intercept	1.13	0.002
	Stimulus (Sent>Pic)	-0.35	0.215
	Task (Sem>Perc)	1.16	<0.001
	Stimulus:Task	-0.11	0.823

To investigate individual brain regions comprising the language network, we conducted follow-up analyses on individual fROIs' activity (FDR-corrected for the number of regions) (**Figure 4-2, B**). These revealed a significant semantic > perceptual task effect in all fROIs (**Table 4-1**). The sentences > pictures stimulus type effect was observed in two fROIs, located in anterior and posterior left temporal lobe. The interaction between task and stimulus type was not significant in any fROI, although, numerically, responses to sentences during the semantic task were stronger than responses to any other condition

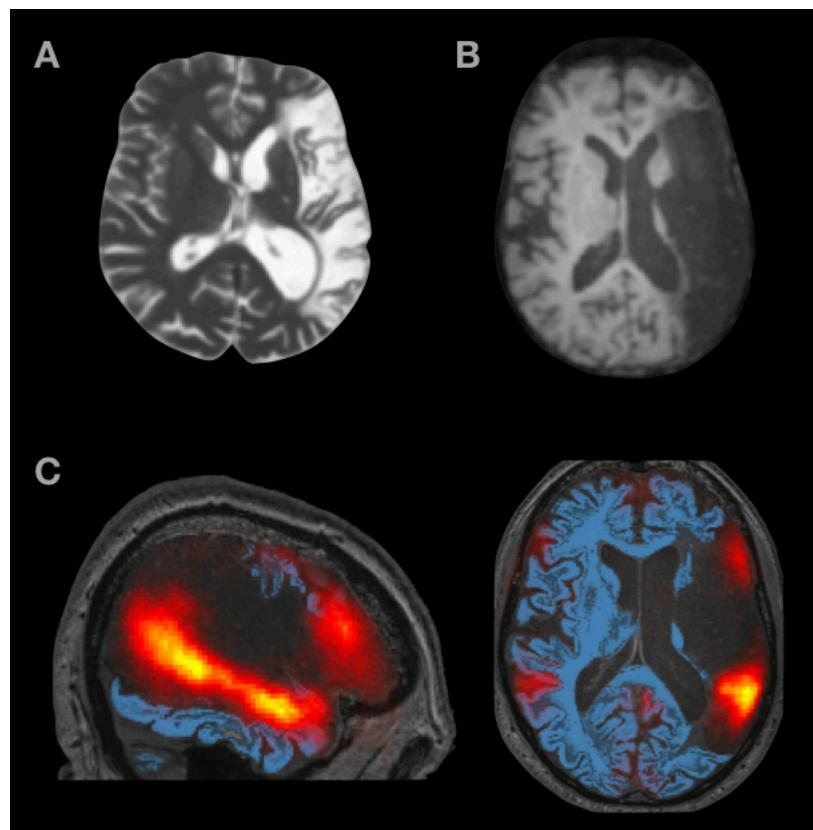
in all except the left angular gyrus fROI. We conclude that sensitivity to the semantic task is a general property of all regions in the language network rather than an effect driven by a subset of regions.

To facilitate the comparison of our results with prior neuroimaging studies, we also performed a random effects whole-brain group analysis (**Supplemental figure 4-1**), which yielded results similar to the fROI-based analyses described above. Specifically, we found that the semantic > perceptual contrast for both sentences and pictures activates left-lateralized frontal and temporal regions that overlap with the language parcels (used to constrain the definition of individual language fROIs). The extent of semantics-evoked activation in the left lateral temporal areas was weaker for pictures than sentences (the opposite was true on the ventral surface of the left temporal lobe). Note, however, that these results should be interpreted with caution, since group analyses might conflate functionally distinct regions that are anatomically close (Nieto-Castañón & Fedorenko, 2012), especially in association cortex, which tends to be functionally heterogeneous (Blank et al., 2017; Braga et al., 2019; Fedorenko & Kanwisher, 2009; Frost & Goebel, 2012; Tahmasebi et al., 2012; Vázquez-Rodríguez et al., 2019).

Overall, the first experiment revealed that the language network is strongly and significantly recruited for semantic processing of events presented not only verbally (through sentences), but also nonverbally (through pictures). Specifically, the language network is active when we interpret pictures that depict agent-patient interactions and relate them to stored world knowledge. It is worth noting, however, that responses to the semantic task are stronger for sentences than for pictures (as shown by the interaction between task and stimulus type at the network level; **Figure 4-2, A**), suggesting that the language network may play a less important role in nonverbal semantic processing. To test whether the engagement of the language network is *necessary* for comprehending visually presented events, we turn to behavioral evidence from individuals with global aphasia.

#### 4.4.2 Experiment 2: Is the language network required for a nonverbal event semantics task?

We examined two individuals with global aphasia (S.A. and P.R.). Both had suffered large vascular lesions that resulted in extensive damage to left perisylvian cortex, including the language network (see **Figure 4-3** for lesion images, including a probabilistic map of the language network based on fMRI data from neurotypical participants, overlayed onto P.R.'s MRI).



**Figure 4-3.** Structural MRI images from (A) S.A. and (B) P.R. (C) Probabilistic language activation overlap map overlaid on top of P.R.'s structural MRI image. The heatmap values range from 0.01 (red) to 0.5 (yellow) and correspond to proportions of individuals (in a set of  $n=220$ ) that show a significant language localizer (sentences > nonwords) effect in that voxel. As can be seen, the lesion covers most left-hemisphere areas with voxels that likely belong to the language network.

Both individuals were severely agrammatic (**Table 4-2**). Whereas they had some residual lexical comprehension ability, scoring well on tasks involving word-picture matching

and synonym matching across spoken and written modalities, their lexical production was impaired. Both failed to correctly name a single item in a spoken picture-naming task. S.A. displayed some residual written word production ability, scoring 24 out of 60 in a written picture-naming task. P.R., however, performed poorly in the written task, correctly naming just 2 out of 60 items.

S.A. and P.R.'s syntactic processing was severely disrupted. They scored at or below chance in the reversible spoken and written sentence comprehension tasks (sentence-picture matching), which included active sentences, e.g. "the man kills the lion", and passive sentences, e.g. "the man is killed by the lion". They also scored near chance in written grammaticality judgment assessments. The patients' comprehension performance was impaired regardless of whether the sentences were presented visually or auditorily, indicating that the impairment was linguistic rather than perceptual. To determine whether the sentence comprehension impairments could be explained by working memory deficits, we evaluated the patients' phonological working memory by means of a digit span test (using a recognition paradigm that did not require language production). The patients' working memory span was somewhat reduced: S.A. and P.R. had the scores of 3 and 4 items, respectively, compared to the neurotypical age-matched controls who had an average score of 6.4 ( $SD=0.6$ ; see Zimmerer et al., 2019). However, even such reduced working memory span should have been sufficient for processing the simple subject-verb-object sentences that were used in the syntactic assessments, as well as in the critical task described below. Thus, S.A. and P.R.'s difficulties with linguistic tasks could not be attributed to phonological working memory problems.

Importantly and in line with prior arguments (Fedorenko & Varley, 2016), S.A. and P.R. performed relatively well on nonverbal reasoning tasks, which included measures of fluid intelligence (Raven's Standard/Colored Progressive Matrices; Raven & Raven, 2003), object semantics (Pyramids and Palm Trees test; Howard & Patterson, 1992), and visual working memory (Visual Pattern Test; Della Sala et al., 1999), indicating that the extensive brain damage in these patients did not ubiquitously affect all cognitive abilities (**Table 4-3**). Such a selective impairment of linguistic skills allowed us to examine the causal role of language in nonverbal event semantics.

**Table 4-2.** Results of linguistic assessments for participants with global aphasia.

<b>Lexical Tests</b>	<b>Chance Score</b>	<b>S.A.</b>	<b>P.R.</b>
ADA spoken word picture matching	16.5	60/66*	61/66*
ADA written word picture matching	16.5	62/66*	66/66*
ADA spoken synonym matching	80	123/160*	113/160*
ADA written synonym matching	80	121/160*	145/160*
PALPA 54 spoken picture naming	n/a	0/60	0/60
PALPA 54 written picture naming	n/a	24/60	2/60

<b>Syntactic Tests</b>			
Comprehension of spoken reversible sentences	50	49/100	38/100
Comprehension of written reversible sentences	50	42/100	49/100
Written grammaticality judgments	20	26/40*	21/40

<b>Verbal Working Memory</b>			
PALPA 13 digit span (recognition)	n/a	3 items	4 items

\* indicates above chance performance (p < .05)

The tests were taken from the Action for Dysphasic Adults (ADA) Auditory Comprehension Battery (S. Franklin et al., 1992) and the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA; Kay et al., 1992) or developed for the purpose of the study.

**Table 4-3.** Results of non-linguistic assessments for participants with global aphasia.

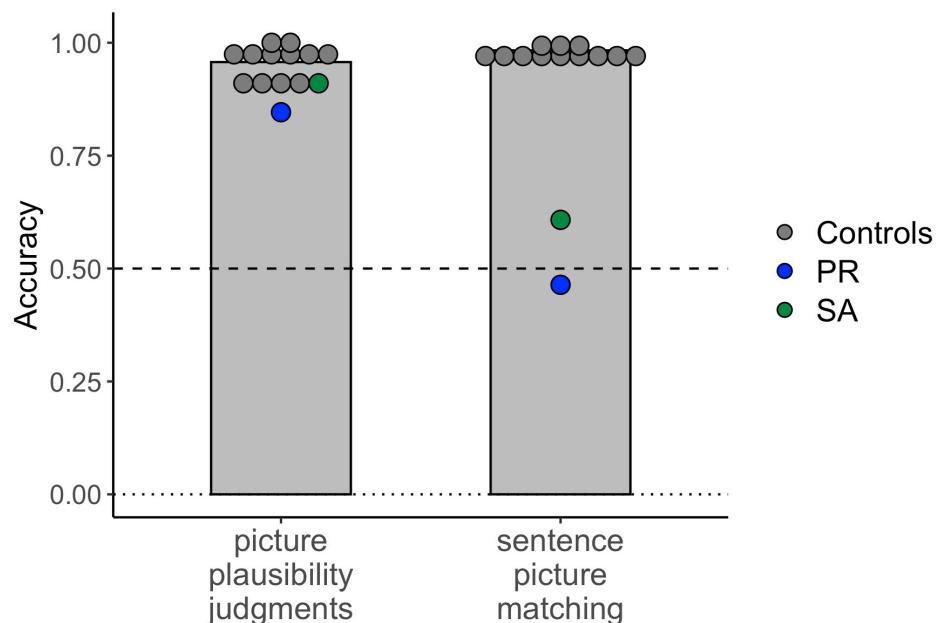
<b>Reasoning Tests</b>	<b>S.A.</b>	<b>P.R.</b>
Raven's Colored Progressive Matrices	36/36	34/36
Raven's Standard Progressive Matrices	53/60	36/60
Pyramids and Palm Trees (3 picture version)	50/52	47/52
Visual Pattern Test	11.5 (90 <sup>th</sup> percentile*)	8.6 (40 <sup>th</sup> percentile*)

\* percentiles are calculated with respect to adults in the same age range with no neurological impairment

To test whether global aphasia affects general event semantics, we measured S.A. and P.R.'s performance on two tasks: (1) the picture plausibility task, identical to the pictures/semantic-task condition from Experiment 1, and (2) a sentence-picture matching task, during which participants saw a picture together with a sentence in which the agent and the patient either matched the picture or were switched ("a cop is arresting a criminal" vs. "a criminal is arresting a cop"); participants had to indicate whether or not the sentence matched the picture. The sentence-picture matching task was similar to the reversible sentence comprehension task in **Table 4-2**, with the exception that the pictures were identical to the pictures from the plausibility task and all sentences used

active voice. For each task, patient performance was compared with the performance of 12 age-matched controls (58-78 years (mean 65.5 years) for the picture plausibility task; 58-78 years (mean 64.7 years) for the sentence-picture matching task).

The results showed a clear difference in performance between the picture plausibility task and the sentence-picture matching task (**Figure 4-4**), despite the fact that both tasks used the same set of pictures. Both individuals with global aphasia and control participants performed well above chance when judging picture plausibility. Neurotypical controls had a mean accuracy of 95.7% ( $SD = 3.8\%$ ). Aphasia patients had mean accuracies of 91.0% (S.A.; 1.2  $SD$  below average) and 84.6% (P.R.; 3.0  $SD$  below average); the exact binomial test showed that performance of both patients was above chance (S.A.,  $p < .001$ , 95% CI [.82, .96]; P.R.,  $p < .001$ , 95% CI [.75, .92]). Although their performance was slightly below the level of the controls, the data indicate that both patients were able to process complex semantic (agent-patient) relations to evaluate the plausibility of depicted events.



**Figure 4-4.** Individuals with profound aphasia perform well on picture plausibility judgment task but fail on the sentence-picture matching task. Patient accuracies are indicated in blue (PR) and green (SA); average controls' performance is shown as gray bars; individual controls' performance (N=12) is shown as gray dots. The dotted line indicates chance performance.

In the sentence-picture matching task, control participants performed close to ceiling, with a mean accuracy of 98.3% ( $SD = 1.1\%$ ). In contrast, both patients were severely impaired: S.A. had a mean accuracy of 60.8% and P.R. had a mean accuracy of 46.4%. Exact binomial test showed that P.R.'s performance was at chance ( $p = .464$ , 95% CI [.38, .55]), while S.A.'s performance was above chance ( $p = .009$ , 95% CI [.53, .69]) but still drastically lower than that of the controls. This result concurs with S.A.'s and P.R.'s poor performance on the reversible sentence comprehension tasks, which had a similar setup but used different materials. However, it stands in stark contrast with the participants' ability to interpret agent-patient interactions in pictures. The Crawford-Howell (1998) *t*-test indicated a significant dissociation between the picture plausibility task and the sentence-picture matching task for both individuals (S.A.,  $t(11) = 18.00$ ,  $p < .001$ ; P.R.,  $t(11) = 24.20$ ,  $p < .001$ ). This dissociation held for both hit rate and false alarm rate (**Supplemental figure 4-2**).

The findings from Experiment 2 demonstrate that, in spite of severe linguistic impairments, individuals with global aphasia are able to access information about event participants depicted in a visual scene, the action taking place between them, the roles they perform in the context of this action, and the real-world plausibility of these roles, indicating that none of these processes require the presence of a functional language network.

## 4.5 Discussion

The relationship between language and thought has been long debated, both in neuroscience (e.g., Binder & Desai, 2011; Bookheimer, 2002; Fedorenko & Varley, 2016; Friederici, 2020) and other fields (e.g., Carruthers, 2002; Hauser et al., 2002; Vygotsky, 1934; Winograd, 1976). Here, we ask whether language-responsive regions of the brain are essential for a core component of thought: processing combinatorial semantic representations. We demonstrate that left-hemisphere language regions are active during the semantic processing of events shown as pictures, although the semantic processing of events shown as sentences elicits a stronger response. We further show that the

language network is not essential for nonverbal event semantics, given that the two individuals with global aphasia, who lack most of their left-hemisphere language network, can still evaluate the plausibility of visually presented events. Our study advances the field in three ways: i) it explores relational semantic processing in the domain of events, moving beyond the semantics of single objects—the focus of most prior neuroscience work on conceptual processing; ii) it evaluates neural overlap between verbal and nonverbal semantics in fMRI at the level of individual participants; and iii) it provides causal evidence in support of a dissociation between language and nonverbal event semantics. In the remainder of the manuscript, we discuss the implications of our results.

#### **4.5.1 The language network is not required for nonverbal event semantics**

Semantic processing of events is a complex, multi-component process. For instance, deciding whether an event is plausible requires one to (1) identify the relevant event participants, (2) determine the action taking place between them, (3) decipher the role that each event participant is performing (in our task, agent vs. patient), and finally, (4) estimate the likelihood that a given participant would be the agent/ patient of the relevant action. Whereas the first three components can, at least in part, be attributed to input-specific processes (e.g., high-level vision), establishing plausibility cannot be solely attributed to perception: in order to decide whether a cop arresting a criminal is more likely than a criminal arresting a cop, participants need to draw on their world knowledge. We demonstrate that this highly abstract process can proceed even when the language network is severely impaired, thus providing strong evidence that a functional language network is not required for nonverbal semantic processing.

The functional dissociation between language-based and vision-based semantic judgments of events accords with the fact that both non-human animals and preverbal infants are capable of complex event processing (Seed & Tomasello, 2010; Spelke, 1976) and that specialized neural mechanisms, distinct from the language network, have been associated with visual understanding of actions (Fang et al., 2016; Häberling et al., 2016; Tarhan & Konkle, 2020) and interactions between animate and/or inanimate entities

(Fischer et al., 2016; Walbrin et al., 2018). These neural mechanisms are either bilateral or right-lateralized, which constitutes further evidence of their dissociation from language, which is typically left-lateralized.

Our results are also consistent with reports of a dissociation between verbal and nonverbal semantic processing of single objects in patients with aphasia (e.g., Antonucci & Reilly, 2008; Bi et al., 2011; Chertkow et al., 1997; Jefferies & Lambon Ralph, 2006; Lambon Ralph et al., 2010) and semantic dementia (e.g., Binney et al., 2016; Gorno-Tempini et al., 2004; Mion et al., 2010; Snowden et al., 2018; Thompson et al., 2003). Those studies typically report that linguistic impairments arise as a result of left hemisphere damage, whereas nonverbal semantic processing deficits are considered to be caused by either bilateral (Lambon Ralph et al., 2017) or right-lateralized lesions (Gainotti, 2011, 2015). Our work contributes to this literature by showing that the language-semantics dissociation holds not only for single concepts but also for combinatorial event-level representations (see also Colvin et al., 2019; Dickey & Warren, 2015). Although we only test two individuals with global aphasia, these data provide an important contribution to the field because of the unique nature of the impairment in these individuals: large-scale disruption of multiple linguistic functions and relatively preserved nonverbal cognition. To test the generalizability of our findings, future work should evaluate a larger sample of individuals with such a dissociation and comprehensively assess both verbal and nonverbal semantic processing of objects, actions, and events.

If language is not essential for event semantics, why is the language network active during a nonverbal event semantics task? It is possible that neurotypical participants partially recode pictorial stimuli into a verbal format (Greene & Fei-Fei, 2014; Trueswell & Papafragou, 2010), which could provide access to linguistic representations as an additional source of task-relevant information (Connell & Lynott, 2013). Indeed, text-based computational models developed in recent years have been shown to successfully perform a wide range of “semantic” tasks, such as inference, paraphrasing, and question answering (T. B. Brown et al., 2020; Devlin et al., 2018, among others). Even simple n-gram models can be used to determine the probability of certain events by, e.g., estimating the probability that the phrase “is arresting” directly follows “cop” vs.

“criminal”. Such language-based semantic information is distinct from non-language-based world knowledge (Clark, 2004; Lucy & Gauthier, 2017), and both kinds of information can be flexibly used depending on task demands (Willits et al., 2015). As a result, it is possible that linguistic resources (housed in the language network) provide an additional source of information when neurotypical individuals determine visual event plausibility. The absence of this additional information source may account for the small decrement in performance observed in participants with aphasia relative to the control participants.

One might speculate that this “language-based” semantic processing route plays a primary role in neurotypical participants, whereas patients with aphasia rely on some alternative route that arose due to the functional reorganization of the brain postinjury. However, we consider this possibility unlikely. Past behavioral evidence from experiments in neurotypical individuals shows that verbal recoding of visual information is relatively slow and can only occur *after* semantic information has been retrieved from the picture (Potter et al., 1986; Potter & Faulconer, 1975). Furthermore, participants do not typically generate covert verbal labels for visually presented objects unless instructed to do so (Dahan et al., 2001; Magnuson et al., 2003; Rehrig et al., 2020; cf. Meyer et al., 2007) or unless the task imposes memory demands (Pontillo et al., 2015). Our stimuli depicted complex two-participant events, making verbal recoding even more effortful than recoding of single objects and, therefore, unlikely to occur during a task that does not require linguistic label generation (Papafragou et al., 2008). Finally, even if individuals with aphasia did rely on a compensatory (e.g., right-hemisphere-mediated) mechanism for semantic processing, it would still indicate that brain mechanisms outside of the core left-hemisphere language network are capable of supporting combinatorial semantics, thus underscoring our claim that language and nonverbal event semantics are neurally dissociable.

Future work should further investigate the nature of the language network’s responses to nonverbal stimuli. Although some studies, like ours, have reported that the left-hemisphere language regions have stronger responses to sentences than content-matched pictures (Amit et al., 2017), others have reported the opposite preference (Jouen et al.,

2015). The divergent result in Jouen et al. (2015) is most likely due to differences in the analytic approach, namely, in the use of ROIs derived from group analyses as opposed to functionally defined fROIs. Task demands could also contribute to the difference in results: Jouen et al. used a one-back memory task (no condition-specific behavioral results reported), whereas we used a plausibility judgment task that had similar accuracies and reaction times between the sentences and pictures. The fact that we found an interaction between input type (sentences vs. pictures) and task also indicates that task effects on activity in the language network merit additional investigation (although see Cheung et al., 2020, for evidence that task demands often have little effect on the responses of the language regions to verbal stimuli). The task effects observed in our study cannot be explained by task difficulty: the participants' accuracies for the semantic vs. perceptual task were not significantly different; the reaction times were slightly faster for the perceptual task, but the effect size was small (0.06s, with average trial RT=1.21s) and therefore unlikely to fully account for the neural effect. Moreover, the language network is not generally driven by task difficulty (Diacheck et al., 2020) and shows strong, consistent responses even in the absence of task (Baldassano et al., 2018; Brennan et al., 2016; Huth et al., 2016; Scott et al., 2017; Shain et al., 2020; Wehbe et al., 2014, among others). Thus, future work needs to explore the effects of task content rather than task difficulty per se.

#### **4.5.2 Implications for theories of semantics in the brain**

In this paper, we focused on the *role of the language network* in nonverbal event semantics, not on the question of which cognitive and neural mechanisms support modality-invariant event processing (we report those analyses in other work; Ivanova et al., in prep.). Nonetheless, current results also bear on general theories of semantic processing in the mind and brain.

Many current theories of semantics highlight broad anatomical areas implicated in linguistic processing as putative semantic hubs. Those include left angular gyrus (e.g., Binder & Desai, 2011), left inferior frontal cortex (e.g., Hagoort & van Berkum, 2007), and the anterior temporal lobes (e.g., Patterson et al., 2007). However, the areas in question

are large patches of cortex that are structurally and functionally heterogeneous: as a result, simply because a visual-semantic study reports activation within the ‘LIFG’, or the ‘angular gyrus’, does not mean that the *language-responsive* portions of those broad areas are at play (see, e.g., Fedorenko & Blank, 2020 for discussion).

In the current study, the language-responsive fROIs that we defined within left angular gyrus, left inferior frontal cortex, and left anterior temporal lobe all responded more strongly during the semantic task than during the perceptual task, for both sentences and pictures. Although this pattern is consistent with evidence of their general involvement in semantic processing, it goes against some of the specific claims made in the literature. For example, our results are inconsistent with the claim that the angular gyrus is the *primary* region involved in event semantics (Binder & Desai, 2011; cf. Williams et al., 2017) given that other regions show a similar functional response profile. That said, the fROI in the angular gyrus was the only one that showed numerically stronger responses to pictures than to sentences, consistent with evidence of its involvement in processing (at least some) semantically meaningful nonverbal stimuli (Amit et al., 2017; Baldassano et al., 2017; Fairhall & Caramazza, 2013; Handjaras et al., 2017; Pritchett et al., 2018). Our results also provide some evidence that a portion of the left anterior temporal lobe (ATL) is engaged in processing event-level representations in verbal stimuli (Jackson et al., 2015; Teige et al., 2019; cf. Lewis et al., 2015; Schwartz et al., 2011; Xu et al., 2018, who claim that the ATL is involved in retrieving property-level but not event-level information). Finally, we observed that the ATL language fROI responded more strongly to sentences than to pictures, which might speak against its role as an amodal semantic hub. Note, however, that this fROI encompasses only a small fraction of left ATL; it therefore remains possible that some other parts of the ATL — especially its ventral/ventromedial portions — have a modality-invariant response profile (Lambon Ralph et al., 2017; Visser et al., 2012).

In addition, our findings contribute to the body of work on the neural representation of agent/patient relationships. Previous experiments attempting to localize brain regions that support thematic role processing have attributed the processing of agent/patient relations to the left hemisphere. Frankland and Greene (2015, 2020) used sentence stimuli

to isolate distinct areas in left superior temporal sulcus (STS) that are sensitive to the identity of the agent vs. the patient. Wang et al. (2016) found that the same (or nearby) STS regions also contained information about thematic roles in videos depicting agent-patient interactions. However, the latter study identified a number of other regions that were sensitive to thematic role information, including clusters in the right posterior middle temporal gyrus and right angular gyrus, suggesting that left STS is not the only region implicated in thematic role processing. A similar distributed pattern was also reported in a neuropsychological study (Wu et al., 2007), which found that lesions to mid-STS led to difficulties in extracting thematic role information from both sentences and pictures; however, deficits in visual agent-patient processing were additionally associated with lesions in anterior superior temporal gyrus, supramarginal gyrus, and inferior frontal cortex, which casts further doubt on the unique role of left STS in agent-patient relation processing. In sum, the evidence to date suggests that parts of the left STS may play a role in processing linguistic information, including thematic relations (Frankland & Greene, 2015, 2020) and verb argument structure (Elli et al., 2019; Williams et al., 2017), but additional brain regions support the processing of event participant roles in nonverbal stimuli.

Finally, our results are generally consistent with a distributed view of semantic representations (McClelland & Rogers, 2003; Tyler & Moss, 2001). Multiple recent studies found that semantic information is not uniquely localized to any given brain region but rather distributed across the cortex (e.g., Anderson et al., 2017; Huth et al., 2016; Pereira et al., 2018; X. Wang et al., 2018). Distributing information across a network of regions in both left and right hemispheres enables the information to be preserved in case of brain damage (Schapiro et al., 2013), which would explain why patients with global aphasia preserve the ability to interpret visually presented events. That said, the findings reported here do not speak to the question of whether such representations rely primarily on sensorimotor areas (Barsalou, 2008; Pulvermuller, 1999) or on associative areas (Mahon, 2015; Mahon & Caramazza, 2008).

#### **4.5.3 Implications for neuroimaging studies of amodal semantics**

The non-causal nature of the language network activation during a nonverbal semantic task has important implications for the study of amodal/multimodal concept representations. A significant body of work has aimed to isolate “amodal” representations of concepts by investigating the overlap between regions active during the viewing of verbal and nonverbal stimuli (Bright et al., 2004; Devereux et al., 2013; Fairhall & Caramazza, 2013; Handjaras et al., 2017; Sevostianov et al., 2002; Thierry & Price, 2006; Vandenberghe et al., 1996; Visser et al., 2012; A. D. Wagner et al., 1997). Most of these overlap-based studies have attributed semantic processing to frontal, temporal, and/or parietal regions within the left hemisphere. Our work, however, demonstrates that, even though meaningful linguistic and visual stimuli evoke overlapping activity in left-lateralized frontal and temporal regions, conceptual information about events persists even when most of these regions are damaged. Thus, overlapping areas of activation for verbal and nonverbal semantic tasks observed in brain imaging studies do not necessarily play a causal role in amodal event semantics.

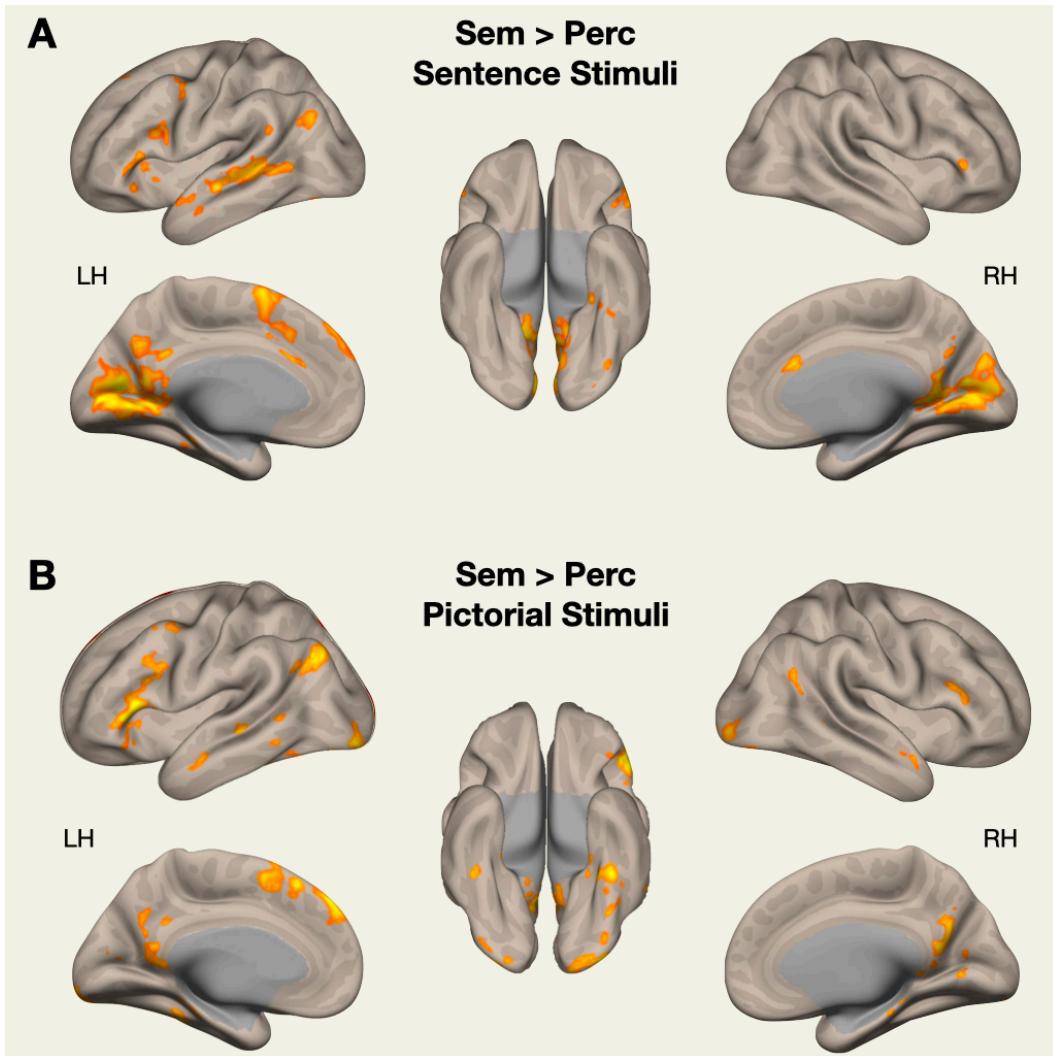
Overall, our study emphasizes the importance of investigating combinatorial semantic processing using both verbal and nonverbal stimuli. Our results show that semantic processing of visually presented events does not require the language network, drawing a sharp distinction between language and nonverbal event semantics and highlighting the necessity to characterize the relationship between them in greater detail using a combination of brain imaging and patient evidence.

#### **4.6 Acknowledgments**

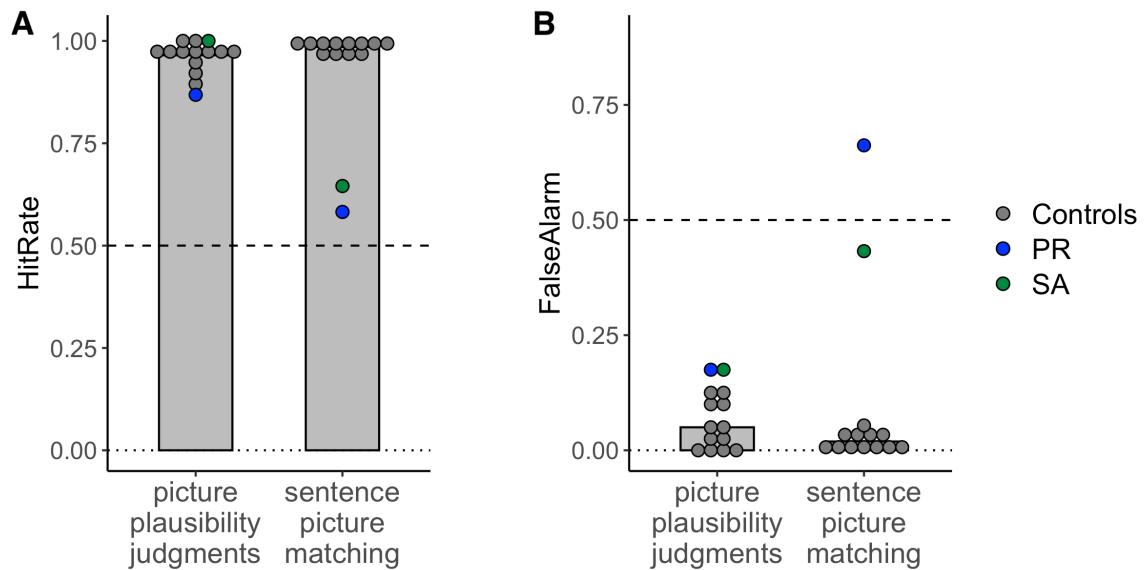
We would like to acknowledge the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT, and its support team (Steve Shannon and Atsushi Takahashi). We thank Birgit Zimmerer for creating the picture stimuli used in both experiments, Chloe Bustin for norming the stimuli, Lily Jordan for help with the behavioral piloting of the fMRI experiment, and EvLab members for their help with fMRI

data collection. E.F. was supported by NIH awards R00-HD057522, R01-DC016607, and R01-DC016950, by a grant from the Simons Foundation to the Simons Center for the Social Brain at MIT, and by funds from BCS and the McGovern Institute for Brain Research at MIT. R.V. was supported by Arts and Humanities Research Council and Alzheimer's Society awards.

## 4.7 Appendix



**Supplemental figure 4-1.** Whole-brain random effects group analysis (Holmes&Friston, 1998) for Semantic > Perceptual task contrast, conducted separately on the Experiment 1 data from sentence trials (a) and picture trials (b). The analysis was conducted using the spm\_ss toolbox (available at [http://www.nitrc.org/projects/spm\\_ss](http://www.nitrc.org/projects/spm_ss)), which interfaces with SPM and the CONN toolbox (<https://www.nitrc.org/projects/conn>). The results were thresholded at  $p=0.001$ , and resulting clusters were FDR-corrected at  $p=0.05$ .



**Supplemental figure 4-2.** Hit rate (A) and false alarm rate (B) for Experiment 2 tests. Error bars indicate standard error of the mean. The Crawford-Howell test indicated a significant dissociation between the two tests for both hit rate (S.A.:  $t(11) = 18.95$ ,  $p < .001$ ; P.R.:  $t(11) = 19.59$ ,  $p < .001$ ) and false alarm rate (S.A.:  $t(11) = 12.55$ ,  $p < .001$ ; P.R.:  $t(11) = 20.31$ ,  $p < .001$ ).

# Chapter 5

## Amodal semantics outside the language network

### Collaborators:

Carina Kauf, Hope H. Kean, Tanya Goldhaber, Zachary Mineroff, Zuzanna Balewski, Alfonso Nieto-Castañón, Rosemary Varley, Nancy Kanwisher, Evelina Fedorenko

### 5.1 Abstract

The language system is often implicated in the representation and manipulation of abstract semantic knowledge. However, this view is not consistent with a large body of evidence suggesting that language processing is neurally distinct from the rest of cognition. Here, we describe a set of brain regions, separate from the language network, that is engaged in task-based semantic processing of both linguistic and non-linguistic stimuli. In three fMRI experiments (49 sessions across 41 participants), participants viewed blocks of events presented as either sentences or pictures. In half of the blocks, participants performed a semantic task; in the other half, they performed a low-level perceptual task, like tracking the motion of the stimulus on the screen. The stimuli and the exact tasks varied across experiments. We found that a set of brain regions, located primarily in left lateral prefrontal cortex, left temporo-parietal cortex, and right cerebellum, responded selectively to semantic tasks for both sentences and pictures. These regions are spatially and functionally separate from the classic language, multiple demand (MD), and default mode networks (DMN), exhibiting a unique response profile. Our results show that semantic processing is distinct from both linguistic and domain-general cognitive processing and engages its own neural machinery.

## 5.2 Introduction

The ability to flexibly leverage information to achieve specific goals is a key component of intelligent behavior. A particularly valuable type of information is semantic knowledge — generalized, abstract information about entities, actions, and ideas. Here, we investigate the brain basis of flexible, task-driven semantic processing.

A vital property of semantic processing is its generalizability. For instance, we can recognize an event “a girl is singing” regardless of the identity of the girl and the song and regardless of whether we hear this event or see it. Further, with the emergence of language, we do not even need to perceive an event to activate its representation: a string of letters describing it is sufficient. Generalizability allows us to integrate information across different types of input so that it can be used in a wide range of future situations. Thus, we investigate semantic processing of events presented in two different formats: sentences vs. pictures.

Our goal is to determine whether any brain regions are selectively engaged in semantic processing regardless of input format. To do so, we conduct a whole-brain search for regions that respond to both sentence and picture semantics and then, in held-out data, determine whether the magnitude of their response is biased toward a particular input format. A putative semantic brain region/network should respond strongly during semantic processing of both sentences and pictures. Following the terminology in much prior work, we will use the term “amodal” to refer to brain regions that are invariant to verbal vs. pictorial presentation format (e.g., Coccia et al., 2004; Fairhall & Caramazza, 2013; Gainotti, 2011; Patterson et al., 2007; Pobric et al., 2010).

A substantial number of studies report overlapping activations/representations for sentences and pictures in left-lateralized frontal and/or temporal/temporo-parietal brain regions broadly resembling the language network (e.g., Devereux et al., 2013; Fairhall & Caramazza, 2013; Handjaras et al., 2017; Krieger-Redwood et al., 2015; Shinkareva et al., 2011; Vandenberghe et al., 1996; Visser et al., 2012). As a result of this and other work, frontal and temporal language regions are sometimes argued to carry out so-called

“semantic control”, i.e., selection of relevant semantic information in response to a task (e.g., Davey et al., 2016; Jefferies, 2013; Lambon Ralph et al., 2017; Martin & Chao, 2001; Whitney et al., 2011, 2012; Zhang et al., 2022). However, most studies do not explicitly compare activity evoked by the semantic task with activity evoked by linguistic processing alone, making it impossible to determine whether these processes rely on the same brain regions in a given participant. In recent work, we showed that, if the language regions are defined using an independent localizer task, they respond primarily to sentences and weakly or not at all to pictures and movies (except the language fROI in left angular gyrus; e.g., Benn et al., 2021; Shain, Paunov, Chen et al., 2022; Sueoka et al., 2022). These findings raise the possibility that previously reported ‘semantic’ regions are actually distinct from the language regions, even though they may broadly fall within the same anatomical areas.

Furthermore, the majority of past studies of semantic processing have focused on single objects, so the results may not generalize to actions (e.g., Wurm & Caramazza, 2019) or events (Baldassano et al., 2018; Z. Hu et al., 2019; Jouen et al., 2015; Thierry & Price, 2006). We define an event as an action along with the entity or entities participating in that action. Of key importance to event representations are *relations* between participating entities, which determine their event (thematic) roles. For instance, an event “a cat is chasing a dog” requires us to identify not only the participating entities (cat and dog), but also the interaction taking place between them (chasing), as well as the role that each participant plays in that interaction (chasing vs. being chased). Thus, semantic processing of events requires object and action processing but is not reducible to them (Dresang et al., 2019).

The few studies that have investigated event processing across modalities implicate a set of regions known as the default mode network (DMN; Baldassano et al., 2018; Z. Hu et al., 2019; Jouen et al., 2015; cf. Thierry & Price, 2006). The DMN has been hypothesized to support construction of structured mental situation models (Baldassano et al., 2017; Hassabis & Maguire, 2009) and encode specific event schemas (Baldassano et al., 2018), making it a prime candidate for semantic processing of events. That said, some studies have indicated that different brain regions within the DMN might belong to functionally

distinct subnetworks (Andrews-Hanna et al., 2010; Braga et al., 2019; Braga & Buckner, 2017; Deen & Freiwald, 2022; DiNicola et al., 2020; Wen et al., 2020), only some of which might be engaged in situation model construction. It is also unclear whether the DMN gets engaged only in response to complex, long-lasting stimuli (as in, e.g., Baldassano et al., 2017, 2018) or is also active during brief event presentations (such as observing a scene for 1-2 seconds).

Finally, an important factor that should be considered in any study of semantic processing is the presence of an external task. Studies of other cognitive domains, such as inner speech (Hurlburt et al., 2016) and conscious stimulus processing (Frässle et al., 2014; Tsuchiya et al., 2015), have shown that the presence of an external task (e.g., following the experimenter's instructions about when to press a button) has a strong effect on the observed activation patterns. In particular, cognitive tasks activate a set of frontal and parietal regions known as the multiple demand (MD) network (Duncan, 2010; Duncan & Owen, 2000). The MD network is active during many different cognitively challenging tasks (Assem, Glasser, et al., 2020; Fedorenko et al., 2013; Hugdahl et al., 2015; Shashidhara, Mitchell, et al., 2019). Typically, activity in the MD network and the DMN is negatively correlated: external tasks activate MD regions and deactivate DMN regions. However, some studies also reported DMN recruitment during semantic tasks (Evans et al., 2020; Jung et al., 2021; Lanzoni et al., 2020; X. Wang et al., 2021; Zhang et al., 2022).

Thus, three previously characterized networks might contribute to semantic processing of events: the fronto-temporal language network, the domain-general multiple demand network, and the default mode network (see Xu et al., 2016, 2017, for an account of semantic processing that implicates all three). Thus, in addition to localizing semantic regions using a whole-brain analysis, we also test response to semantic processing of events in these three networks.

Our critical experiments probe task-driven semantic processing of events presented as sentences versus pictures. By introducing design variations across the three experiments, we test whether our results generalize across semantic tasks (plausibility vs. reversibility),

event type (animate-animate or animate-inanimate interactions), and picture type (photos vs. line drawings).

A key distinguishing feature of our work compared to previous investigations of amodal semantics is the use of the individual-subject functional localization method (Fedorenko et al., 2010; Saxe, Brett, et al., 2006). This approach stands in contrast to traditional group-averaging analyses, whereby neural responses are averaged across participants on a voxel-by-voxel basis, and the resulting activation clusters are interpreted via ‘reverse inference’ from anatomy (e.g., Fedorenko, 2021; Poldrack, 2006, 2011). Group analyses tend to overestimate overlap in cases of nearby functionally distinct areas (Nieto-Castañón & Fedorenko, 2012), which is problematic for studies that aim to establish regions of shared activations across input modalities. This consideration is particularly important when analyzing responses in association cortex, where different functional regions vary in their precise locations across individuals (Blank et al., 2017; Fedorenko & Kanwisher, 2009; Frost & Goebel, 2012; Shashidhara, Spronkers, et al., 2019; Tahmasebi et al., 2012; Vázquez-Rodríguez et al., 2019) and often lay side by side in the form of ‘interdigitated networks’ (Braga et al., 2019, 2020; Deen & Freiwald, 2022; DiNicola et al., 2020; Fedorenko, Duncan, et al., 2012). The use of functional localization allows us to distinguish between semantic, linguistic, and domain-general executive processes even if they recruit brain regions that are adjacent to each other, thus enabling powerful and generalizable inferences about the neural basis of task-driven semantics.

## 5.3 Method

### 5.3.1 Participants

We collected data from 14 participants for Experiment 1 (11 women, 3 men, mean age = 25.9 years, SD = 8.10), 18 participants for Experiment 2 (7 women, 11 men, mean age = 21.0 years, SD = 1.68), and 26 participants for Experiment 3 (11 women, 13 men, mean age = 23.4 years, SD = 4.73). The participants were recruited from MIT and the surrounding Cambridge/Boston, MA, community and paid for their participation. All were native

speakers of English, had normal or corrected to normal vision, and no history of language impairment. Two participants from Experiment 1, one from Experiment 2, and three from Experiment 3 were excluded due to low data quality; two participants from Experiment 2 and one participant from Experiment 3 were further excluded due to low behavioral task performance, leaving a total of 12, 15, and 22 participants for Experiments 1, 2, and 3, respectively. The protocol for the study was approved by MIT's Committee on the Use of Humans as Experimental Subjects (COUHES). All participants gave written informed consent in accordance with protocol requirements.

	<b>SENTENCES</b>	<b>PICTURES</b>	<b>TASKS</b>
Expt. 1	A man is peeling a carrot.		PLAUSIBLE OR IMPLAUSIBLE?
	A man is peeling a candle.		MOVING LEFT OR RIGHT?
Expt. 2	A man is eating a clementine.		REVERSIBLE OR IRREVERSIBLE?
	A woman is putting staples into a stapler.		3 OR 4 DIRECTION CHANGES?
Expt. 3	The cop is arresting the criminal.		PLAUSIBLE OR IMPLAUSIBLE?
	The criminal is arresting the cop.		MOVING LEFT OR RIGHT?

**Figure 5-1.** Sample stimuli used for the three critical experiments. In Experiments 1 and 3, sentences and pictures described the same events (but varied in presentation details; see **Section 5.3.2**). Half of the events were plausible, and half were implausible (in Experiment 1, plausibility was manipulated by replacing the object; in Experiment 3, plausibility was manipulated by switching the agent and the patient). In Experiment 2, sentences and pictures described different events. All events were plausible, but some were reversible (the action could be undone) and some were irreversible.

### 5.3.2 Critical experiments

The critical experiments in all three studies had a 2x2 blocked design: in each block, the stimuli were either sentences or pictures, and the task was either semantic (main) or

perceptual (control). Thus, each experiment had four conditions: sentences + semantic task (SENT\_SEM), sentences + perceptual task (SENT\_PERC), pictures + semantic task (PIC\_SEM), and pictures + perceptual task (PIC\_PERC). To test the generalizability of our findings, we varied the nature of the stimuli and tasks across experiments (**Figure 5-1**). During behavioral piloting, the semantic and perceptual tasks within each experiment were adjusted to be approximately matched in difficulty.

### 5.3.2.1 Experiment 1

The materials consisted of 120 pairs of color photographs depicting interactions of people with everyday objects (except one pair that depicted animate-animate interactions), as well as 120 pairs of ‘corresponding’ sentences that described the same interactions. Half of the stimuli depicted/described plausible scenarios (e.g., “peeling a carrot”), and half of the stimuli depicted/described implausible scenarios (e.g., “peeling a candle”). The photographs in each pair were highly similar except for the object used. Furthermore, materials were created in quadruplets, so that any given object was used in a plausible version in one pair, and in an implausible version in another pair (for example, for the peeling a candle / carrot, there was another pair: lighting up a carrot / candle).

The semantic task was to determine whether a given event was plausible or implausible. The perceptual task was to determine whether the stimulus was moving left or right (see below).

Before the experiment, participants completed a practice run outside the fMRI scanner with 8 picture and 8 sentence stimuli (which were never shown inside the scanner). During the experiment, each of the 480 unique stimuli (240 pictures and 240 sentences) was presented twice: once in the semantic task, and once in the perceptual task, for a total of 960 trials. The 960 trials were split into 5 subsets of 192 trials each (48 for each of the four conditions), corresponding to five runs, with the constraint that any given event (e.g., a man peeling a carrot) occurred only once within a run (i.e., in only one of the four conditions). The 192 trials in each run were further grouped into 16 blocks (4 per condition) of 12 trials each, ensuring that the plausible and the implausible versions of

the same event did not occur within the same block (within or across the four conditions). The exact grouping of the trials varied across participants.

Each trial lasted 2 s and consisted of the stimulus presented for 1.4 s followed by 0.6 s of fixation. During stimulus presentation (across all four conditions), the stimulus appeared on the screen (centered vertically, with the horizontal position drawn uniformly from an 80 pixel range around the vertical midline of the screen) and then moved diagonally toward the top left or top right corner. The picture and sentence stimuli moved at the velocity of (1,10) and (2,10) pixels in (x,y)-direction, respectively (the speed was chosen so as to approximately match the semantic and perceptual conditions for difficulty, based on behavioral piloting). Each block was preceded by a 2 s instructions screen to tell participants whether to perform the semantic or the perceptual task on the pictures/sentences that follow, and to remind them which button to press for which response: “1 – plausible; 2 – implausible” or “1 – left; 2 – right”. The instructions were additionally displayed in small font in the top right corner throughout each block.

Each experimental block lasted 26 s (2 s instructions + 12 trials \* 2 s each). Each run consisted of 16 experimental blocks (4 per condition) and 5 fixation blocks each 16 s in duration, for a total duration of 496 s (8 min 16 s). A fixation block appeared at the beginning of each run, and after each set of four experimental blocks. Condition order was counterbalanced across runs and participants and was palindromic within each run (to counteract scanner drift). Each participant completed five runs for a total of 20 blocks per condition per subject. For consistency with the other two experiments, each of which only included two runs, we here focus on the first two runs for each participant (the responses for all five runs were attenuated due to adaptation but showed a similar pattern).

### 5.3.2.2 Experiment 2

The materials consisted of color photographs and sentences depicting/describing interactions of people with everyday objects (except two photos and two sentences that depicted / described animate-animate interactions). Unlike in Experiment 1, the scenarios

in the pictures and sentences were different (to minimize the possibility of cross-coding: i.e., recalling the verbal representation for a picture if a sentence version of the same event has been previously encountered, or recalling the pictorial representation for a sentence if a picture version has been previously encountered). Further, all scenarios were plausible (to ensure that responses in Experiment 1 are not driven by the unexpected nature of half of the stimuli), but some depicted/described irreversible actions, i.e., actions, the effects of which cannot be undone (e.g., eating a clementine or peeling a carrot) and others depicted/described reversible actions (e.g., putting staples into a stapler or inserting a key into a lock). A subset of the plausible photographs from Experiment 1 ( $N=112$ ) were reused for this study along with newly created sentences ( $N=112$ ).

The semantic task was to determine whether a given action was reversible or irreversible. The perceptual task was to determine whether the movement direction of the stimulus changed 3 or 4 times (two participants completed an earlier version of the task where the direction changed either 5 or 7 times).

Before the experiment, participants completed a practice run outside the scanner with 8 picture and 8 sentence stimuli (not used inside the scanner). For the main experiment, we created two lists, where each stimulus was assigned either to the semantic condition or to the perceptual condition. Across the two lists, we balanced the ratio of actions irreversibly affecting the object to those reversibly doing so and the split of these event types across modalities (i.e., each list consisted of 83 irreversible events, and 141 reversible events; 43 of the 83 irreversible events were presented as pictures and 40 were presented as sentences). Each participant was assigned one of the two lists. Each list was split into two subsets of 112 trials each (28 for each of the four conditions), corresponding to two runs. The 112 trials in each run were further grouped into 16 blocks (4 per condition) of 7 trials each, ensuring that each block contained between 1 and 4 irreversible events and that at most 3 trials in a row changed direction the same number of times.

Each trial lasted 2 s and consisted of the stimulus presented for 1.4 s followed by a prompt (“RESPOND”) presented in the center of the screen in red font for 0.6 s. During stimulus

presentation (across all four conditions), the stimulus appeared in the center of the screen and then randomly changed directions. The picture and sentence stimuli moved at the velocity of (1,10) and (2,10) pixels in (x,y)-direction, respectively (the speed was chosen so as to approximately match the semantic and perceptual conditions for difficulty, based on behavioral piloting). Each block was preceded by a 2 s instructions screen to tell participants whether to perform the semantic or the perceptual task on the pictures/sentences that follow, and to remind them which button to press for which response: “1 - irreversible; 2 – reversible” or “1 – odd number of direction changes; 2 – even number of direction changes”. The instructions were additionally displayed in small font in the bottom left corner throughout each block.

Each experimental block lasted 16 s (2 s instructions + 7 trials \* 2 s each). Each run consisted of 16 experimental blocks (4 per condition) and 5 fixation blocks each 16 s in duration, for a total duration of 336 s (5 min 36 s). A fixation block appeared at the beginning of each run, and after each set of four experimental blocks. As in Experiment 1, condition order was counterbalanced across runs and participants, and was palindromic within each run. Each participant completed two runs.

### 5.3.2.3 Experiment 3

The data from this experiment are the same as those described in Chapter 4. The materials consisted of 40 pairs of line drawings depicting interactions between two animate entities, as well as 40 pairs of ‘corresponding’ sentences that described the same interactions. As in Experiment 1, half of the stimuli depicted/described plausible scenarios (e.g., “a cop is arresting a criminal”), and half of the stimuli depicted/described implausible scenarios (e.g., “a criminal is arresting a cop”). The implausible events were created by swapping the agent and the patient of the plausible events. The semantic task was to determine whether a given event was plausible or implausible, and the perceptual task was to determine whether the stimulus was moving left or right (as in Experiment 1).

As in Experiment 2, we created two lists, where each stimulus was assigned either to the semantic condition or to the perceptual condition. Each list was split into two subsets of

80 trials each (20 for each of the four conditions), corresponding to two runs, ensuring that the same event (in picture vs. sentence format) does not occur in the same run. The 80 trials in each run were further grouped into 8 blocks (2 per condition) of 10 trials, ensuring that at most 3 plausible or implausible events appeared in a row. Movement direction of the stimuli was assigned randomly.

Each trial lasted 2 s and consisted of the stimulus presented for 1.5 s followed by 0.5 s of fixation. During stimulus presentation (across all four conditions), the stimulus appeared in the center of the screen and then moved horizontally toward the left or right. The picture and sentence stimuli both moved at the velocity of 0.025 pixels/screen in positive or negative x-direction (the speed was chosen to approximately match the semantic and perceptual conditions for difficulty, based on behavioral piloting). Each block was preceded by a 2 s instructions screen to tell participants whether to perform the semantic or the perceptual task on the pictures/sentences that follow, and to remind them which button to press for which response: "1 - plausible; 2 – implausible" or "1 - left; 2 – right". The instructions were additionally displayed in small font in the bottom left corner throughout each block.

Each experimental block lasted 22 s (2 s instructions + 10 trials \* 2 s each). Each run consisted of 8 experimental blocks (2 per condition) and 3 fixation blocks each 22 s in duration, for a total duration of 242 s (4 min 2 s). A fixation block appeared at the beginning of each run, and after each set of four experimental blocks. Condition order was counterbalanced across runs and participants and was palindromic within each run. Each participant completed two runs.

### **5.3.3 Localizer experiments**

In addition to the critical experiment, all participants completed the language localizer experiment (sentence and nonword list reading), used to identify language-responsive regions in individual participants (Fedorenko et al., 2010). Participants in Studies 2 and 3 also completed a spatial working memory task (Fedorenko et al., 2013), used to identify multiple demand and default mode regions in individual participants.

### **5.3.3.1 Language localizer**

Participants from Experiment 1 read sentences (“THE DOG CHASED THE CAT ALL DAY LONG”), word lists (derived by scrambling the words across sentences; “THE FOR JUICE UP GARDEN AROUND TRIES LILY”), and lists of unconnected, pronounceable nonwords (“CRON DACTOR DID MAMP FAMBED BLALK THE MALVITE”) in a blocked design. Each stimulus consisted of eight words/nonwords. For details of how the language materials were constructed, see Fedorenko et al. (2010). Stimuli were presented in the center of the screen, one word/nonword at a time, at the rate of 350 ms per word/nonword. Each stimulus was preceded by a 300 ms blank screen and followed by a 1350 ms screen showing the memory probe and then a blank screen for another 350 ms, for a total trial duration of 4.8 s. Participants were instructed to perform a memory probe task at the end of each trial (deciding whether a word/nonword appeared in the string just read). Condition order was counterbalanced across runs. Experimental blocks lasted 24 s (with 5 trials per block), and fixation blocks lasted 16 s. Each run (consisting of 3 fixation blocks and 12 experimental blocks) lasted 336 s. Each participant completed 5 runs; we here report data from the first three runs, to approximately match the amount of data from each condition with the other version of the language localizer.

For participants from Experiments 2 and 3, the language localizer task was the same as that described in Section 2.5.2.3. Multiple studies show that the language localizer is invariant to changes in materials, task, and presentation modality, so both localizers should identify the same network.(e.g., Cheung et al., 2020; Diacheck et al., 2020; Fedorenko et al., 2010; e.g., Lipkin et al., 2022; Malik-Moraleda, Ayyash, et al, 2022; Scott et al., 2017).

### **5.3.3.2 Multiple demand and default mode localizer**

The multiple demand localizer task was the same as that described in Section 2.5.2.2. The hard > easy contrast has been previously shown to reliably activate bilateral frontal and parietal MD regions (Assem, Glasser, et al., 2020; Blank et al., 2014b; Fedorenko et al.,

2013). Numerous studies have shown that the same brain regions are activated by diverse executively-demanding tasks (Duncan & Owen, 2000; Fedorenko et al., 2013; Hugdahl et al., 2015; Shashidhara, Mitchell, et al., 2019; Woolgar et al., 2011). The easy > hard contrast has been shown to localize default mode regions (Blank & Fedorenko, 2020; Mineroff et al., 2018), which are generally attenuated in the presence of an external task (e.g., Buckner et al., 2008; Buckner & DiNicola, 2019; Fox et al., 2005; Raichle, 2015).

### **5.3.4 fMRI data acquisition**

The data acquisition procedure was the same as that described in Section 2.5.3.

### **5.3.5 fMRI data preprocessing**

The preprocessing procedure was the same as that described in Section 2.5.4.

### **5.3.6 First-level analysis**

First-level analysis procedure was the same as that described in Section 2.5.5. For the localizer tasks, we modeled entire blocks. For the critical tasks, we modeled entire blocks (sentences, semantic task; pictures, semantic task; sentences, perceptual task; pictures, perceptual task) with the exception of the instruction reading phase, which was modeled as a separate condition.

### **5.3.7 Defining functional regions of interest (fROIs)**

The critical analyses were restricted to individually defined functional regions of interest (fROIs). These fROIs were defined using the group-constrained subject-specific (GcSS) approach (Fedorenko et al., 2010; Julian et al., 2012), where a set of ‘parcels’ (masks) is combined with each individual subject’s activation map for the relevant contrast to constrain the definition of individual fROIs.

The parcels delineate the expected gross locations of activations for a given contrast and are sufficiently large to encompass the extent of variability in the locations of individual

activations. We used two types of parcels: (a) semantic parcels defined based on the critical experiments from this study, and (b) three sets of parcels for well-established and characterized large-scale networks defined based on fMRI activation maps from the relevant localizer experiments: the language network, the multiple demand (MD) network, and the default mode network (DMN), all of which might contribute to semantic processing. The semantic parcels were generated based on the data from 30 participants from this experiment (10 from each critical experiment, in order to provide a balanced representation of the three experiments). We used a subset of the data to be able to test whether our results generalize to left-out participants. The parcels were generated using a conjunction contrast of SENT\_SEM > SENT\_PERC (a contrast that targets semantic processing in sentences) and PIC\_SEM > PIC\_PERC (a contrast that targets semantic processing in pictures). The language network parcels were generated from previously collected language localizer data from 220 participants (using the sentences>nonwords contrast). The multiple demand and default mode network parcels were from previously collected spatial working memory task data from 197 participants (using the hard>easy and easy>hard contrasts, respectively). For details of the parcel definition procedure, see Fedorenko et al (2010).

Within each parcel, we selected the top 10% most responsive voxels, based on the *p*-values for the contrast(s) of interest. This top n% approach ensures that the fROIs can be defined in every participant, thus enabling us to generalize the results to the entire population (Nieto-Castañón & Fedorenko, 2012). The contrasts used to define the fROIs were the same as those used to create the parcels. For a simple contrast, we ranked the voxels based on the *p*-value for that contrast; for a conjunction contrast (i.e., SENT\_SEM > SENT\_PERC and PIC\_SEM > PIC\_PERC), we ranked the voxels based on the larger of the two *p*-values (which is equivalent to a soft “and” conjunction contrast but allows us to specify the exact number of voxels to be selected).

### 5.3.8 Examining the functional response profiles of fROIs

fROI response estimation procedure was the same as that described in Section 3.3.7.

One participant completed only one run of the multiple demand localizer task; therefore, we did not estimate the strength of their responses to the hard and easy multiple demand localizer conditions but ensured that the whole-brain activation maps for the hard>easy contrast looked as expected.

### 5.3.9 Statistical analyses

To compare responses across conditions, we ran linear mixed-effect regression models with participant as a random intercept (for analysis of responses at the network level, we also included fROI as a random intercept). The analysis was run using the *lmer* function from the *lme4* R package (Bates et al., 2015); statistical significance of the effects was evaluated using the *lmerTest* package (Kuznetsova et al., 2017). For most analyses, we jointly analyzed the data from the three critical experiments while adding *Experiment* and *Experiment\*Condition* terms to the regression model to account for potential inter-experiment variability. All follow-up analyses used FDR correction for multiple comparisons; *p*-values were adjusted based on the number of groupings used in each case (e.g., number of fROIs for within-network fROI analyses, number of fROIs times the number of experiments for evaluating response consistency across experiments, two for within-network hemisphere comparisons).

### 5.3.10 Overlap analysis

To determine whether our newly defined semantic fROIs overlap with fROIs from previously established networks, we calculated the overlap coefficient between each fROI pair (using the formula: number of voxels shared / number of voxels in the smaller fROI). As a control, we also calculated the overlap between fROIs defined using the same contrast, the same experiment, and the same parcel, but using data from two different runs.

## 5.4 Results

### 5.4.1 Multiple brain regions respond to semantic task across sentences and pictures

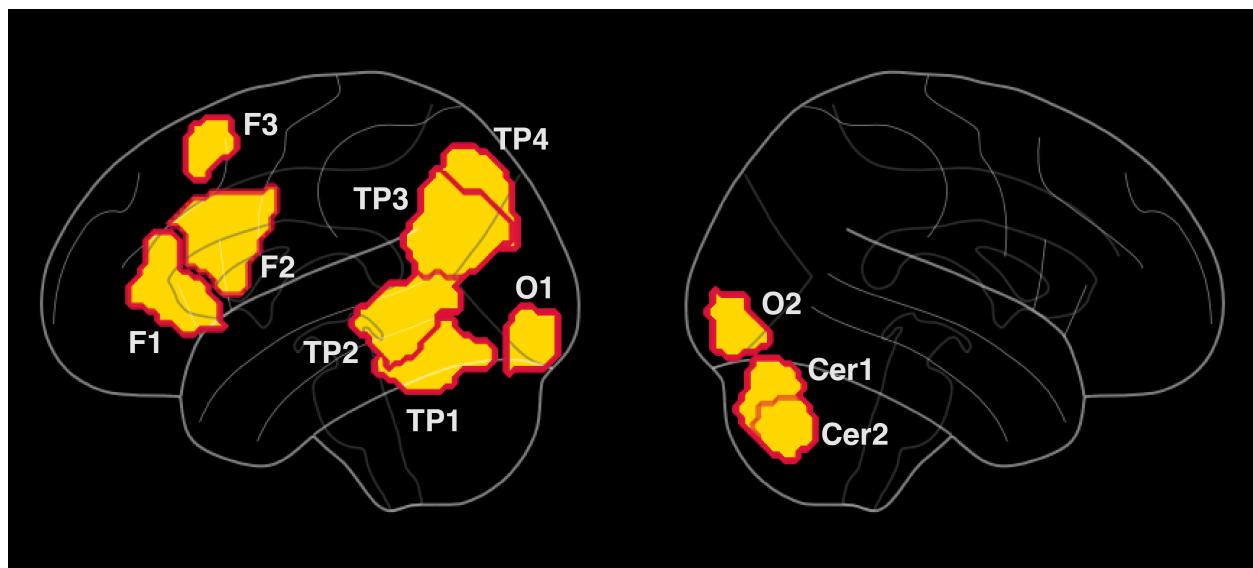
A whole-brain group-constrained subject-specific (GSS) analysis identified 11 brain regions (**Figure 5-2**) that respond more strongly to semantic tasks compared to perceptual tasks for both sentences and pictures (see **Section 5.3 (Method)** for analysis details). Three of them are located in left frontal cortex (F1-F3), four are located in left posterior temporal and left inferior parietal cortex (TP1-4), two are located in occipital cortex (left hemisphere: O1, right hemisphere: O2), and two are located in the right cerebellum (Cer1 and Cer2).

Statistical analyses of response magnitudes in the semantic fROIs (**Figure 5-3**), showed above-baseline responses to the semantic tasks (all  $p < .001$ ). Further, the response to the semantic tasks was stronger than the response to the perceptual tasks (all  $p < .001$ ). Most fROIs had equally strong responses to sentences and pictures, with the exception of TP1 (sentences < pictures:  $\beta = -0.81$ ,  $SE = 0.17$ ,  $p < .001$ ) and TP2 (sentences > pictures:  $\beta = 0.57$ ,  $SE = 0.21$ ,  $p = .038$ ). There was no interaction between task and stimulus type for any fROI. The stronger response to semantic tasks cannot be explained by task difficulty, as shown in the behavioral analyses (**Supplemental figure 5-1**). Overall, this response pattern indicates strong selectivity for semantic over perceptual tasks regardless of stimulus type.

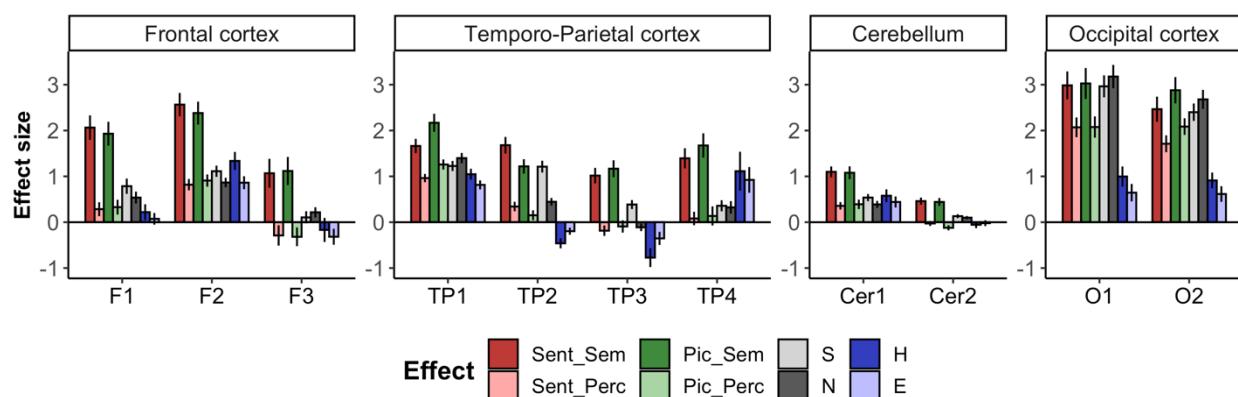
No semantic fROI exhibited a significant difference between hard and easy working memory task, a pattern markedly different from the domain-general task-responsive regions that comprise the Multiple Demand network (Fedorenko et al., 2013). Only two fROIs exhibited significantly higher responses to sentence reading than to nonword reading: TP2 ( $\beta = 0.72$ ,  $SE = 0.15$ ,  $p < .001$ ) and TP3 ( $\beta = 0.52$ ,  $SE = 0.17$ ,  $p = .011$ ); in other fROIs, there was no significant difference, indicating that these regions have a distinct response profile from the language network. The response to sentences during the semantic task was stronger than the response to passive sentence reading in all fROIs except O1 and O2 (n.s. for O1 and O2;  $p \leq .001$  for all others), suggesting that the response is at least

partially task-driven rather than being fully determined by the semantic content of the stimulus. The full table of results can be found in the Appendix (**Supplemental table 5-1**).

Finally, the response pattern is consistent between participants whose data was used to define the parcels and a new set of participants (**Supplemental figure 5-2**).

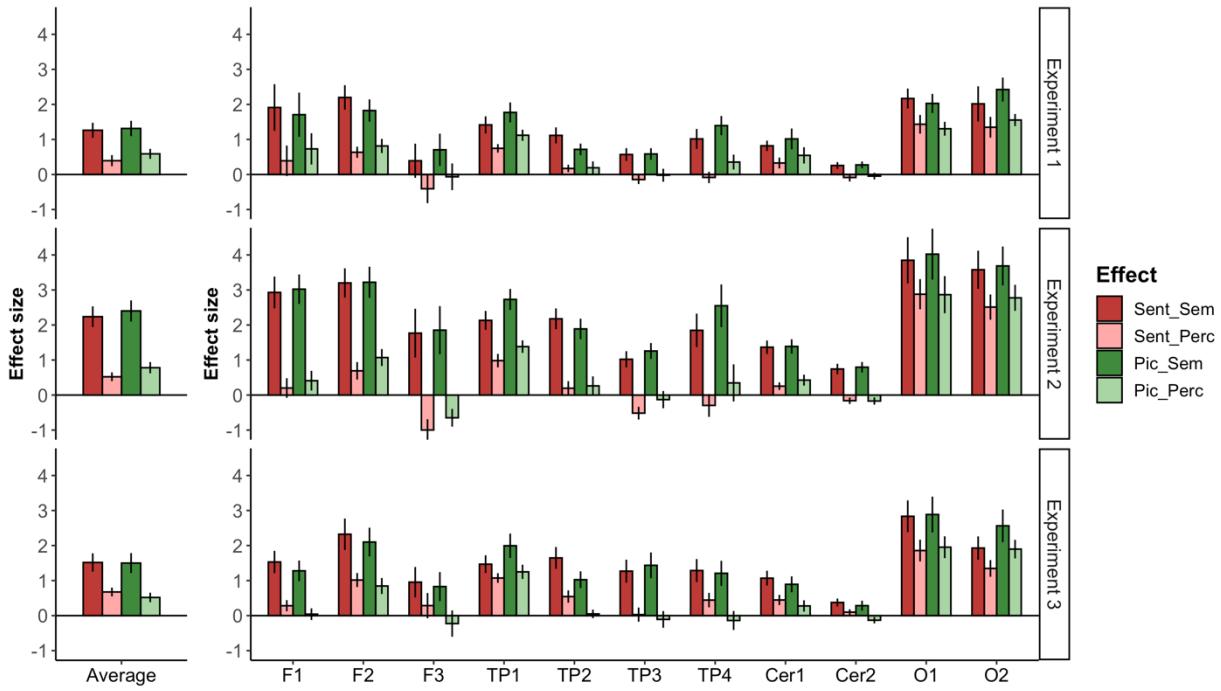


**Figure 5-2.** A left-lateralized set of parcels marking the approximate location of brain regions selective for semantic tasks for both sentences and pictures of events (individual fROIs are 10% of these parcels and their exact locations vary across individuals). F=frontal cortex, TP=temporal and parietal cortex, O=occipital cortex, Cer=cerebellum.



**Figure 5-3.** Responses to conditions from the critical experiments (Sent=sentences, Pic=pictures, Sem=semantic task, Perc=perceptual task; responses to each condition are combined across experiments), the language localizer (S=sentence reading, N=nonword

reading), and the multiple demand localizer (H=hard working memory task, E=easy working memory task). Here and below, error bars show the standard error of the mean.



**Figure 5-4.** Response patterns across the semantic fROIs are remarkably consistent across experiments. Left: responses averaged across the fROIs. Right: responses for individual fROIs. Sent=sentences, Pic=pictures, Sem=semantic task, Perc=perceptual task.

### 5.4.2 The responses are consistent across experiments

Selectivity for the semantic task was remarkably consistent across fROIs and experiments (**Figure 5-4**), reaching significance in all comparisons except the F3 fROI in Experiment 1 ( $p=.053$ ). Stimulus selectivity varied slightly between experiments, with

- three fROIs showing a preference for sentences over pictures in Experiment 3 (Cer2:  $\beta=0.23$ , SE=0.09,  $p=.040$ , F3:  $\beta=0.51$ , SE=0.18,  $p=.037$ , TP2:  $\beta=0.49$ , SE=0.15,  $p=.011$ )
- two fROIs showing a preference for pictures over sentences in Experiment 2 (F2:  $\beta=-0.38$ , SE=0.11,  $p=.011$ , TP3:  $\beta=-0.39$ , SE=0.14,  $p=.037$ ).

- two fROIs showing an interaction between stimulus type and task, with strongest responses to sentences during the semantic task in Experiment 1 (F2:  $\beta=0.55$ , SE=0.14, p=.009, TP2:  $\beta=-0.42$ , SE=0.10, p=.007).

In addition, TP1's preference for pictures over sentences, shown in the main analysis, reached significance in Experiments 1 and 2 (Expt1:  $\beta=-0.37$ , SE=0.08, p=.004, Expt2:  $\beta=-0.40$ , SE=0.10, p=.008) but not in Experiment 3. The full results are shown in **Supplemental table 5-2**.

Overall, despite some experiment-specific trends, the fROIs consistently show a preference for semantic task and are, in most cases, invariant to stimulus type (sentences vs. pictures).

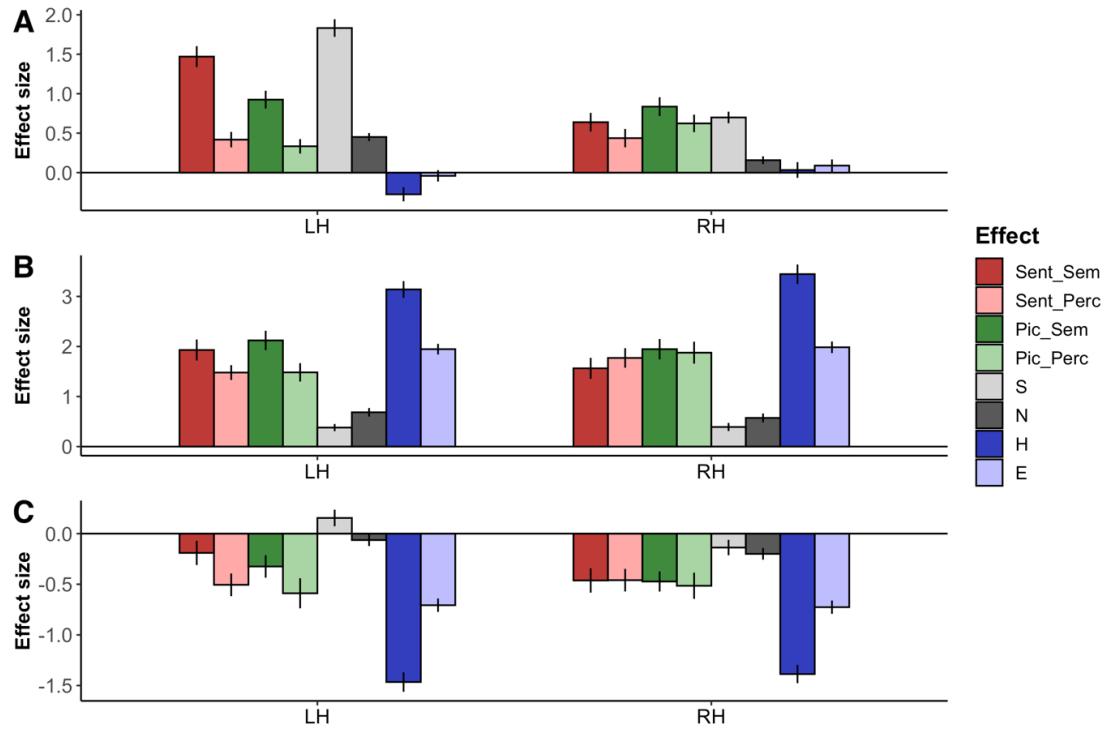
### 5.4.3 The response to semantic tasks in other cognitive networks

Next, we tested the response to semantic processing of sentences and pictures in other networks that may contribute to semantic processing: the language network (Fedorenko, Behr, et al., 2011; known to respond to meaningful linguistic input; e.g., Fedorenko et al., 2010; Scott et al., 2017), the multiple demand network (known to respond to cognitively demanding tasks; e.g., Duncan, 2010; Fedorenko et al., 2013; Shashidhara, Mitchell, et al., 2019; Woolgar et al., 2011), and the default mode network (previously reported to contribute to semantic tasks; e.g., Evans et al., 2020; Jung et al., 2021; Lanzoni et al., 2020; X. Wang et al., 2021; Zhang et al., 2022). As shown in **Figure 5-5** and elaborated below, the results show that the profile described above differs from the profiles in each of these three networks.

*The language network.* Consistent with our definition criteria and replicating much prior work, the language regions in both hemispheres showed stronger responses to sentence reading than to nonword reading (LH:  $\beta=1.28$ , SE=0.07, p<.001, RH:  $\beta=0.48$ , SE=0.06, p<.001). In addition, the right hemisphere, but not the left hemisphere language fROIs showed a stimulus-type-invariant response to event semantics. In both hemispheres, the response to semantic tasks (aggregated across sentences and pictures) was significantly above baseline (LH:  $\beta=1.20$ , SE=0.20, p<.001, RH:  $\beta=0.70$ , SE=0.14, p<.001) and above the

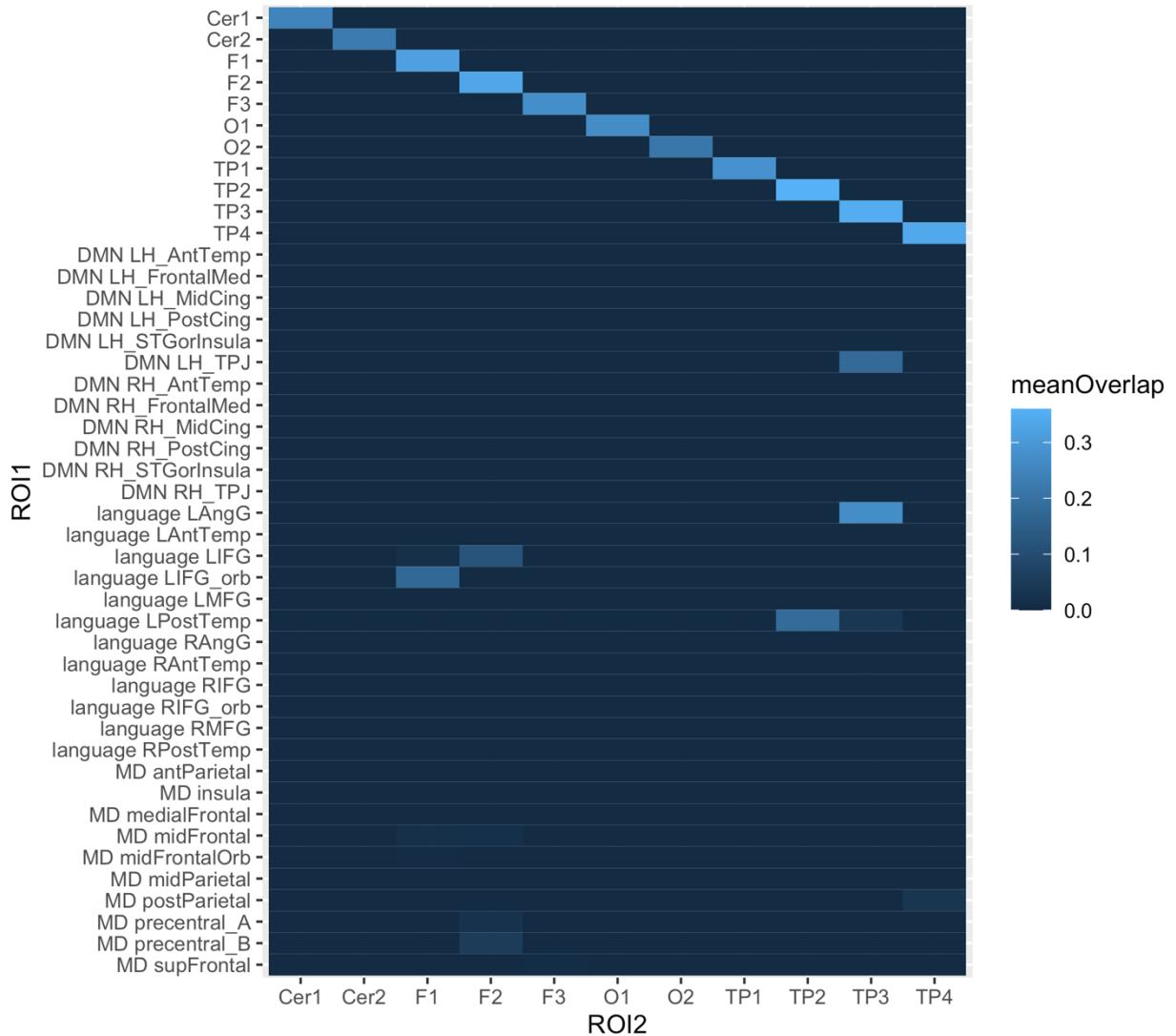
response to perceptual tasks (LH:  $\beta=1.60$ , SE=0.11,  $p<.001$ , RH:  $\beta=0.38$ , SE=0.10,  $p<.001$ ). The response to sentences was stronger than the response to pictures in the left hemisphere but weaker in the right (LH:  $\beta=0.63$ , SE=0.11,  $p<.001$ , RH:  $\beta=-0.41$ , SE=0.10,  $p<.001$ ). Finally, the interaction between task and stimulus type was significant in the left hemisphere, with sentences evoking a stronger response for the semantic task ( $\beta=0.46$ , SE=0.11,  $p<.001$ ), but not significant in the right hemisphere. Overall, although the right hemisphere language regions show a clear preference for sentences over nonwords, their response to sentences and pictures is similar in magnitude (and stronger during the semantic task).

*The multiple demand network.* Consistent with our definition criteria, the multiple demand regions responded more strongly to the hard than to the easy working memory task (LH:  $\beta=1.19$ , SE=0.07,  $p<.001$ , RH:  $\beta=1.46$ , SE=0.07,  $p<.001$ ) and to nonword reading compared to sentence reading (LH:  $\beta=-0.31$ , SE=0.07,  $p<.001$ , RH:  $\beta=-0.18$ , SE=0.07,  $p=.013$ ). The response to the semantic task was above baseline in both hemispheres (LH:  $\beta=2.16$ , SE=0.19,  $p<.001$ , RH:  $\beta=1.78$ , SE=0.19,  $p<.001$ ), but only the left hemisphere regions exhibited a preference for semantic over perceptual tasks ( $\beta=1.16$ , SE=0.12,  $p<.001$ ). Both hemispheres had an overall stronger response to pictures than to sentences (LH:  $\beta=-0.27$ , SE=0.12,  $p=.028$ , RH:  $\beta=-0.53$ , SE=0.13,  $p<.001$ ). Neither hemisphere showed an interaction between task and stimulus type for the critical experiments. Thus, left hemisphere multiple demand regions also exhibit a preference for semantic vs. perceptual tasks regardless of input type. Importantly, however unlike the semantic fROIs, these regions' response is much stronger during the (non-semantic) spatial working memory task).



**Figure 5-5.** Response to conditions of interest in previously characterized networks (averaged across fROIs in each hemisphere). (A) The language network, (B) The multiple demand network, (C) The default mode network.

*The default mode network.* Consistent with our definition criteria, the default mode regions responded less strongly to the hard than to the easy spatial working memory task (LH:  $\beta=-0.76$ , SE=0.07,  $p<.001$ , RH:  $\beta=-0.66$ , SE=0.06,  $p<.001$ ). The left hemisphere also had stronger responses to sentence reading than to nonword reading ( $\beta=0.22$ , SE=0.07,  $p=.002$ ). The response to the semantic tasks in the critical experiments was below baseline (LH:  $\beta=-0.27$ , SE=0.10,  $p=.020$ , RH:  $\beta=-0.49$ , SE=0.11,  $p<.001$ ). The left hemisphere regions had higher responses to the semantic than to the perceptual task ( $\beta=0.55$ , SE=0.12,  $p<.001$ ), whereas the right hemisphere regions showed no difference. There was no effect of stimulus type nor an interaction between task and stimulus type in either hemisphere. Thus, left hemisphere default mode regions show a stimulus-type-invariant preference for semantic vs. perceptual tasks, although their response remains below baseline.



**Figure 5-6.** Overlap between semantic fROIs and fROIs from three other cognitive networks (number of shared voxels divided by the number of voxels in the smaller fROI). Overlap within the events parcels (e.g., F1 to F1) is calculated using data from different runs (and is therefore a conservative estimate, as it only uses a subset of the data and the resulting fROIs are noisy).

#### 5.4.4 Semantic fROIs are largely distinct from known functional networks

To determine whether any of the semantic fROIs overlap spatially with the fROIs in the three previously characterized networks – language, MD, and DMN – we calculated voxel overlap between them (**Figure 5-6**). We observed a) substantial overlap between

TP3, LTPJ (DMN), and LAngG (language), b) partial overlap between TP2 and LPostTemp (language), c) partial overlap between F1 and LIFG\_orb (language), and d) partial overlap between F2 and LIFG (language). Overlap with other fROIs was negligible, indicating that most event semantics fROIs are distinct from regions of previously characterized networks.

Finally, preliminary analyses on a subset of participants who participated in both Experiment 2 and Experiment 3 ( $n=7$ ) show that fROIs defined using data from two different event semantics paradigms overlap but less so than fROIs defined using data from the same paradigm (**Supplemental figure 5-3**).

## 5.5 Discussion

We describe a set of brain regions that are active during semantic tasks regardless of stimulus type (sentences vs. pictures). Although a subset of these regions was implicated in semantic tasks before (e.g., Jefferies, 2013; Lambon Ralph et al., 2017), our key contribution here is to show that these regions are distinct both from the language network and from the domain-general multiple demand network. They are also distinct from the default mode network, which has been implicated in event semantics and semantic tasks more broadly. Thus, the regions we describe exhibit a distinct functional signature.

Below, we discuss our findings in the context of other research work, focusing on anatomical areas that have been discussed in work on semantic processing.

### 5.5.1 Inferior frontal gyrus (IFG)

The nature and extent of functional specialization within the IFG has been subject of much debate (e.g., Friederici et al., 2011; Fuster, 1991; Koechlin & Jubault, 2006; Matchin, 2018; Tettamanti & Weniger, 2006; Tremblay & Dick, 2016; Zaccarella & Friederici, 2017). Much of this debate is complicated by the fact that (a) the IFG is functionally

heterogeneous and (b) the location of the functionally specialized subregions in individual brains cannot be inferred from anatomy (e.g., Amunts et al., 1999, 2010; Fedorenko, Duncan, et al., 2012; Fedorenko & Blank, 2020; Tahmasebi et al., 2012; Vázquez-Rodríguez et al., 2019). Thus, the traditional approach of averaging activation maps across participants leads to the blurring of functionally specialized regions (Nieto-Castañón & Fedorenko, 2012; Shashidhara et al., 2020).

Prior work has shown that left lateral prefrontal cortex (including both inferior and middle prefrontal gyri) contains closely juxtaposed regions for linguistic and domain-general executive processing (Braga et al., 2020; Fedorenko, Duncan, et al., 2012). We here build upon that work to demonstrate a third kind of selectivity within this cortical region: high responses to semantic tasks. Our findings are consistent with prior reports of neural responses to semantic tasks (or ‘semantic control’) within the IFG (e.g., Badre et al., 2005; Davey et al., 2016; Martin & Chao, 2001; Thompson-Schill et al., 1998; Whitney et al., 2011, 2012; Zhang et al., 2022), but importantly, highlight that semantic regions’ response profile is distinct from that of both the language regions and the multiple demand regions.

### 5.5.2 Posterior temporal cortex

Like the IFG, posterior temporal cortex (specifically, posterior middle temporal gyrus, or pMTG) has been implicated in semantic control (e.g., Jefferies, 2013; Lambon Ralph et al., 2017) and in action representation across domains (e.g., Bedny et al., 2014; Bottini et al., 2020; Wurm & Caramazza, 2019). However, it is also considered to be a classic language processing site (Geschwind, 1970; Wernicke, 1874), responding to language input regardless of external task (e.g., Cheung et al., 2020; Diacheck et al., 2020; Fedorenko et al., 2010, 2020a; Lerner et al., 2011; cf. Tremblay & Dick, 2016 for a discussion of variability in the definition of “Wernicke’s area”).

Here, we take a step toward reconciling these accounts. We show that pMTG encompasses a fROI (TP2) that is responsive to both sentences and pictures during the semantic task (with a small bias toward sentences), but is also engaged during passive sentence processing, indicating its broad sensitivity to meaning rather than just task-

driven semantics. A preference for sentences over nonwords is a hallmark of the language regions, and indeed, this fROI partially overlaps with the posterior temporal fROI belonging to the language network (as defined with the language localizer). In addition, we find another fROI (TP1) located ventrally to TP2, which is sensitive to semantic task but does not show a preference for passive sentence reading over nonword reading. Unlike TP2, TP1 has a bias toward pictorial stimuli, suggesting that posterior temporal lobe may exhibit a selectivity gradient for sentence vs. picture semantics (as reported also by Devereux et al., 2013; Popham et al., 2021).

### 5.5.3 Angular gyrus (AG) and intraparietal sulcus (IPS)

The AG and, sometimes, IPS are also commonly implicated in semantic processing generally (Binder & Desai, 2011; Bonner et al., 2013; Kuhnke et al., 2022; A. R. Price et al., 2015) and/or in event/thematic processing in particular (Humphreys et al., 2021; Schwartz et al., 2011; Xu et al., 2018; cf. Humphreys, Halai, et al., 2022; Humphreys, Jung, et al., 2022). Here, in accordance with previous reports, we find semantically responsive fROIs within left AG (TP3) and left IPL (TP4). The AG semantic fROI overlaps with the language and default mode fROIs within the AG, in line with its response profile: high responses to semantic tasks for both sentences and picture, a preference for sentence reading over nonword reading, and deactivation in response to a spatial working memory task (modulated by task difficulty). The finding that the same fROI has functional signatures of semantic, language, and default mode regions is important given the fact that AG is, in general, highly structurally and functionally heterogeneous (e.g., Caspers et al., 2006; Choi et al., 2006; Humphreys & Tibon, 2022; Niu & Palomero-Gallagher, 2022; Seghier, 2013; Uddin et al., 2010) and highlights the value of functional localization for distinguishing between closely adjacent specialized regions (such as the IFG fROIs) and regions that exhibit several functional signatures (such as the AG semantic fROI).

#### **5.5.4 Cerebellum**

We found semantic responses in the right cerebellum, in line with accounts of cerebellar involvement in cognitive function in general (e.g., Ito, 2008; Jacobi et al., 2021; Rapoport et al., 2000; Schmahmann, 2019; Schmahmann et al., 2019; Strick et al., 2009) and language in particular (e.g., Fedorenko, Behr, et al., 2011; Highnam & Bleile, 2011; Mariën & Borgatti, 2018; Murdoch, 2010; Smet et al., 2007; Starowicz-Filip et al., 2017), recent work showing cerebellum's sensitivity to high-level conceptual information (LeBel et al., 2021), as well as reports of functional connectivity between cerebellar regions and cortical networks, including multiple demand and default mode networks (Buckner et al., 2011; Guell, Gabrieli, et al., 2018; Guell, Schmahmann, et al., 2018; for a review, see Habas, 2021). As in cortical regions, we show that the functional response profile of the fROIs we described (Cer1 and Cer2) is clearly distinct from both passive language and general executive function response profiles. Thus, our findings contribute to the body of work aiming to establish functional subdivisions within the cognitive cerebellum.

#### **5.5.5 Anterior temporal lobe (ATL)**

We did not find activity in ATL, a region commonly implicated in semantic processing (e.g., Bemis & Pylkkänen, 2013; Boylan et al., 2015; Humphreys et al., 2021; Lambon Ralph et al., 2017; Patterson et al., 2007; Schwartz et al., 2011; Visser et al., 2010, 2012). Although some accounts implicate the ATL in semantic representation rather than control (e.g., Lambon Ralph et al., 2017), we would still expect it to be more engaged during semantic than during perceptual tasks because performing semantic judgments would necessarily require activating the relevant representations. There are (at least) two possible explanations for why we did not observe responses in the ATL: conceptual and methodological. The conceptual explanation is that the ATL is engaged specifically in object and/or taxonomic knowledge (for review, see Mirman et al., 2017; cf. Jefferies et al., 2020) and therefore would not be active in a task that involves accessing event knowledge. The methodological explanation is that ventral ATL, which is most commonly implicated in general conceptual processing, has low signal-to-noise ratio in our data due to fMRI signal distortion (Visser et al., 2012). Thus, future work should

further probe the relationship between the ATL and the semantic regions described in this work.

### 5.5.6 Posterior medial parietal cortex

Finally, we did not observe semantic-task-evoked activity in posterior medial parietal cortex (precuneus and posterior cingulate areas), which was previously implicated in processing event semantics across input types (Baldassano et al., 2017, 2018; Fairhall & Caramazza, 2013). One possible reason is that this region tracks temporal input over long timescales (e.g., Blank & Fedorenko, 2020; Caucheteux et al., 2021; Chang et al., 2022; Lerner et al., 2011) and might not be sensitive to brief event presentations employed in our study. Another reason might be the fact that default mode network might not, in fact, be amodal. Some evidence indicates that the regions commonly associated with the default mode network, including posterior medial parietal cortex, might include distinct domain-specific networks (Braga et al., 2019; Deen & Freiwald, 2022; DiNicola et al., 2020), one of which is responsible for spatial processing and one for social processing. Under this account, our critical experiments (two of which focus primarily on physical interactions, and one focuses on social interactions) would activate different ‘default mode’ subnetworks, and so we would not observe a shared activation pattern.

Future work should investigate the extent of semantic domain specificity within the semantic regions. The activations pattern we observed generalize across experiments that focus on animate-animate interactions and those that mainly tackle animate-inanimate interactions (**Figure 5-4**), but the majority of our participants did only one critical experiment, thus leaving much room for mapping out semantic specialization patterns within the semantic parcels (see, e.g., Deniz et al., 2019; Huth et al., 2016; Popham et al., 2021, for evidence of such specialization). Another aspect that remains to be investigated in depth is the nature of semantic processing. Most fMRI studies of semantic processing — including our own — only employ one or two semantic tasks at a time (or focus on task-free processing altogether), but the nature of the task, i.e., the specific question being asked, has a substantial effect on brain activity (Toneva et al., 2020) and should be

considered further. Finally, resting state and naturalistic functional connectivity data can be used to determine whether these regions form a coherent network.

The ability to sift through world knowledge to retrieve relevant information is a remarkable cognitive feat. Thus, studies of brain regions specialized for semantic processing provide an important contribution to a mechanistic understanding of human cognition.

## **5.6 Acknowledgments**

We would like to acknowledge the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT and its support team. We thank the following individuals for help with the creation and norming of the stimuli, creation of experimental scripts, and data collection: Irina Shklyar, Eyal Dechter, Michael Behr, Jenelle Feather, Lily Jordan, and Matt Siegelman. We are grateful to Molly Potter, Ed Vul, Matt Lambon Ralph, Idan Blank for helpful discussions about semantics. This work was partially funded by NIH awards R00-HD057522 from NIDCD and R01-DC016607 from NICHD. EF was additionally supported by NIH awards R01-DC016950 from NICHD and U01-NS121471 from NINDS, as well as by research funds from the McGovern Institute for Brain Research, the Brain and Cognitive Sciences department, and the Simons Center for the Social Brain.

## 5.7 Appendix

### 5.7.1 Behavioral results

The behavioral results for the three critical experiments are shown in **Supplemental figure 5-1**.

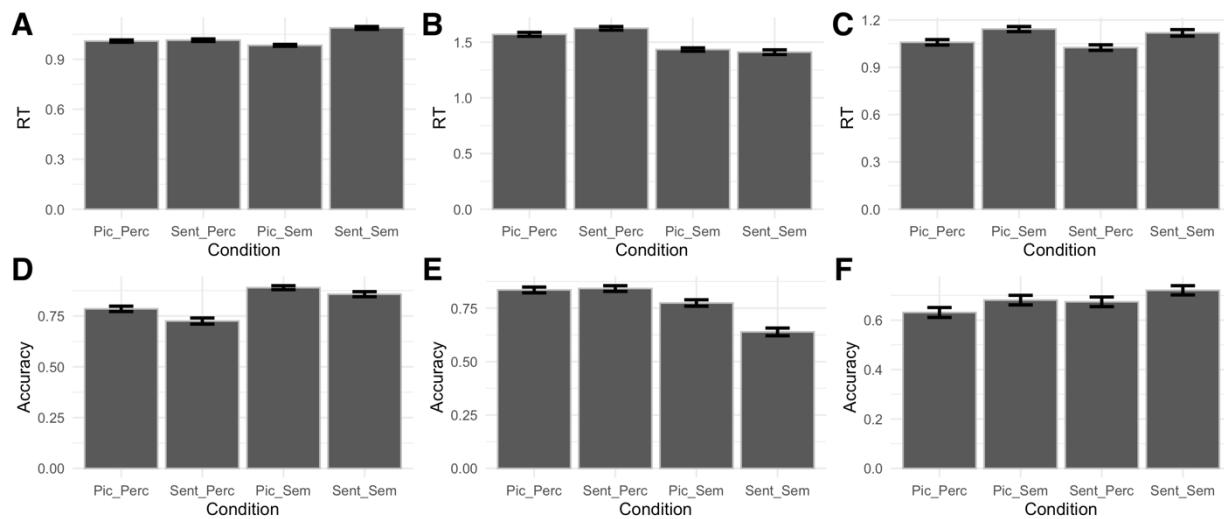
*Experiment 1.* Average response rate was 0.80. Overall mean reaction time was 1.02s, SD=0.22 (SENT\_SEM: 1.09s, SD=0.24; SENT\_PERC: 1.01s, SD=0.23, PIC\_SEM: 0.98s, SD=0.18, PIC\_PERC: 1.01s, SD=0.21). RTs were slightly higher for the semantic task than for the perceptual task ( $\beta=0.02$ , SE=0.01,  $p<.001$ ) and for sentence stimuli compared to picture stimuli ( $\beta=0.02$ , SE=0.01,  $p<.001$ ), with an interaction between stimulus type and task ( $\beta=0.1$ , SE=0.01,  $p<.001$ ). Overall mean accuracy was 0.82, SD=0.39 (SENT\_SEM: 0.86, SD=0.35; SENT\_PERC: 0.73, SD=0.45, PIC\_SEM: 0.89, SD=0.31, PIC\_PERC: 0.79, SD=0.41). Accuracy was higher for the semantic task than for the perceptual task ( $\beta=0.88$ , SE=0.09,  $p<.001$ ) and lower for sentence stimuli compared to picture stimuli ( $\beta=-0.28$ , SE=0.09,  $p<.001$ ), with no interaction between stimulus type and task.

*Experiment 2.* Average response rate was 0.89. Overall mean reaction time was 1.51s, SD=0.49 (SENT\_SEM: 1.41s, SD=0.56; SENT\_PERC: 1.62s, SD=0.44, PIC\_SEM: 1.43s, SD=0.42, PIC\_PERC: 1.57s, SD=0.49). RTs for the semantic task were slightly lower than for the perceptual task ( $\beta=-0.17$ , SE=0.02,  $p<.001$ ), with an interaction between stimulus type and task ( $\beta=-0.08$ , SE=0.03,  $p=0.016$ ). There was no main effect of stimulus type on reaction time. Overall mean accuracy was 0.77, SD=0.42 (SENT\_SEM: 0.64, SD=0.48; SENT\_PERC: 0.84, SD=0.36, PIC\_SEM: 0.77, SD=0.42, PIC\_PERC: 0.84, SD=0.37). Accuracy was lower for the semantic task than for the perceptual task ( $\beta=-0.79$ , SE=0.09,  $p<.001$ ) and for sentence stimuli compared to picture stimuli ( $\beta=-0.33$ , SE=0.09,  $p<.001$ ), with an interaction between stimulus type and task ( $\beta=-0.75$ , SE=0.19,  $p<.001$ ).

*Experiment 3.* Average response rate was 0.90. Overall mean reaction time was 1.09s, SD=0.52 (SENT\_SEM: 1.12s, SD=0.6; SENT\_PERC: 1.03s, SD=0.51, PIC\_SEM: 1.14s, SD=0.47, PIC\_PERC: 1.06s, SD=0.5). RTs for the semantic task were slightly higher than for the perceptual task ( $\beta=0.09$ , SE=0.02,  $p<.001$ ). There was no significant effect of

stimulus type (sentences vs. pictures) and no interaction between stimulus type and task. Overall mean accuracy was 0.68, SD=0.47 (SENT\_SEM: 0.81, SD=0.39; SENT\_PERC: 0.78, SD=0.42, PIC\_SEM: 0.75, SD=0.43, PIC\_PERC: 0.73, SD=0.44). Accuracy was higher for the semantic task than for the perceptual task ( $\beta=0.16$ , SE=0.07,  $p=0.031$ ) There was no significant effect of stimulus type and no interaction between stimulus type and task. For 16 participants, accuracy data from one run was erased because of a bug in the script.

Overall, there is no consistent trend in difficulty patterns across experiments, suggesting that the observed neural response patterns cannot be explained by between-condition differences in difficulty.

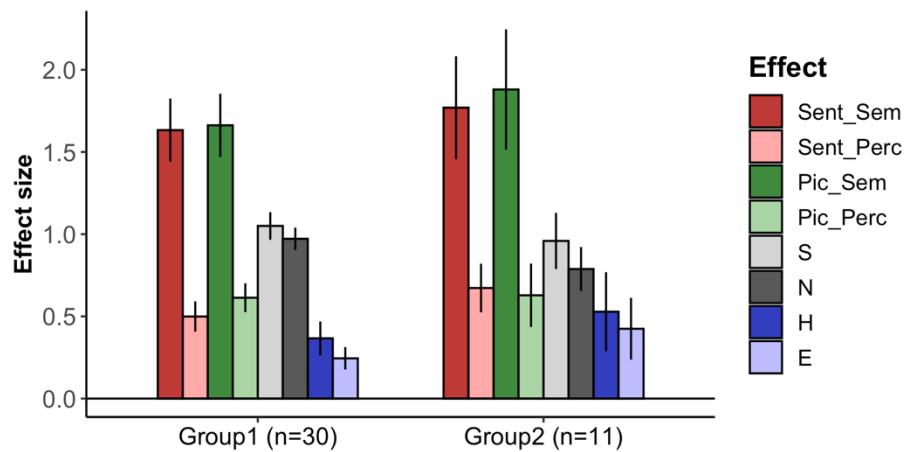


**Supplemental figure 5-1.** Behavioral results for the critical experiments. (A-C) Response times for Experiments 1, 2 and 3 respectively. (D-F) Accuracies for Experiments 1, 2, and 3, respectively.

### 5.7.2 The responses generalize to left-out participants

We use independent sets of data (different runs of the same task) when we define a fROI using data from a given experiment and measure its responses to the conditions from the same experiment. In addition, to test how well our results generalize to a new set of participants, we split the participants into two groups: the first group ( $n=30$ , 10 from each

experiment) were used to define the parcels in **Supplemental figure 5-2**, and the second group ( $n=11$ ) was left out. The results were consistent across the two groups: we observed a preference for semantic over perceptual tasks (both groups:  $p<.001$ ), no overall effect of stimulus type, no interaction between stimulus type and task, no preference for sentences over nonwords or for hard vs. easy versions of the working memory task.

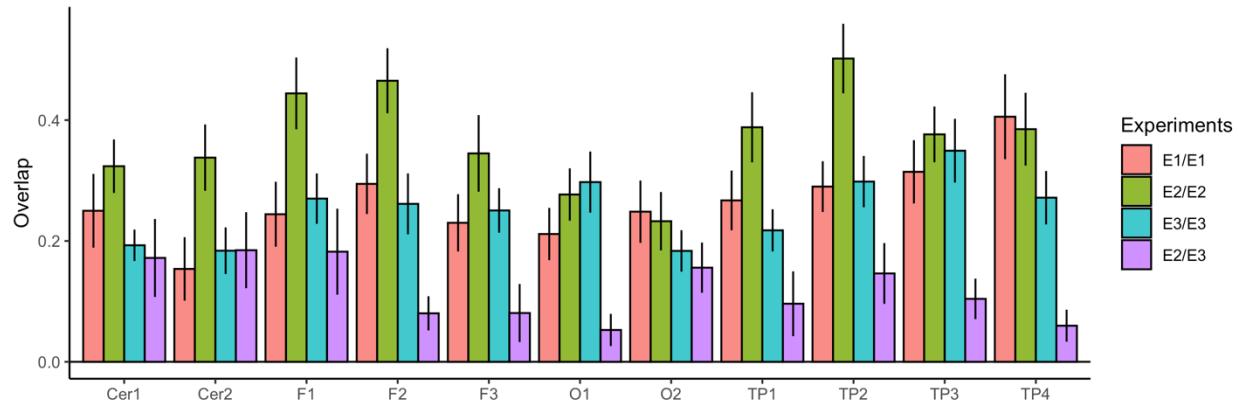


**Supplemental figure 5-2.** The results generalize to a new group of participants (whose data was not used for parcel definition). Sent=sentences, Pic=pictures, Sem=semantic task, Perc=perceptual task, S=sentence reading, N=nonword reading, H=hard working memory task, E=easy working memory task.

### 5.7.3 Different event semantics experiments recruit co-located but distinct cortical regions

We report preliminary analyses where we investigate the consistency of activations evoked by different event semantics experiments. We calculated the overlap between fROIs defined using different experimental runs, as well as between-experiment overlap, calculated from the data of 7 participants who completed both Experiment 2 and Experiment 3. We find that, for most fROIs, the between-experiment overlap is significantly lower than any of the within-experiment overlap values, indicating that variation in the experiment setup does affect which exact brain areas get recruited (**Supplemental figure 5-3**). That said, as shown in **Figure 5-4**, the response patterns of fROIs defined within the same parcel are generally consistent across experiments,

indicating that the GSS functional localization method enables us to capture similar response patterns across experiments despite differences in the exact locus of activation.



**Supplemental figure 5-3.** Overlap between fROIs defined using data from different runs of the same experiment (E1/E1, E2/E2, E3/E3) or using data from different experiments (E2/E3). Most cortical fROIs have lower overlap values for fROIs defined using data from different experiments compared to within-experiment overlap.

## 5.7.4 Supplemental tables

**Supplemental table 5-1.** Mixed-effect linear regression results for individual semantic fROIs. P values are FDR-corrected for the number of regions (n=11). Sem – semantic task, Perc – perceptual task, Sent – sentences (critical experiments), Pic – pictures (critical experiments), S – sentence reading (language localizer), N – nonword reading (language localizer), H – hard working memory task (multiple demand localizer), E – easy working memory task (multiple demand localizer). E1, E2, E3 – critical experiments.

fROI	Regression Term	Beta	SE	p value
F2	Intercept	2.51	0.14	<.001 ***
	S-N	0.2	0.2	0.484
	H-E	0.48	0.23	0.226
	Task Sem>Perc	3.26	0.29	<.001 ***
	StimType Sent>Pic	0.06	0.29	0.971
	Task:StimType	0.32	0.29	0.951
	Sent_Sem>S	1.51	0.2	<.001 ***
	SemTask>WMTask	3.23	0.33	<.001 ***
	E2>E3	0.82	0.17	<.001 ***
	E1>E3	-0.49	0.21	0.04 *
	S>N:E2>E3	-0.06	0.29	0.946
	H-E:E2>E3	0.04	0.23	0.994
	Sem>Perc:E2>E3	1.39	0.41	0.001 **
	Sent>Pic:E2>E3	-0.46	0.41	0.616
	Task:StimType:E2>E3	0.04	0.41	0.992
	Sent_Sem>S:E2>E3	0.81	0.29	0.009 **
	SemTask>WMTask:E2>E3	0.99	0.33	0.004 **
	S-N:E1>E3	-0.21	0.31	0.843
	Sem>Perc:E1>E3	-0.69	0.43	0.178
	Sent>Pic:E1>E3	0.13	0.43	0.802
	Task:StimType:E1>E3	0.23	0.43	0.915
	Sent_Sem>S:E1>E3	-0.35	0.31	0.349
F1	Intercept	2.09	0.16	<.001 ***
	S-N	0.18	0.2	0.484
	H-E	0.16	0.23	0.609
	Task Sem>Perc	3.44	0.28	<.001 ***
	StimType Sent>Pic	0.02	0.28	0.971
	Task:StimType	0.22	0.28	0.951
	Sent_Sem>S	1.39	0.2	<.001 ***
	SemTask>WMTask	4.04	0.32	<.001 ***
	E2>E3	0.94	0.18	<.001 ***
	E1>E3	-0.28	0.24	0.245
	S>N:E2>E3	-0.31	0.28	0.946
	H-E:E2>E3	0.08	0.23	0.994
	Sem>Perc:E2>E3	1.9	0.4	<.001 ***
	Sent>Pic:E2>E3	-0.32	0.4	0.616
	Task:StimType:E2>E3	-0.11	0.4	0.992
	Sent_Sem>S:E2>E3	1.17	0.28	<.001 ***
	SemTask>WMTask:E2>E3	1.19	0.32	<.001 ***
	S-N:E1>E3	-0.14	0.3	0.843
	Sem>Perc:E1>E3	-0.94	0.42	0.055
	Sent>Pic:E1>E3	-0.15	0.42	0.802
	Task:StimType:E1>E3	0.32	0.42	0.915
	Sent_Sem>S:E1>E3	-0.28	0.3	0.377
TP3	Intercept	1.04	0.11	<.001 ***

	S-N	0.52	0.17	0.011	*
	H-E	-0.42	0.19	0.226	
	Task Sem>Perc	2.34	0.24	<.001	***
	StimType Sent>Pic	-0.26	0.24	0.583	
	Task:StimType	-0.01	0.24	0.951	
	Sent_Sem>S	0.6	0.17	<.001	***
	SemTask>WMTask	3.63	0.27	<.001	***
	E2>E3	0.1	0.14	0.459	
	E1>E3	-0.46	0.17	0.026	*
	S>N:E2>E3	0.03	0.24	0.946	
	H-E:E2>E3	-0.01	0.19	0.994	
	Sem>Perc:E2>E3	0.58	0.34	0.101	
	Sent>Pic:E2>E3	-0.36	0.34	0.616	
	Task:StimType:E2>E3	0.16	0.34	0.992	
	Sent_Sem>S:E2>E3	0.17	0.24	0.47	
	SemTask>WMTask:E2>E3	-0.02	0.27	0.939	
	S-N:E1>E3	0.1	0.25	0.843	
	Sem>Perc:E1>E3	-1.02	0.36	0.022	*
	Sent>Pic:E1>E3	0.13	0.36	0.802	
	Task:StimType:E1>E3	0.12	0.36	0.915	
	Sent_Sem>S:E1>E3	-0.36	0.25	0.348	
TP2	Intercept	1.44	0.1	<.001	***
	S-N	0.72	0.15	<.001	***
	H-E	-0.27	0.17	0.311	
	Task Sem>Perc	2.38	0.21	<.001	***
	StimType Sent>Pic	0.57	0.21	0.038	*
	Task:StimType	0.3	0.21	0.951	
	Sent_Sem>S	0.5	0.15	0.001	**
	SemTask>WMTask	4.06	0.24	<.001	***
	E2>E3	0.62	0.12	<.001	***
	E1>E3	-0.52	0.14	0.004	**
	S>N:E2>E3	0.06	0.21	0.946	
	H-E:E2>E3	-0.06	0.17	0.994	
	Sem>Perc:E2>E3	1.22	0.3	<.001	***
	Sent>Pic:E2>E3	-0.35	0.3	0.616	
	Task:StimType:E2>E3	0.05	0.3	0.992	
	Sent_Sem>S:E2>E3	0.62	0.21	0.009	**
	SemTask>WMTask:E2>E3	0.88	0.24	<.001	***
	S-N:E1>E3	-0.28	0.22	0.843	
	Sem>Perc:E1>E3	-0.92	0.32	0.022	*
	Sent>Pic:E1>E3	-0.19	0.32	0.802	
	Task:StimType:E1>E3	0.12	0.32	0.915	
	Sent_Sem>S:E1>E3	-0.29	0.22	0.348	
Cer1	Intercept	1.1	0.1	<.001	***
	S-N	0.14	0.1	0.332	
	H-E	0.14	0.12	0.365	
	Task Sem>Perc	1.42	0.14	<.001	***
	StimType Sent>Pic	-0.09	0.14	0.842	
	Task:StimType	0.06	0.14	0.951	
	Sent_Sem>S	0.54	0.1	<.001	***
	SemTask>WMTask	1.33	0.16	<.001	***
	E2>E3	0.36	0.1	<.001	***
	E1>E3	-0.18	0.15	0.243	
	S>N:E2>E3	0.02	0.14	0.946	
	H-E:E2>E3	0.01	0.12	0.994	
	Sem>Perc:E2>E3	0.65	0.2	0.002	**
	Sent>Pic:E2>E3	-0.11	0.2	0.654	

	Task:StimType:E2>E3	0.09	0.2	0.992	
	Sent_Sem>S:E2>E3	0.36	0.14	0.017	*
	SemTask>WMTTask:E2>E3	0.46	0.16	0.006	**
	S-N:E1>E3	-0.04	0.15	0.843	
	Sem>Perc:E1>E3	-0.46	0.21	0.055	
	Sent>Pic:E1>E3	-0.33	0.21	0.802	
	Task:StimType:E1>E3	-0.04	0.21	0.915	
	Sent_Sem>S:E1>E3	-0.38	0.15	0.121	
TP4	Intercept	1.53	0.18	<.001	***
	S-N	0.06	0.24	0.801	
	H-E	0.19	0.28	0.609	
	Task Sem>Perc	2.89	0.34	<.001	***
	StimType Sent>Pic	-0.5	0.34	0.382	
	Task:StimType	-0.17	0.34	0.951	
	Sent_Sem>S	1.04	0.24	<.001	***
	SemTask>WMTTask	1.47	0.39	<.001	***
	E2>E3	0.66	0.21	0.002	**
	E1>E3	-0.32	0.26	0.243	
	S>N:E2>E3	-0.02	0.34	0.946	
	H-E:E2>E3	-0.02	0.28	0.994	
	Sem>Perc:E2>E3	1.45	0.48	0.004	**
	Sent>Pic:E2>E3	-0.84	0.48	0.616	
	Task:StimType:E2>E3	0.11	0.48	0.992	
	Sent_Sem>S:E2>E3	0.74	0.34	0.037	*
	SemTask>WMTTask:E2>E3	1.03	0.39	0.01	**
	S-N:E1>E3	0.15	0.36	0.843	
	Sem>Perc:E1>E3	-0.75	0.51	0.192	
	Sent>Pic:E1>E3	-0.32	0.51	0.802	
	Task:StimType:E1>E3	0.23	0.51	0.915	
	Sent_Sem>S:E1>E3	-0.62	0.36	0.319	
F3	Intercept	1.13	0.2	<.001	***
	S-N	-0.07	0.23	0.801	
	H-E	0.14	0.27	0.658	
	Task Sem>Perc	2.85	0.33	<.001	***
	StimType Sent>Pic	-0.15	0.33	0.89	
	Task:StimType	-0.03	0.33	0.951	
	Sent_Sem>S	0.94	0.23	<.001	***
	SemTask>WMTTask	3.26	0.38	<.001	***
	E2>E3	0.78	0.22	<.001	***
	E1>E3	-0.58	0.29	0.077	
	S>N:E2>E3	0.03	0.33	0.946	
	H-E:E2>E3	-0.05	0.27	0.994	
	Sem>Perc:E2>E3	2.41	0.47	<.001	***
	Sent>Pic:E2>E3	-0.28	0.47	0.654	
	Task:StimType:E2>E3	0.3	0.47	0.992	
	Sent_Sem>S:E2>E3	0.77	0.33	0.028	*
	SemTask>WMTTask:E2>E3	1.34	0.38	<.001	***
	S-N:E1>E3	0.18	0.35	0.843	
	Sem>Perc:E1>E3	-1.28	0.49	0.03	*
	Sent>Pic:E1>E3	-0.5	0.49	0.802	
	Task:StimType:E1>E3	0.06	0.49	0.915	
	Sent_Sem>S:E1>E3	-0.76	0.35	0.165	
TP1	Intercept	1.95	0.12	<.001	***
	S-N	-0.2	0.12	0.332	
	H-E	0.23	0.14	0.311	
	Task Sem>Perc	1.65	0.17	<.001	***
	StimType Sent>Pic	-0.81	0.17	<.001	***

	Task:StimType	-0.18	0.17	0.951	
	Sent_Sem>S	0.49	0.12	<.001	***
	SemTask>WMTask	2.31	0.2	<.001	***
	E2>E3	0.75	0.12	<.001	***
	E1>E3	-0.36	0.18	0.077	
	S>N:E2>E3	-0.04	0.18	0.946	
	H-E:E2>E3	-0.05	0.14	0.994	
	Sem>Perc:E2>E3	0.84	0.25	0.001	**
	Sent>Pic:E2>E3	-0.19	0.25	0.616	
	Task:StimType:E2>E3	-0.02	0.25	0.992	
	Sent_Sem>S:E2>E3	0.69	0.18	<.001	***
	SemTask>WMTask:E2>E3	1.02	0.2	<.001	***
	S-N:E1>E3	-0.1	0.19	0.843	
	Sem>Perc:E1>E3	-0.33	0.26	0.254	
	Sent>Pic:E1>E3	0.08	0.26	0.802	
	Task:StimType:E1>E3	0.2	0.26	0.915	
	Sent_Sem>S:E1>E3	-0.21	0.19	0.349	
O1	Intercept	2.96	0.23	<.001	***
	S-N	-0.25	0.25	0.484	
	H-E	0.35	0.28	0.365	
	Task Sem>Perc	1.83	0.35	<.001	***
	StimType Sent>Pic	-0.01	0.35	0.971	
	Task:StimType	-0.04	0.35	0.951	
	Sent_Sem>S	0.09	0.25	0.705	
	SemTask>WMTask	5.1	0.4	<.001	***
	E2>E3	1.43	0.24	<.001	***
	E1>E3	-0.86	0.34	0.036	*
	S>N:E2>E3	0.1	0.35	0.946	
	H-E:E2>E3	0	0.28	0.994	
	Sem>Perc:E2>E3	0.29	0.49	0.551	
	Sent>Pic:E2>E3	-0.15	0.49	0.762	
	Task:StimType:E2>E3	-0.14	0.49	0.992	
	Sent_Sem>S:E2>E3	0.56	0.35	0.118	
	SemTask>WMTask:E2>E3	1.34	0.4	0.001	**
	S-N:E1>E3	-0.27	0.37	0.843	
	Sem>Perc:E1>E3	-0.37	0.52	0.522	
	Sent>Pic:E1>E3	0.28	0.52	0.802	
	Task:StimType:E1>E3	0.06	0.52	0.915	
	Sent_Sem>S:E1>E3	-0.07	0.37	0.843	
Cer2	Intercept	0.46	0.05	<.001	***
	S-N	0.03	0.08	0.787	
	H-E	-0.04	0.09	0.682	
	Task Sem>Perc	1.07	0.11	<.001	***
	StimType Sent>Pic	0.07	0.11	0.842	
	Task:StimType	-0.06	0.11	0.951	
	Sent_Sem>S	0.34	0.08	<.001	***
	SemTask>WMTask	1.17	0.12	<.001	***
	E2>E3	0.3	0.06	<.001	***
	E1>E3	-0.19	0.07	0.026	*
	S>N:E2>E3	-0.01	0.11	0.946	
	H-E:E2>E3	-0.01	0.09	0.994	
	Sem>Perc:E2>E3	0.8	0.15	<.001	***
	Sent>Pic:E2>E3	-0.12	0.15	0.616	
	Task:StimType:E2>E3	0	0.15	0.992	
	Sent_Sem>S:E2>E3	0.32	0.11	0.008	**
	SemTask>WMTask:E2>E3	0.55	0.12	<.001	***
	S-N:E1>E3	0.02	0.11	0.843	

	Sem>Perc:E1>E3	-0.41	0.16	0.03	*
	Sent>Pic:E1>E3	-0.13	0.16	0.802	
	Task:StimType:E1>E3	0.08	0.16	0.915	
	Sent_Sem>S:E1>E3	-0.17	0.11	0.348	
O2	Intercept	2.71	0.19	<.001	***
	S-N	-0.31	0.21	0.332	
	H-E	0.29	0.24	0.365	
	Task Sem>Perc	1.59	0.29	<.001	***
	StimType Sent>Pic	-0.72	0.29	0.051	
	Task:StimType	-0.04	0.29	0.951	
	Sent_Sem>S	0.15	0.21	0.502	
	SemTask>WMTask	4.43	0.34	<.001	***
	E2>E3	1.23	0.2	<.001	***
	E1>E3	-0.49	0.29	0.129	
	S>N:E2>E3	0.15	0.29	0.946	
	H-E:E2>E3	-0.08	0.24	0.994	
	Sem>Perc:E2>E3	0.39	0.41	0.384	
	Sent>Pic:E2>E3	0.35	0.41	0.616	
	Task:StimType:E2>E3	0.2	0.41	0.992	
	Sent_Sem>S:E2>E3	1.02	0.29	0.002	**
	SemTask>WMTask:E2>E3	2	0.34	<.001	***
	S-N:E1>E3	-0.24	0.31	0.843	
	Sem>Perc:E1>E3	-0.05	0.44	0.912	
	Sent>Pic:E1>E3	0.11	0.44	0.802	
	Task:StimType:E1>E3	-0.16	0.44	0.915	
	Sent_Sem>S:E1>E3	-0.33	0.31	0.356	

**Supplemental table 5-2.** Mixed-effect linear regression results for individual semantic fROIs for each critical experiment. P values are FDR-corrected for the number of regions (n=11) and experiments (n=3). Same abbreviations as in **Supplemental table 5-1**.

fROI	Experiment	Regression Term	Beta	SE	p value
F1	Experiment 1	(Intercept)	0.73	0.42	0.182
		TaskSem	0.98	0.32	0.009 **
		StimTypeSent	-0.34	0.13	0.052
	Experiment 2	TaskSem:StimTypeSent	0.54	0.16	0.055
		(Intercept)	0.41	0.28	0.248
		TaskSem	2.61	0.5	<.001 ***
F1	Experiment 3	StimTypeSent	-0.21	0.13	0.198
		TaskSem:StimTypeSent	0.12	0.16	0.797
		(Intercept)	0.04	0.16	0.851
	Experiment 1	TaskSem	1.24	0.3	<.001 ***
		StimTypeSent	0.24	0.12	0.129
		TaskSem:StimTypeSent	0.01	0.16	0.972
F2	Experiment 1	(Intercept)	0.81	0.17	0.001 **
		TaskSem	1.01	0.3	0.007 **
		StimTypeSent	-0.18	0.1	0.155
	Experiment 2	TaskSem:StimTypeSent	0.55	0.14	0.009 **
		(Intercept)	1.07	0.24	0.001 **
		TaskSem	2.15	0.47	<.001 ***
F2	Experiment 3	StimTypeSent	-0.38	0.11	0.011 *
		TaskSem:StimTypeSent	0.36	0.14	0.112
		(Intercept)	0.85	0.21	0.001 **
	Experiment 1	TaskSem	1.25	0.35	0.002 **
		StimTypeSent	0.17	0.14	0.354
		TaskSem:StimTypeSent	0.05	0.19	0.955
F3	Experiment 1	(Intercept)	-0.07	0.36	0.884
		TaskSem	0.77	0.37	0.053
		StimTypeSent	-0.34	0.18	0.153
	Experiment 2	TaskSem:StimTypeSent	0.03	0.24	0.972
		(Intercept)	-0.65	0.26	0.057
		TaskSem	2.5	0.79	0.008 **
F3	Experiment 3	StimTypeSent	-0.35	0.21	0.198
		TaskSem:StimTypeSent	0.27	0.27	0.722
		(Intercept)	-0.23	0.35	0.692
	Experiment 1	TaskSem	1.06	0.29	0.002 **
		StimTypeSent	0.51	0.18	0.037 *
		TaskSem:StimTypeSent	-0.39	0.26	0.441
TP1	Experiment 1	(Intercept)	1.12	0.15	<.001 ***
		TaskSem	0.65	0.16	0.003 **
		StimTypeSent	-0.37	0.08	0.004 **
	Experiment 2	TaskSem:StimTypeSent	0.02	0.06	0.955
		(Intercept)	1.38	0.18	<.001 ***
		TaskSem	1.35	0.24	<.001 ***
TP1	Experiment 3	StimTypeSent	-0.4	0.1	0.008 **
		TaskSem:StimTypeSent	-0.2	0.08	0.148
		(Intercept)	1.25	0.21	<.001 ***
	Experiment 1	TaskSem	0.75	0.18	<.001 ***
		StimTypeSent	-0.18	0.17	0.412
		TaskSem:StimTypeSent	-0.35	0.16	0.148
TP2	Experiment 1	(Intercept)	0.19	0.18	0.46
		TaskSem	0.53	0.19	0.018 *
		StimTypeSent	-0.02	0.13	0.932

		TaskSem:StimTypeSent	0.42	0.1	0.007	**
TP2	Experiment 2	(Intercept)	0.26	0.27	0.485	
		TaskSem	1.63	0.33	<.001	***
		StimTypeSent	-0.07	0.15	0.739	
TP2	Experiment 3	TaskSem:StimTypeSent	0.36	0.13	0.107	
		(Intercept)	0.05	0.12	0.738	**
		TaskSem	0.98	0.25	0.002	**
		StimTypeSent	0.49	0.15	0.011	*
TP3	Experiment 1	TaskSem:StimTypeSent	0.13	0.13	0.722	
		(Intercept)	-0.02	0.17	0.905	**
		TaskSem	0.61	0.16	0.003	**
		StimTypeSent	-0.12	0.1	0.372	
TP3	Experiment 2	TaskSem:StimTypeSent	0.11	0.12	0.759	
		(Intercept)	-0.13	0.25	0.729	
		TaskSem	1.39	0.28	<.001	***
		StimTypeSent	-0.39	0.14	0.037	*
TP3	Experiment 3	TaskSem:StimTypeSent	0.15	0.13	0.647	
		(Intercept)	-0.11	0.23	0.729	
		TaskSem	1.54	0.35	<.001	***
		StimTypeSent	0.14	0.15	0.477	
TP4	Experiment 1	TaskSem:StimTypeSent	-0.3	0.21	0.441	
		(Intercept)	0.35	0.19	0.163	
		TaskSem	1.04	0.22	<.001	***
		StimTypeSent	-0.44	0.11	0.008	**
TP4	Experiment 2	TaskSem:StimTypeSent	0.06	0.15	0.955	
		(Intercept)	0.35	0.5	0.681	
		TaskSem	2.2	0.47	<.001	***
		StimTypeSent	-0.64	0.24	0.046	*
TP4	Experiment 3	TaskSem:StimTypeSent	-0.06	0.2	0.955	
		(Intercept)	-0.14	0.24	0.729	
		TaskSem	1.35	0.29	<.001	***
		StimTypeSent	0.58	0.17	0.008	**
Cer1	Experiment 1	TaskSem:StimTypeSent	-0.5	0.23	0.148	
		(Intercept)	0.54	0.24	0.086	
		TaskSem	0.47	0.14	0.004	**
		StimTypeSent	-0.22	0.19	0.407	
Cer1	Experiment 2	TaskSem:StimTypeSent	0.02	0.1	0.972	
		(Intercept)	0.43	0.15	0.033	*
		TaskSem	0.96	0.19	<.001	***
		StimTypeSent	-0.17	0.09	0.153	
Cer1	Experiment 3	TaskSem:StimTypeSent	0.15	0.1	0.441	
		(Intercept)	0.28	0.16	0.176	
		TaskSem	0.62	0.17	0.002	**
		StimTypeSent	0.17	0.09	0.153	
Cer2	Experiment 1	TaskSem:StimTypeSent	0	0.13	0.972	
		(Intercept)	-0.05	0.09	0.729	
		TaskSem	0.32	0.13	0.025	*
		StimTypeSent	-0.04	0.05	0.497	
Cer2	Experiment 2	TaskSem:StimTypeSent	0.02	0.06	0.955	
		(Intercept)	-0.17	0.1	0.176	
		TaskSem	0.97	0.18	<.001	***
		StimTypeSent	0.01	0.08	0.932	
Cer2	Experiment 3	TaskSem:StimTypeSent	-0.06	0.11	0.891	
		(Intercept)	-0.13	0.09	0.248	
		TaskSem	0.42	0.15	0.009	**
		StimTypeSent	0.23	0.09	0.04	*
		TaskSem:StimTypeSent	-0.14	0.12	0.647	

O1	Experiment 1	(Intercept)	1.3	0.2	<.001	***
		TaskSem	0.72	0.17	<.001	***
		StimTypeSent	0.13	0.17	0.541	
		TaskSem:StimTypeSent	0.01	0.19	0.972	
O1	Experiment 2	(Intercept)	2.86	0.51	<.001	***
		TaskSem	1.16	0.34	0.004	**
		StimTypeSent	0.01	0.26	0.96	
		TaskSem:StimTypeSent	-0.19	0.25	0.79	
O1	Experiment 3	(Intercept)	1.95	0.32	<.001	***
		TaskSem	0.93	0.34	0.013	*
		StimTypeSent	-0.1	0.23	0.745	
		TaskSem:StimTypeSent	0.05	0.29	0.972	
O2	Experiment 1	(Intercept)	1.55	0.18	<.001	***
		TaskSem	0.87	0.26	0.004	**
		StimTypeSent	-0.21	0.24	0.497	
		TaskSem:StimTypeSent	-0.2	0.26	0.79	
O2	Experiment 2	(Intercept)	2.78	0.37	<.001	***
		TaskSem	0.91	0.27	0.006	**
		StimTypeSent	-0.26	0.2	0.322	
		TaskSem:StimTypeSent	0.16	0.14	0.667	
O2	Experiment 3	(Intercept)	1.9	0.27	<.001	***
		TaskSem	0.67	0.28	0.023	*
		StimTypeSent	-0.55	0.22	0.052	
		TaskSem:StimTypeSent	-0.08	0.25	0.955	

# Chapter 6

## Conclusion

### 6.1 Thesis summary

My thesis work supports the hypothesis that the mechanisms underlying language processing are distinct from mechanisms underlying other cognitive functions.

In Chapter 2, I examined brain responses to computer code, an input whose structure is very similar to natural language. Based on this similarity, I hypothesized that the language network might contribute to computer code comprehension. My collaborators and I conducted two experiments testing the brain response to two different programming languages: Python and ScratchJr. A key feature of our design was a distinction between computer code comprehension (the process of extracting meaning from code based on semantic and syntactic rules, which I hypothesized to be the most language-like) and analyzing program content (mentally simulating the piece of code to determine the output). The language network responded weakly to computer code comprehension for Python and not at all for ScratchJr; however, another set of brain regions, known as the multiple demand network, responded robustly to computer code comprehension in both programming languages. I concluded that computer programmers recruit their multiple demand brain regions to process computer code, and the language regions remain functionally specialized for language.

In Chapter 3 and Chapter 4, I examined the role of the language network in processing non-linguistic stimuli (pictures) during tasks that required accessing their semantic content. I found no response in the language regions to object semantics (e.g., deciding whether the animal presented on the screen is dangerous) and weak response to combinatorial event semantics (e.g., deciding whether a criminal arresting a cop is a plausible event). I furthermore showed that the language network's engagement in event semantics is not necessary: individuals with global aphasia, whose language network

function is severely impaired, can nevertheless perform an event semantics task at a level close to that of neurotypical controls. Finally, in Chapter 5, I showed that another set of regions, distinct from the language network, does respond strongly to both sentences and pictures during semantic tasks, which constitutes further evidence that linguistic and semantic processing are spatially segregated in the human brain.

## 6.2 Language processing outside of the language network

In preceding chapters, I often equate linguistic processing with activity in the language network. Is this a reasonable approach? After all, activity in other regions might contribute to aspects of language processing not covered by the language localizer, such as discourse or pragmatics.

In this section, I show that, indeed, real-life language use engages not only the language network, but also other brain networks. Those networks support specific aspects of language, including error correction, reasoning about the effect of an utterance on other people, and connecting information across sentences. These aspects might be plausibly construed as essential parts of language processing; however, evidence to date shows that the networks supporting them process both linguistic and non-linguistic inputs, meaning that, unlike the core language network, they cannot be considered language-specific. Thus, we should view language as a cognitive faculty relying on a specialized module (the language network) that is tightly integrated with — and dependent upon — multiple downstream processing mechanisms.

### 6.2.1 Multiple demand network: hard language and external tasks

In Chapter 2, I have described the multiple demand network — a set of domain-general brain regions that are active during diverse cognitively challenging tasks (Cole & Schneider, 2007; Duncan, 2010; Duncan & Owen, 2000; Hugdahl et al., 2015; Shashidhara, Mitchell, et al., 2019). Multiple demand and language networks are robustly dissociable, as evident from both neuroimaging studies (Assem, Blank, et al., 2020; Blank et al., 2014a; Diachek et al., 2020; Fedorenko et al., 2013; Mineroff et al., 2018; Shain et al., 2020) and

studies of individuals with brain damage (Gläscher et al., 2010; MacGregor et al., 2022; Woolgar et al., 2010, 2018a).

The multiple demand network contributes to language use in several ways. One, it gets engaged in cases when language use is effortful, such as comprehension of acoustically degraded, ambiguous, or syntactically unusual utterances (e.g., Kuperberg et al., 2003; MacGregor et al., 2022; Nieuwland et al., 2012; Wild et al., 2012; for reviews, see Novick et al., 2010; Fedorenko, 2014). Two, it is recruited in the presence of an external task, such as answering a question, deciding whether a given word was present in a sentence, or naming a picture on demand (Cheung et al., 2020; Diacheck et al., 2020; Hu, Small et al., 2021). Three, it is required for processing certain types of content, such as math statements or logic problems, even when the problem is presented linguistically (Amalric & Dehaene, 2016, 2019; Monti et al., 2007, 2009, 2012). Thus, real-life language use involves a dynamic interplay between language and multiple demand networks.

That said, the multiple demand network is not dependent on language regions: if the input to a challenging tasks is non-linguistic, it activates the multiple demand network, but the language regions do not get engaged (e.g., Fedorenko et al., 2013; see also Chapters 2 and 3 of this thesis). In contrast to the view that language and reasoning rely on a shared processing system, we see that language processing interacts with domain-general reasoning mechanisms in specific, well-defined ways.

### 6.2.2 Theory of mind network: pragmatics and social inference

A wealth of neuroscientific evidence shows that the human brain has dedicated machinery for processing social information (e.g., Adolphs, 1999, 2009; Deen et al., 2015; Isik et al., 2017; Lee Masson & Isik, 2021; Saxe, 2006; Tarhan & Konkle, 2020; Walbrin et al., 2018). Perhaps the most relevant to our current discussion is the theory of mind network, a set of brain regions that are engaged when their owner is trying to infer another person's mental state (Fletcher et al., 1995; Gallagher et al., 2000; Jacoby et al., 2016; Saxe, Moran, et al., 2006; Saxe & Kanwisher, 2003; Saxe & Powell, 2006). Similar to the multiple demand network, the theory of mind network constitutes a separate

cognitive module that is anatomically and functionally distinct from the language network (Apperly et al., 2006; Deen et al., 2015; Paunov et al., 2022; Shain, Paunov, X. Chen et al., 2022; Varley et al., 2001; Varley & Siegal, 2000; Willems et al., 2011). However, it too contributes to specific aspects of language understanding. First, it is engaged during nonliteral language comprehension, such as jokes, sarcasm, indirect speech, and conversational implicature (Feng et al., 2017, 2021; Hauptman et al., 2022; Jang et al., 2013; Spotorno et al., 2012; van Ackeren et al., 2012; for a review, see Hagoort & Levinson, 2014) — in other words, in situations where understanding the meaning of an utterance requires inferring the intentions of the speaker. Second, just like other functionally specialized brain modules, it is engaged when processing semantic content that is specifically related to its domain: narratives that require inferring the mental state of the characters engage the theory of mind network regardless of whether the actual stimuli are texts or as movies (Jacoby et al., 2016; Paunov et al., 2022), and texts that require inferring the characters' intentions evoke greater activity than those that do not (Ferstl & von Cramon, 2002; Fletcher et al., 1995; Saxe & Powell, 2006). Thus, successful language understanding and use rely on our broader, non-language-specific social inference skills.

### 6.2.3 The default mode network: narratives and situation modeling

As discussed in Chapter 1, the language network is sensitive to linguistic regularities that apply to units of different 'grain size', from phonemes to morphemes to words to phrases and clauses (Blank & Fedorenko, 2020; Fedorenko, Behr, et al., 2011; Fedorenko et al., 2020a; Fedorenko, Nieto-Castañon, et al., 2012; T. I. Regev et al., 2021). However, it is not sensitive to structure above the clause level (Blank & Fedorenko, 2020; Caucheteux et al., 2021; Jacoby & Fedorenko, 2020; Lerner et al., 2011; Yeshurun, Swanson, et al., 2017; Yeshurun, Nguyen, et al., 2017), suggesting that it does not carry out the computations required to integrate information across sentences.

What brain network integrates individual sentences to build an overall situation model of the described events? Neuroimaging evidence points to the network that includes medial parietal and prefrontal regions, often referred to as the default mode (or default) network (Andrews-Hanna, 2012; Buckner et al., 2008; Buckner & DiNicola, 2019; Greicius

et al., 2003). This network is recruited for coherent texts more than for disconnected sentences (e.g., Ferstl et al., 2008; Ferstl & von Cramon, 2002; Simony et al., 2016), encodes specific event schemas, such as going to a restaurant (Baldassano et al., 2018), and has a long temporal integration window, which enables it to integrate information over multiple sentences (e.g., Blank & Fedorenko, 2020; Caucheteux et al., 2021; Jain et al., 2020; Jain & Huth, 2018; Lerner et al., 2011). Just like the multiple demand and the theory of mind networks, the default mode network is not language-specific: it exhibits similar information processing signatures when people process both verbal and nonverbal narratives (Baldassano et al., 2017, 2018).

Recent work has suggested that the default mode network may consist of two distinct interdigitated sub-networks (Braga et al., 2020; Deen & Freiwald, 2022; DiNicola et al., 2020). One of them appears to correspond to the theory of mind network (Saxe & Kanwisher, 2003) discussed in [Section 6.2.2](#). The exact contributions of the other sub-network remain debated, with different proposals linking its functions to episodic projection (placing oneself into the past, when remembering things, or into the future, when imagining things; Buckner et al., 2008), scene construction and situation modeling (Hassabis & Maguire, 2009) or spatial cognition in general (Deen & Freiwald, 2022). Based on the evidence so far, two options are likely: either discourse-level information is processed by the default mode network, with theory of mind network getting engaged when social inference is required, or the information is passed in parallel to the two networks, one of which constructs a physical situation model, and the other constructs a social situation model.

In either case, we can conclude that, in the human brain, the task of tracking information over multiple sentences is allocated to a system (or systems) that is not only separate from the core language network but is also shared between linguistic and non-linguistic inputs. The fact that people's memory for precise linguistic forms is relatively poor (Lombardi & Potter, 1992; Potter & Lombardi, 1990, 1998; cf. Gurevich et al., 2010) aligns well with the idea that the language input, once processed, is then passed on to downstream, non-language-specific regions, which are tasked with the goal of integrating sentence-level information into an overall situation model.

#### **6.2.4 Semantic demand fROIs and other brain regions**

Finally, in Chapter 5, I describe a set of regions that are active during a semantic task during both sentence and picture processing. Although we still know little, their involvement pattern so far resembles the pattern described for other networks discussed in this section: they are not language-specific (or biased toward language inputs) but contribute to language use in certain situations. Thus, it appears that the language network interfaces with additional neural machinery dedicated to semantic processing in the presence of an external task (i.e., semantic demand). Whether the semantic demand regions form a network or can be broken down into different groups remains to be established.

This list of extra-linguistic brain regions/networks is, of course, not exhaustive. Subcortical brain structures that may contribute the language processing include the hippocampus (Blank, Duff, et al., 2016; Duff & Brown-Schmidt, 2012, 2017) and the basal ganglia (e.g., Bohsali & Crosson, 2016; Crosson et al., 2007; Damasio, 1983). Whole-brain encoding models of brain activity reveal language-evoked responses in a multitude of cortical regions outside the classical language areas (Caucheteux & King, 2022; Huth et al., 2016; T. M. Mitchell et al., 2008; Pereira et al., 2018; Schrimpf et al., 2021), many of which might process domain-specific information conveyed via language. The overall picture, however, remains the same: contextualized, real-life language processing requires an interaction between the core language network and multiple other brain regions, which contribute specific aspects to language understanding and use but are not themselves dependent on language regions.

### **6.3 “Language as a foundation for human cognition” revisited**

How does the neuroscience-informed view of language processing contribute to our understanding of the relationship between language and other cognitive functions? First, it enables us to build a cognitive ontology (Lenartowicz et al., 2010; Poldrack et al., 2011; C. J. Price & Friston, 2005) of basic processes that contribute to language understanding. We now know that syntax, lexical semantics, and combinatorial semantics rely on shared

processing mechanisms (see [Section 1.3](#)), but narrative-level processing and many aspects of pragmatics rely on separate cognitive modules (see [Section 6.2](#)). Second, once we know which cognitive/neural modules contribute to language processing, we can determine whether each of them is language-specific. We are therefore able to establish that the mechanisms dedicated to word- and sentence-level processing rely on a specialized set of brain regions, whereas the mechanisms dedicated to narrative-level and pragmatic processing rely on non-language-specific mechanisms. And finally, we can apply this knowledge to hypotheses describing the language-thought relationship by (a) using our knowledge of cognitive modules to better define what we mean by ‘language’ and (b) providing evidence in favor or against the interdependence between language and other aspects of human cognition.

In this thesis, I have shown that the system that carries out the core aspects of language processing (lexical semantics, syntax, combinatorial semantics, etc.) does not support other cognitive functions. I have tested domains that are most likely to rely on the same system as language: programming languages ([Chapter 2](#)) and pictures with semantic content during semantic tasks ([Chapter 3](#) and [Chapter 4](#)). The language network showed a preference for sentences over all other types of stimuli. Furthermore, in cases where I observed some response in the language network to non-linguistic input, I also found evidence against the claim that this response is essential to the process in question: for computer programming ([Chapter 2](#)), the activity was observed for only one of the two programming languages tested, and for event semantics ([Chapter 4](#)), the language network was not causally implicated in successful task performance.

The non-language-specific semantic demand regions ([Chapter 5](#)) do constitute evidence of shared processing mechanisms between linguistic and non-linguistic input; however, most of them are only active in the presence of an external task and do not respond to passive sentence reading. As such, they do not constitute part of the primary language processing pipeline. In fact, it would be counterintuitive to label these regions as responsible for ‘language’, when we have evidence of another system that gets activated for core aspects of language processing regardless of specific language, input modality, or presence of a task. Thus, the existence of amodal semantic demand regions is not

evidence that language underlies semantic processing, but rather that the language network interacts closely with other cognitive modules (as discussed in [Section 6.2](#)).

The “language as a foundation for thought” view posits that linguistic representations and/or computations are accessed during non-linguistic processing (e.g., Carruthers, 2002; Chomsky, 2007; Gauker, 2011; Hauser et al., 2002). We see that this is not the case: linguistic representations and computations, whether lexical or syntactic, reside within a functionally specialized module that remains untapped during other functions that require syntactic combination (music, math, computer programming), semantic combination (interpreting a visually presented event), or accessing semantic knowledge in general (semantic categorization). Moreover, the language network relies on other cognitive modules during real-life language use (e.g., the multiple demand network when the input is noisy or on the theory of mind network when the input requires non-literal interpretation), which goes against the view that it serves as a foundation for other functions. Thus, cognitive neuroscience provides a clear perspective on the nature of the relationship between language and broader human cognition.

## 6.4 Caveat: development is different

As mentioned in [Chapter 1](#), I focus on the relationship between linguistic and non-linguistic processing mechanisms in adult, fully developed human brains. The functional interdependence between language and other cognitive functions during development is a different matter. Evidence points to a tight coupling between learning number words and learning numbers themselves (e.g., Bull et al., 2011; Frank et al., 2008; Pitt et al., 2022), between language experience and the development of theory of mind (e.g., Moeller & Schick, 2006; Peterson, 2016; Peterson & Siegal, 1995; Richardson et al., 2020), and possibly even between language experience and situation modeling (Peterson & Slaughter, 2006). Thus, although the networks discussed in [Section 6.2](#) are independent from language in adulthood, their development does, at least to some extent, depend on the amount and content of the language input.

So far, evidence of a relationship between language and other cognitive functions during development has come primarily from cases of language deprivation in childhood (most common of which is deaf or hard of hearing children with delayed access to sign language and/or hearing aids) and from studies of cross-linguistic differences (e.g., numeric cognition in speakers of languages without exact number terms). Today, we have another putative source of information about the role of language in development: computational models. Artificial neural networks trained exclusively on language input acquire large amounts of information about the world (T. B. Brown et al., 2020; Chowdhery et al., 2022; Grand et al., 2022; Petroni et al., 2019; Tshitoyan et al., 2019; Unger & Fisher, 2021, among others), although their cognitive capabilities remain brittle (Elazar, Kassner, et al., 2021; Elazar, Zhang, et al., 2021; Ettinger, 2020; A. Patel et al., 2021; Schuster & Linzen, 2022; Talmor et al., 2020, among others). In general, large language models provide a fertile ground for dynamically simulating interactions and interdependencies between different cognitive modules and are likely to provide useful insights on the relationship between language and cognitive functions throughout development.

## 6.5 Implications for inner thought

Philosophers and linguists who claim language to be essential for thought often refer specifically to inner thought — a sequence of mental representations that people produce either unprompted (e.g., when daydreaming or reminiscing) or in response to an external need (e.g., when figuring out an answer to a question). The work described in this thesis primarily addresses the latter. I asked: do people rely on language processing mechanisms when accessing a specific, externally cued aspect of world knowledge? The results show that they do not, highlighting a dissociation between language and task-driven cognitive processing. But what about unprompted inner thought?

Neuroimaging evidence so far suggests that inner thought relies on mechanisms outside the language network. Activity in the language network is low at rest (when unprompted thinking is most likely to occur) compared to sentence reading or even to nonword reading (Fedorenko, Behr, et al., 2011; Fedorenko et al., 2010). Furthermore, activity patterns of different language regions are less correlated at rest than during story

listening (Blank et al., 2014a). In contrast, a set of regions known as the default mode network is most active at rest and gets deactivated during external tasks, suggesting that those regions do contribute to internally oriented thought (Andrews-Hanna, 2012; Buckner et al., 2008; Greicius et al., 2003).

Furthermore, evidence from individuals with global aphasia (see Section 1.3) indicates that they can continue to think and reason even in the absence of language. Internal thinking is (by definition) hard to evaluate externally, but evidence from individuals who have experienced aphasia suggests that inner thought can proceed even in the absence of language.

Tom Lubbock, an art critic, documented the deterioration of his language skill caused by a brain tumor (Lubbock, 2010, 2012). At some point, watching his language decline, he wrote: “I think that loss of speech, and of understanding of speech, and of understanding of writing, and of coherent writing – these losses will amount to the loss of my mind. <...> Mind means talking to oneself. There wouldn't be any secret mind surviving in me.” And yet, a few months later, he noted (with his wife now writing for him): “I find my brain is still busy, moving, thinking. I am surprised. My language to describe things in the world is very small, limited. My thoughts when I look at the world are vast, limitless and normal, same as they ever were.” Throughout this period, he continued to write art reviews, despite increasing difficulties with word retrieval. Even though his language skills were deteriorating, he still had thoughts about art that he wanted to express.

Researchers Lecours and Jeanette (1980) described the case of Brother John, a French monk who experienced transient periods of aphasia (affecting both production and comprehension) with no impairment in thought. Oftentimes, Brother John did not even realize he was undergoing an episode of aphasia unless he tried to read or listen to the radio. His planning skills remained intact: when he was experiencing an episode, he would indicate to his colleagues to call the researchers and would continuously test himself (by recording his speech or trying to write) at regular intervals until they arrived. Just as Tom Lubbock, Brother John could think about the world in all its complexity and vastness even in the absence of language.

Thus, as with many other cognitive capacities, neuroimaging studies and evidence from individuals with aphasia converge to show independence of language and inner thought.

Many would find this conclusion surprising, stating that they often experience an inner voice, which constantly packages their thoughts, memories and plans into a verbal frame. To them, language is obviously an important part of inner thought. However, the experience of language-like inner thought varies greatly between individuals: some think primarily in language, others do not experience inner speech at all, and others are somewhere in between (e.g., Heavey & Hurlburt, 2008; Hurlburt et al., 2013; Klinger & Cox, 1987; Roebuck & Lupyan, 2020). Thus, it appears that language can contribute but is not necessary for inner thought, just like it contributes to but is not necessary for event understanding (see [Chapter 4](#)).

For people who perceive their inner thought as highly verbal, losing the ability to use language might alter their conscious experience. For instance, a neuroscientist Jill Bolte Taylor experienced a stroke which, for some time, left her with aphasia (Taylor, 2006, 2008). She reports that, during the stroke: “my brain chatter, my left hemisphere brain chatter went totally silent. Just like someone took a remote control and pushed the mute button and — total silence.” She also reports changes in her sense of self and her ability to plan, but it is unclear whether they are linked to language or to other left-lateralized brain networks. The latter seems more likely given cases like that of Brother John, who experienced no impairment in his planning skills and, as far as we know, no changes in his sense of self.

The extent to which individual variability in inner thought affects the language-thought relationship is an open question subject to future research. One interesting conjecture is that a researcher’s subjective experience of their inner thoughts might influence their theorizing about the role of language in cognition. Perhaps people whose thoughts have more verbal content are more likely to claim that language is foundational for thought. Like Tom Lubbock, they might perceive their mind and their sense of self to be tied to words; unlike Tom Lubbock, they might never realize that, even in the absence of words, their thoughts would still be there.



# References

- Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences*, 3(12), 469–479.
- Adolphs, R. (2009). The Social Brain: Neural Basis of Social Knowledge. *Annual Review of Psychology*, 60, 693–716. <https://doi.org/10.1146/annurev.psych.60.110707.163514>
- Allamanis, M., Barr, E. T., Devanbu, P., & Sutton, C. (2018). A Survey of Machine Learning for Big Code and Naturalness. *ArXiv:1709.06182 [Cs]*. <http://arxiv.org/abs/1709.06182>
- Altshuler, D., Parsons, T., & Schwarzschild, R. (2019). *A course in semantics*.
- Amalric, M., & Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. *Proceedings of the National Academy of Sciences of the United States of America*, 113(18), 4909–4917. <https://doi.org/10.1073/pnas.1603205113>
- Amalric, M., & Dehaene, S. (2019). A distinct cortical network for mathematical knowledge in the human brain. *NeuroImage*, 189, 19–31. <https://doi.org/10.1016/j.neuroimage.2019.01.001>
- Amit, E., Hoeflin, C., Hamzah, N., & Fedorenko, E. (2017). An asymmetrical relationship between verbal and visual thinking: Converging evidence from behavior and fMRI. *NeuroImage*, 152, 619–627. <https://doi.org/10.1016/j.neuroimage.2017.03.029>
- Amunts, K., Lenzen, M., Friederici, A. D., Schleicher, A., Morosan, P., Palomero-Gallagher, N., & Zilles, K. (2010). Broca's Region: Novel Organizational Principles and Multiple Receptor Mapping. *PLoS Biology*, 8(9), e1000489. <https://doi.org/10.1371/journal.pbio.1000489>
- Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H. B., & Zilles, K. (1999). Broca's region revisited: Cytoarchitecture and intersubject variability. *The Journal of Comparative Neurology*, 412(2), 319–341. [https://doi.org/10.1002/\(sici\)1096-9861\(19990920\)412:2<319::aid-cne10>3.0.co;2-7](https://doi.org/10.1002/(sici)1096-9861(19990920)412:2<319::aid-cne10>3.0.co;2-7)
- Anderson, A. J., Binder, J. R., Fernandino, L., Humphries, C. J., Conant, L. L., Aguilar, M., Wang, X., Doko, D., & Raizada, R. D. S. (2017). Predicting Neural Activity Patterns Associated with Sentences Using a Neurobiologically Motivated Model of Semantic Representation. *Cerebral Cortex*, 27(9), 4379–4395. <https://doi.org/10.1093/cercor/bhw240>
- Anderson, A. J., Lalor, E. C., Lin, F., Binder, J. R., Fernandino, L., Humphries, C. J., Conant, L. L., Raizada, R. D. S., Grimm, S., & Wang, X. (2019). Multiple Regions of a Cortical Network Commonly Encode the Meaning of Words in Multiple Grammatical Positions of Read Sentences. *Cerebral Cortex (New York, N.Y.: 1991)*, 29(6), 2396–2411. <https://doi.org/10.1093/cercor/bhy110>
- Andrews-Hanna, J. R. (2012). The Brain's Default Network and Its Adaptive Role in Internal Mentation. *The Neuroscientist*, 18(3), 251–270. <https://doi.org/10.1177/1073858411403316>
- Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-Anatomic Fractionation of the Brain's Default Network. *Neuron*, 65(4), 550–562. <https://doi.org/10.1016/j.neuron.2010.02.005>
- Antonucci, S. M., & Reilly, J. (2008). Semantic Memory and Language Processing: A Primer. *Seminars in Speech and Language*, 29(01), 005–017. <https://doi.org/10.1055/s-2008-1061621>
- Apperly, I. A., Samson, D., Carroll, N., Hussain, S., & Humphreys, G. (2006). Intact first-and second-order false belief reasoning in a patient with severely impaired grammar. *Social Neuroscience*, 1(3–4), 334–348.
- Ardila, A., & Rosselli, M. (2002). Acalculia and Dyscalculia. *Neuropsychology Review*, 12(4), 179–231. <https://doi.org/10.1023/A:1021343508573>
- Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage*, 26(3), 839–851. <https://doi.org/10.1016/j.neuroimage.2005.02.018>
- Asoulin, E. (2016). Language as an instrument of thought. *Glossa: A Journal of General Linguistics*, 1(1), Article 1. <https://doi.org/10.5334/gjgl.34>

- Assem, M., Blank, I. A., Mineroff, Z., Ademoğlu, A., & Fedorenko, E. (2020). Activity in the fronto-parietal multiple-demand network is robustly associated with individual differences in working memory and fluid intelligence. *Cortex*, 131, 1–16. <https://doi.org/10.1016/j.cortex.2020.06.013>
- Assem, M., Glasser, M. F., Van Essen, D. C., & Duncan, J. (2020). A Domain-General Cognitive Core Defined in Multimodally Parcellated Human Cortex. *Cerebral Cortex*, 30(8), 4361–4380. <https://doi.org/10.1093/cercor/bhaa023>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Badre, D., Poldrack, R. A., Paré-Bagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable Controlled Retrieval and Generalized Selection Mechanisms in Ventrolateral Prefrontal Cortex. *Neuron*, 47(6), 907–918. <https://doi.org/10.1016/j.neuron.2005.07.023>
- Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, 95(3), 709–721.e5. <https://doi.org/10.1016/j.neuron.2017.06.041>
- Baldassano, C., Hasson, U., & Norman, K. A. (2018). Representation of Real-World Event Schemas during Narrative Perception. *Journal of Neuroscience*, 38(45), 9689–9699. <https://doi.org/10.1523/JNEUROSCI.0251-18.2018>
- Baldo, J. V., Bunge, S. A., Wilson, S. M., & Dronkers, N. F. (2010). Is relational reasoning dependent on language? A voxel-based lesion symptom mapping study. *Brain and Language*, 113(2), 59–64. <https://doi.org/10.1016/j.bandl.2010.01.004>
- Baldo, J. V., Paulraj, S. R., Curran, B. C., & Dronkers, N. F. (2015). Impaired reasoning and problem-solving in individuals with language impairment due to aphasia or language delay. *Frontiers in Psychology*, 6, 1523. <https://doi.org/10.3389/fpsyg.2015.01523>
- Baroni, M., Maguire, S., & Drabkin, W. (1983). The Concept of Musical Grammar. *Music Analysis*, 2(2), 175–208. <https://doi.org/10.2307/854248>
- Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, 59(1), 617–645. <https://doi.org/10.1146/annurev.psych.59.103006.093639>
- Basso, A., & Capitani, E. (1985). Spared musical abilities in a conductor with global aphasia and ideomotor apraxia. *Journal of Neurology, Neurosurgery, and Psychiatry*, 48(5), 407–412.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bautista, A., & Wilson, S. M. (2016). Neural responses to grammatically and lexically degraded speech. *Language, Cognition and Neuroscience*, 31(4), 567–574. <https://doi.org/10.1080/23273798.2015.1123281>
- Bedny, M., Dravida, S., & Saxe, R. (2014). Shindigs, brunches, and rodeos: The neural basis of event words. *Cognitive, Affective & Behavioral Neuroscience*, 14(3), 891–901. <https://doi.org/10.3758/s13415-013-0217-z>
- Bek, J., Blades, M., Siegal, M., & Varley, R. A. (2010). Language and spatial reorientation: Evidence from severe aphasia. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 36(3), 646–658. <https://doi.org/10.1037/a0018281>
- Bek, J., Blades, M., Siegal, M., & Varley, R. A. (2013). Dual-Task Interference in Spatial Reorientation: Linguistic and Nonlinguistic Factors. *Spatial Cognition & Computation*, 13(1), 26–49. <https://doi.org/10.1080/13875868.2011.590622>
- Bemis, D. K., & Pylkänen, L. (2013). Basic Linguistic Composition Recruits the Left Anterior Temporal Lobe and Left Angular Gyrus During Both Listening and Reading. *Cerebral Cortex*, 23(8), 1859–1873. <https://doi.org/10.1093/cercor/bhs170>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300. JSTOR.

- Benn, Y., Ivanova, A. A., Clark, O., Mineroff, Z., Seikus, C., Silva, J. S., Varley, R., & Fedorenko, E. (2021). No evidence for a special role of language in feature-based categorization. *BioRxiv*, 2021.03.18.436075. <https://doi.org/10.1101/2021.03.18.436075>
- Benn, Y., Wilkinson, I. D., Zheng, Y., Kadosh, K. C., Romanowski, C. A. J., Siegal, M., & Varley, R. (2013). Differentiating core and co-opted mechanisms in calculation: The neuroimaging of calculation in aphasia. *Brain and Cognition*, 82(3), 254–264. <https://doi.org/10.1016/j.bandc.2013.04.012>
- Bergersen, G. R., & Gustafsson, J.-E. (2011). Programming skill, knowledge, and working memory among professional software developers from an investment theory perspective. *Journal of Individual Differences*, 32(4), 201–209. <https://doi.org/10.1027/1614-0001/a000052>
- Bermúdez, J. L. (2007). *Thinking without words*. Oxford University Press.
- Bers, M. U. (2018). Coding, playgrounds and literacy in early childhood education: The development of KIBO robotics and ScratchJr. *2018 IEEE Global Engineering Education Conference (EDUCON)*, 2094–2102. <https://doi.org/10.1109/EDUCON.2018.8363498>
- Bers, M. U. (2019). Coding as another language: A pedagogical approach for teaching computer science in early childhood. *Journal of Computers in Education*, 6(4), 499–528. <https://doi.org/10.1007/s40692-019-00147-3>
- Bers, M. U., & Resnick, M. (2015). *The Official ScratchJr Book: Help Your Kids Learn to Code* (1 edition). No Starch Press.
- Berwick, R. C., & Chomsky, N. (2016). *Why only us: Language and evolution*. MIT press.
- Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in Cognitive Sciences*, 17(2), 89–98. <https://doi.org/10.1016/j.tics.2012.12.002>
- Bi, Y., Wei, T., Wu, C., Han, Z., Jiang, T., & Caramazza, A. (2011). The role of the left anterior temporal lobe in language processing revisited: Evidence from an individual with ATL resection. *Cortex*, 47(5), 575–587. <https://doi.org/10.1016/j.cortex.2009.12.002>
- Bickerton, D. (1990). *Language & species* (pp. x, 297). University of Chicago Press.
- Bickerton, D. (1995). *Language and Human Behavior*. University of Washington Press.
- Binder, J. R., & Desai, R. H. (2011). The Neurobiology of Semantic Memory. *Trends in Cognitive Sciences*, 15(11), 527–536. <https://doi.org/10.1016/j.tics.2011.10.001>
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex (New York, N.Y.: 1991)*, 19(12), 2767–2796. <https://doi.org/10.1093/cercor/bhp055>
- Binney, R. J., Henry, M. L., Babiak, M., Pressman, P. S., Santos-Santos, M. A., Narvid, J., Mandelli, M. L., Strain, P. J., Miller, B. L., Rankin, K. P., Rosen, H. J., & Gorno-Tempini, M. L. (2016). Reading words and other people: A comparison of exception word, familiar face and affect processing in the left and right temporal variants of primary progressive aphasia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 82, 147–163. <https://doi.org/10.1016/j.cortex.2016.05.014>
- Blank, I. A., Balewski, Z., Mahowald, K., & Fedorenko, E. (2016). Syntactic processing is distributed across the language system. *NeuroImage*, 127, 307–323. <https://doi.org/10.1016/j.neuroimage.2015.11.069>
- Blank, I. A., Duff, M. C., Brown-Schmidt, S., & Fedorenko, E. (2016). *Expanding the language network: Domain-specific hippocampal recruitment during high-level linguistic processing* (p. 091900). bioRxiv. <https://doi.org/10.1101/091900>
- Blank, I. A., & Fedorenko, E. (2017). Domain-General Brain Regions Do Not Track Linguistic Input as Closely as Language-Selective Regions. *Journal of Neuroscience*, 37(41), 9999–10011. <https://doi.org/10.1523/JNEUROSCI.3642-16.2017>
- Blank, I. A., & Fedorenko, E. (2020). No evidence for differences among language regions in their temporal receptive windows. *NeuroImage*, 116925. <https://doi.org/10.1016/j.neuroimage.2020.116925>

- Blank, I. A., Kanwisher, N., & Fedorenko, E. (2014a). A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *Journal of Neurophysiology*, 112(5), 1105–1118. <https://doi.org/10.1152/jn.00884.2013>
- Blank, I. A., Kanwisher, N., & Fedorenko, E. (2014b). A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations. *Journal of Neurophysiology*, 112(5), 1105–1118. <https://doi.org/10.1152/jn.00884.2013>
- Blank, I. A., Kiran, S., & Fedorenko, E. (2017). Can neuroimaging help aphasia researchers? Addressing generalizability, variability, and interpretability. *Cognitive Neuropsychology*, 34(6), 377–393. <https://doi.org/10.1080/02643294.2017.1402756>
- Blank, I. A., Rohter, S. V., Kiran, S., & Fedorenko, E. (2015). Functional reorganization of the large-scale brain networks that support high-level cognition following brain damage in aphasia. *Front. Psychol*, 6(14), 10–3389.
- Bohsali, A., & Crosson, B. (2016). The Basal Ganglia and Language: A Tale of Two Loops. In J.-J. Soghomonian (Ed.), *The Basal Ganglia: Novel Perspectives on Motor and Cognitive Functions* (pp. 217–242). Springer International Publishing. [https://doi.org/10.1007/978-3-319-42743-0\\_10](https://doi.org/10.1007/978-3-319-42743-0_10)
- Bonner, M. F., Peelle, J. E., Cook, P. A., & Grossman, M. (2013). Heteromodal conceptual processing in the angular gyrus. *NeuroImage*, 71, 175–186. <https://doi.org/10.1016/j.neuroimage.2013.01.006>
- Bookheimer, S. (2002). Functional MRI of Language: New Approaches to Understanding the Cortical Organization of Semantic Processing. *Annual Review of Neuroscience*, 25(1), 151–188. <https://doi.org/10.1146/annurev.neuro.25.112701.142946>
- Bottini, R., Ferraro, S., Nigri, A., Cuccarini, V., Bruzzone, M. G., & Collignon, O. (2020). Brain Regions Involved in Conceptual Retrieval in Sighted and Blind People. *Journal of Cognitive Neuroscience*, 32(6), 1009–1025. [https://doi.org/10.1162/jocn\\_a\\_01538](https://doi.org/10.1162/jocn_a_01538)
- Boylan, C., Trueswell, J. C., & Thompson-Schill, S. L. (2015). Compositionality and the angular gyrus: A multi-voxel similarity analysis of the semantic composition of nouns and verbs. *Neuropsychologia*, 78, 130–141. <https://doi.org/10.1016/j.neuropsychologia.2015.10.007>
- Braga, R. M., & Buckner, R. L. (2017). Parallel Interdigitated Distributed Networks within the Individual Estimated by Intrinsic Functional Connectivity. *Neuron*, 95(2), 457–471.e5. <https://doi.org/10.1016/j.neuron.2017.06.038>
- Braga, R. M., DiNicola, L. M., Becker, H. C., & Buckner, R. L. (2020). Situating the Left-Lateralized Language Network in the Broader Organization of Multiple Specialized Large-Scale Distributed Networks. *Journal of Neurophysiology*. <https://doi.org/10.1152/jn.00753.2019>
- Braga, R. M., Van Dijk, K. R. A., Polimeni, J. R., Eldaief, M. C., & Buckner, R. L. (2019). Parallel distributed networks resolved at high resolution reveal close juxtaposition of distinct regions. *Journal of Neurophysiology*, 121(4), 1513–1534. <https://doi.org/10.1152/jn.00808.2018>
- Brennan, J. R., Stabler, E. P., Van Wagenen, S. E., Luh, W.-M., & Hale, J. T. (2016). Abstract linguistic structure correlates with temporal activity during naturalistic comprehension. *Brain and Language*, 157–158, 81–94. <https://doi.org/10.1016/j.bandl.2016.04.008>
- Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nature Reviews Neuroscience*, 3(3), 243–249. <https://doi.org/10.1038/nrn756>
- Bright, P., Moss, H., & Tyler, L. K. (2004). Unitary vs multiple semantics: PET studies of word and picture processing. *Brain and Language*, 89(3), 417–432. <https://doi.org/10.1016/j.bandl.2004.01.010>
- Broca, P. (1861). Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche du cerveau. *Bulletin de La Société Anthropologique*, 2(1), 235–238.
- Broca, P. (1865). Sur le siège de la faculté du langage articulé (15 juin). *Bulletins de La Société Anthropologique de Paris*, 6, 377–393.

- Brojde, C. L., Porter, C., & Colunga, E. (2011). Words can slow down category learning. *Psychonomic Bulletin & Review*, 18(4), 798–804. <https://doi.org/10.3758/s13423-011-0103-z>
- Brown, P., & Levinson, S. C. (1992). 'Left' and 'Right' in Tenejapa: Investigating a Linguistic and Conceptual Gap. *STUF - Language Typology and Universals*, 45(1–4), 590–611. <https://doi.org/10.1524/stuf.1992.45.14.590>
- Brown, P., & Levinson, S. C. (1993). "Uphill" and "Downhill" in Tzeltal. *Journal of Linguistic Anthropology*, 3(1), 46–74.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language Models are Few-Shot Learners. *ArXiv:2005.14165 [Cs]*. <http://arxiv.org/abs/2005.14165>
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The Brain's Default Network. *Annals of the New York Academy of Sciences*, 1124(1), 1–38. <https://doi.org/10.1196/annals.1440.011>
- Buckner, R. L., & DiNicola, L. M. (2019). The brain's default network: Updated anatomy, physiology and evolving insights. *Nature Reviews Neuroscience*, 20(10), 593–608. <https://doi.org/10.1038/s41583-019-0212-7>
- Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C., & Yeo, B. T. T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(5), 2322–2345. <https://doi.org/10.1152/jn.00339.2011>
- Bull, R., Marschark, M., Sapere, P., Davidson, W. A., Murphy, D., & Nordmann, E. (2011). Numerical Estimation in Deaf and Hearing Adults. *Learning and Individual Differences*, 21(4), 453–457. <https://doi.org/10.1016/j.lindif.2011.02.001>
- Burger, R. A., & Muma, J. R. (1980). Cognitive distancing in mediated categorization in aphasia. *Journal of Psycholinguistic Research*, 9(4), 355–365. <https://doi.org/10.1007/BF01067448>
- Buse, R. P. L., & Weimer, W. R. (2010). Learning a Metric for Code Readability. *IEEE Transactions on Software Engineering*, 36(4), 546–558. <https://doi.org/10.1109/TSE.2009.70>
- Cappelletti, M., Butterworth, B., & Kopelman, M. (2001). Spared numerical abilities in a case of semantic dementia. *Neuropsychologia*, 39(11), 1224–1239. [https://doi.org/10.1016/s0028-3932\(01\)00035-5](https://doi.org/10.1016/s0028-3932(01)00035-5)
- Caramazza, A., Berndt, R. S., & Brownell, H. H. (1982). The semantic deficit hypothesis: Perceptual parsing and object classification by aphasic patients. *Brain and Language*, 15(1), 161–189. [https://doi.org/10.1016/0093-934x\(82\)90054-2](https://doi.org/10.1016/0093-934x(82)90054-2)
- Carlson, G. N., & Tanenhaus, M. K. (1988). Thematic Roles and Language Comprehension. In *Thematic Relations* (pp. 263–288). Brill. [https://doi.org/10.1163/9789004373211\\_015](https://doi.org/10.1163/9789004373211_015)
- Carruthers, P. (2002). The cognitive functions of language. *The Behavioral and Brain Sciences*, 25(6), 657–674; discussion 674–725. <https://doi.org/10.1017/s0140525x02000122>
- Caspers, S., Geyer, S., Schleicher, A., Mohlberg, H., Amunts, K., & Zilles, K. (2006). The human inferior parietal cortex: Cytoarchitectonic parcellation and interindividual variability. *NeuroImage*, 33(2), 430–448. <https://doi.org/10.1016/j.neuroimage.2006.06.054>
- Castelhano, J., Duarte, I. C., Ferreira, C., Duraes, J., Madeira, H., & Castelo-Branco, M. (2019). The role of the insula in intuitive expert bug detection in computer code: An fMRI study. *Brain Imaging and Behavior*, 13(3), 623–637. <https://doi.org/10.1007/s11682-018-9885-1>
- Caucheteux, C., Gramfort, A., & King, J.-R. (2021, November 7). *Model-based analysis of brain activity reveals the hierarchy of language in 305 subjects*. EMNLP 2021 - Conference on Empirical Methods in Natural Language Processing. <https://hal.archives-ouvertes.fr/hal-03361430>
- Caucheteux, C., & King, J.-R. (2022). Brains and algorithms partially converge in natural language processing. *Communications Biology*, 5(1), 1–10. <https://doi.org/10.1038/s42003-022-03036-1>

- Cetron, J. S., Connolly, A. C., Diamond, S. G., May, V. V., Haxby, J. V., & Kraemer, D. J. M. (2019). Decoding individual differences in STEM learning from functional MRI data. *Nature Communications*, 10(1), 1–10. <https://doi.org/10.1038/s41467-019-10053-y>
- Chang, C. H. C., Nastase, S. A., & Hasson, U. (2022). *Information flow across the cortical timescales hierarchy during narrative construction* (p. 2021.12.01.470825). bioRxiv. <https://doi.org/10.1101/2021.12.01.470825>
- Chen, G., Taylor, P. A., & Cox, R. W. (2017). Is the statistic value all we should care about in neuroimaging? *NeuroImage*, 147, 952–959. <https://doi.org/10.1016/j.neuroimage.2016.09.066>
- Chen, S., Conway, B., & Gibson, E. (in prep). *Investigating the effects of color categories on color cognition*.
- Chen, X., Affourtit, J., Ryskin, R., Regev, T. I., Norman-Haignere, S., Jouravlev, O., Malik-Moraleda, S., Kean, H., Varley, R., & Fedorenko, E. (2021). *The human language system does not support music processing* (p. 2021.06.01.446439). bioRxiv. <https://doi.org/10.1101/2021.06.01.446439>
- Chertkow, H., Bub, D., Deaudon, C., & Whitehead, V. (1997). On the status of object concepts in aphasia. *Brain and Language*, 58(2), 203–232. <https://doi.org/10.1006/brln.1997.1771>
- Cheung, C., Ivanova, A. A., Siegelman, M., Pongos, A. L. A., Kean, H. H., & Fedorenko, E. (2020). The effect of task on sentence processing in the brain. *Poster Presentation at the Society for the Neurobiology of Language*.
- Chklovskii, D. B., & Koulakov, A. A. (2004). Maps in the brain: What can we learn from them? *Annual Review of Neuroscience*, 27, 369–392. <https://doi.org/10.1146/annurev.neuro.27.070203.144226>
- Choi, H.-J., Zilles, K., Mohlberg, H., Schleicher, A., Fink, G. R., Armstrong, E., & Amunts, K. (2006). Cytoarchitectonic identification and probabilistic mapping of two distinct areas within the anterior ventral bank of the human intraparietal sulcus. *Journal of Comparative Neurology*, 495(1), 53–69. <https://doi.org/10.1002/cne.20849>
- Chomsky, N. (2007). Biolinguistic Explorations: Design, Development, Evolution. *International Journal of Philosophical Studies*, 15(1), 1–21. <https://doi.org/10.1080/09672550601143078>
- Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., Barham, P., Chung, H. W., Sutton, C., Gehrmann, S., Schuh, P., Shi, K., Tsvyashchenko, S., Maynez, J., Rao, A., Barnes, P., Tay, Y., Shazeer, N., Prabhakaran, V., ... Fiedel, N. (2022). PaLM: Scaling Language Modeling with Pathways. *ArXiv:2204.02311 [Cs]*. <http://arxiv.org/abs/2204.02311>
- Clark, E. V. (2004). How language acquisition builds on cognitive development. *Trends in Cognitive Sciences*, 8(10), 472–478. <https://doi.org/10.1016/j.tics.2004.08.012>
- Coccia, M., Bartolini, M., Luzzi, S., Provinciali, L., & Lambon Ralph, M. A. (2004). Semantic memory is an amodal, dynamic system: Evidence from the interaction of naming and object use in semantic dementia. *Cognitive Neuropsychology*, 21(5), 513–527. <https://doi.org/10.1080/02643290342000113>
- Coco, M. I., Nuthmann, A., & Dimigen, O. (2020). Fixation-related Brain Potentials during Semantic Integration of Object-Scene Information. *Journal of Cognitive Neuroscience*, 32(4), 571–589. [https://doi.org/10.1162/jocn\\_a\\_01504](https://doi.org/10.1162/jocn_a_01504)
- Coetzee, J. P., & Monti, M. M. (2018). At the core of reasoning: Dissociating deductive and non-deductive load. *Human Brain Mapping*, 39(4), 1850–1861. <https://doi.org/10.1002/hbm.23979>
- Cohen, L., Dehaene, S., Chochon, F., Lehéricy, S., & Naccache, L. (2000). Language and calculation within the parietal lobe: A combined cognitive, anatomical and fMRI study. *Neuropsychologia*, 38(10), 1426–1440. [https://doi.org/10.1016/s0028-3932\(00\)00038-5](https://doi.org/10.1016/s0028-3932(00)00038-5)
- Cohn, N. (2020). Your Brain on Comics: A Cognitive Model of Visual Narrative Comprehension. *Topics in Cognitive Science*, 12(1), 352–386. <https://doi.org/10.1111/tops.12421>

- Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of Agents: The processing of semantic roles in visual narrative. *Cognitive Psychology*, 67(3), 73–97. <https://doi.org/10.1016/j.cogpsych.2013.07.002>
- Cole, M. W., & Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *NeuroImage*, 37(1), 343–360. <https://doi.org/10.1016/j.neuroimage.2007.03.071>
- Colvin, M., Warren, T., & Dickey, M. W. (2019). Event Knowledge and Verb Knowledge Predict Sensitivity to Different Aspects of Semantic Anomalies in Aphasia. In K. Carlson, Jr. Clifton Charles, & J. D. Fodor (Eds.), *Grammatical Approaches to Language Processing: Essays in Honor of Lyn Frazier* (pp. 241–259). Springer International Publishing. [https://doi.org/10.1007/978-3-030-01563-3\\_13](https://doi.org/10.1007/978-3-030-01563-3_13)
- Connell, L., & Lynott, D. (2013). Flexible and fast: Linguistic shortcut affects both shallow and deep conceptual processing. *Psychonomic Bulletin & Review*, 20(3), 542–550. <https://doi.org/10.3758/s13423-012-0368-x>
- Corballis, P. M. (2003). Visuospatial processing and the right-hemisphere interpreter. *Brain and Cognition*, 53(2), 171–176. [https://doi.org/10.1016/s0278-2626\(03\)00103-9](https://doi.org/10.1016/s0278-2626(03)00103-9)
- Crawford, J. R., & Howell, D. C. (1998). Comparing an Individual's Test Score Against Norms Derived from Small Samples. *The Clinical Neuropsychologist*, 12(4), 482–486. <https://doi.org/10.1076/clin.12.4.482.7241>
- Crittenden, B. M., Mitchell, D. J., & Duncan, J. (2015). Recruitment of the default mode network during a demanding act of executive control. *eLife*, 4, e06481. <https://doi.org/10.7554/eLife.06481>
- Crosson, B., Benjamin, M., & Levy, I. (2007). Role of the basal ganglia in language and semantics: Supporting cast. In J. Hart & M. A. Kraut (Eds.), *Neural Basis of Semantic Memory* (pp. 219–244). Cambridge University Press. <https://doi.org/10.1017/CBO9780511544965.010>
- Dahan, D., Magnuson, J. S., & Tanerhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, 42(4), 317–367. <https://doi.org/10.1006/cogp.2001.0750>
- Dalbey, J., & Linn, M. (1985). The Demands and Requirements of Computer Programming: A Literature Review. *Journal of Educational Computing Research*, 1. <https://doi.org/10.2190/BC76-8479-YM0X-7FUA>
- Damasio, A. R. (1983). Language and the basal ganglia. *Trends in Neurosciences*, 6, 442–443. [https://doi.org/10.1016/0166-2236\(83\)90213-8](https://doi.org/10.1016/0166-2236(83)90213-8)
- Darwin, C. (1871). *The descent of man and selection in relation to sex*. John Murray.
- Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., Krieger-Redwood, K., Bernhardt, B. C., Smallwood, J., & Jefferies, E. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *NeuroImage*, 137, 165–177. <https://doi.org/10.1016/j.neuroimage.2016.05.051>
- Davidson, D. (1975). Thought and talk. *Mind and Language*, 1975, 7–23.
- Dax, M. (1865). Lésions de la moitié gauche de l'encéphale coïncident avec l'oubli des signes de la pensée (lu à Montpellier en 1836). *Bulletin Hebdomadaire de Médecine et de Chirurgie*, 2(2), 259–262.
- Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain*. WW Norton & Company.
- Deen, B., & Freiwald, W. A. (2022). *Parallel systems for social and spatial reasoning within the cortical apex* (p. 2021.09.23.461550). bioRxiv. <https://doi.org/10.1101/2021.09.23.461550>
- Deen, B., Koldewyn, K., Kanwisher, N., & Saxe, R. (2015). Functional Organization of Social Perception and Cognition in the Superior Temporal Sulcus. *Cerebral Cortex*, 25(11), 4596–4609. <https://doi.org/10.1093/cercor/bhv111>

- Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: A tool for unwelding visuo-spatial memory. *Neuropsychologia*, 37(10), 1189–1199. [https://doi.org/10.1016/s0028-3932\(98\)00159-6](https://doi.org/10.1016/s0028-3932(98)00159-6)
- Deniz, F., Nunez-Elizalde, A. O., Huth, A. G., & Gallant, J. L. (2019). The Representation of Semantic Information Across Human Cerebral Cortex During Listening Versus Reading Is Invariant to Stimulus Modality. *Journal of Neuroscience*, 39(39), 7722–7736. <https://doi.org/10.1523/JNEUROSCI.0675-19.2019>
- Dennett, D. C. (1994). The role of language in intelligence. In *What is Intelligence?* The Darwin College Lectures, ed. Jean Kalfa, Cambridge University Press.
- Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 33(48), 18906–18916. <https://doi.org/10.1523/JNEUROSCI.3809-13.2013>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *ArXiv:1810.04805 [Cs]*. <http://arxiv.org/abs/1810.04805>
- Diachek, E., Blank, I., Siegelman, M., Affourtit, J., & Fedorenko, E. (2020). The Domain-General Multiple Demand (MD) Network Does Not Support Core Aspects of Language Comprehension: A Large-Scale fMRI Investigation. *Journal of Neuroscience*, 40(23), 4536–4550. <https://doi.org/10.1523/JNEUROSCI.2036-19.2020>
- Dickey, M. W., & Warren, T. (2015). The influence of event-related knowledge on verb-argument processing in aphasia. *Neuropsychologia*, 0, 63–81. <https://doi.org/10.1016/j.neuropsychologia.2014.12.003>
- DiNicola, L. M., Braga, R. M., & Buckner, R. L. (2020). Parallel distributed networks dissociate episodic and social functions within the individual. *Journal of Neurophysiology*, 123(3), 1144–1179. <https://doi.org/10.1152/jn.00529.2019>
- Dresang, H. C., Dickey, M. W., & Warren, T. C. (2019). Semantic memory for objects, actions, and events: A novel test of event-related conceptual semantic knowledge. *Cognitive Neuropsychology*, 36(7–8), 313–335. <https://doi.org/10.1080/02643294.2019.1656604>
- Dronkers, N. F., Ludy, C. A., & Redfern, B. B. (1998). Pragmatics in the absence of verbal language: Descriptions of a severe aphasic and a language-deprived adult. *Journal of Neurolinguistics*, 11(1), 179–190. [https://doi.org/10.1016/S0911-6044\(98\)00012-8](https://doi.org/10.1016/S0911-6044(98)00012-8)
- Duff, M. C., & Brown-Schmidt, S. (2012). The hippocampus and the flexible use and processing of language. *Frontiers in Human Neuroscience*, 6. <https://www.frontiersin.org/articles/10.3389/fnhum.2012.00069>
- Duff, M. C., & Brown-Schmidt, S. (2017). Hippocampal Contributions to Language Use and Processing. In D. E. Hannula & M. C. Duff (Eds.), *The Hippocampus from Cells to Systems: Structure, Connectivity, and Functional Contributions to Memory and Flexible Cognition* (pp. 503–536). Springer International Publishing. [https://doi.org/10.1007/978-3-319-50406-3\\_16](https://doi.org/10.1007/978-3-319-50406-3_16)
- Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, 14(4), 172–179. <https://doi.org/10.1016/j.tics.2010.01.004>
- Duncan, J. (2013). The Structure of Cognition: Attentional Episodes in Mind and Brain. *Neuron*, 80(1), 35–50. <https://doi.org/10.1016/j.neuron.2013.09.015>
- Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, 23(10), 475–483. [https://doi.org/10.1016/S0166-2236\(00\)01633-7](https://doi.org/10.1016/S0166-2236(00)01633-7)
- Elazar, Y., Kassner, N., Ravfogel, S., Ravichander, A., Hovy, E., Schütze, H., & Goldberg, Y. (2021). Measuring and Improving Consistency in Pretrained Language Models. *ArXiv:2102.01017 [Cs]*. <http://arxiv.org/abs/2102.01017>

- Elazar, Y., Zhang, H., Goldberg, Y., & Roth, D. (2021). Back to Square One: Artifact Detection, Training and Commonsense Disentanglement in the Winograd Schema. *ArXiv:2104.08161 [Cs]*. <http://arxiv.org/abs/2104.08161>
- Elli, G. V., Lane, C., & Bedny, M. (2019). A Double Dissociation in Sensitivity to Verb and Noun Semantics Across Cortical Networks. *Cerebral Cortex*, 29(11), 4803–4817. <https://doi.org/10.1093/cercor/bhz014>
- Ensmenger, N. (2012). The Digital Construction of Technology: Rethinking the History of Computers in Society. *Technology and Culture*, 53(4), 753–776. <https://doi.org/10.1353/tech.2012.0126>
- Estes, Z., Golonka, S., & Jones, L. L. (2011). Thematic thinking: The apprehension and consequences of thematic relations. In *The psychology of learning and motivation: Advances in research and theory*, Vol. 54 (pp. 249–294). Elsevier Academic Press.
- Ettinger, A. (2020). What BERT is not: Lessons from a new suite of psycholinguistic diagnostics for language models. *ArXiv:1907.13528 [Cs]*. <http://arxiv.org/abs/1907.13528>
- Evans, M., Krieger-Redwood, K., Gonzalez Alam, T. R., Smallwood, J., & Jefferies, E. (2020). Controlled semantic summation correlates with intrinsic connectivity between default mode and control networks. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 129, 356–375. <https://doi.org/10.1016/j.cortex.2020.04.032>
- Fairhall, S. L., & Caramazza, A. (2013). Brain Regions That Represent Amodal Conceptual Knowledge. *Journal of Neuroscience*, 33(25), 10552–10558. <https://doi.org/10.1523/JNEUROSCI.0051-13.2013>
- Fakhoury, S., Ma, Y., Arnaoudova, V., & Adesope, O. (2018). The Effect of Poor Source Code Lexicon and Readability on Developers' Cognitive Load. *Proceedings of the 26th Conference on Program Comprehension*, 286–296. <https://doi.org/10.1145/3196321.3196347>
- Fang, Y., Chen, Q., Lingnau, A., Han, Z., & Bi, Y. (2016). Areas Recruited during Action Understanding Are Not Modulated by Auditory or Sign Language Experience. *Frontiers in Human Neuroscience*, 10, 94. <https://doi.org/10.3389/fnhum.2016.00094>
- Fay, T. (1971). Perceived Hierarchic Structure in Language and Music. *Journal of Music Theory*, 15(1/2), 112–137. <https://doi.org/10.2307/842898>
- Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Frontiers in Psychology*, 5. <https://www.frontiersin.org/article/10.3389/fpsyg.2014.00335>
- Fedorenko, E. (2021). The early origins and the growing popularity of the individual-subject analytic approach in human neuroscience. *Current Opinion in Behavioral Sciences*, 40, 105–112. <https://doi.org/10.1016/j.cobeha.2021.02.023>
- Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, 108(39), 16428–16433. <https://doi.org/10.1073/pnas.1112937108>
- Fedorenko, E., Bers, M. U., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, 108(39), 16428–16433. <https://doi.org/10.1073/pnas.1112937108>
- Fedorenko, E., & Blank, I. A. (2020). Broca's Area Is Not a Natural Kind. *Trends in Cognitive Sciences*. <https://doi.org/10.1016/j.tics.2020.01.001>
- Fedorenko, E., Blank, I. A., Siegelman, M., & Mineroff, Z. (2020a). Lack of selectivity for syntax relative to word meanings throughout the language network. *Cognition*, 203, 104348. <https://doi.org/10.1016/j.cognition.2020.104348>
- Fedorenko, E., Blank, I., Siegelman, M., & Mineroff, Z. (2020b). Lack of selectivity for syntax relative to word meanings throughout the language network. *BioRxiv*, 477851. <https://doi.org/10.1101/477851>
- Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-Selective and Domain-General Regions Lie Side by Side within Broca's Area. *Current Biology*, 22(21), 2059–2062. <https://doi.org/10.1016/j.cub.2012.09.011>

- Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 110(41), 16616–16621. <https://doi.org/10.1073/pnas.1315235110>
- Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: Defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, 104(2), 1177–1194. <https://doi.org/10.1152/jn.00032.2010>
- Fedorenko, E., Ivanova, A., Dhamala, R., & Bers, M. U. (2019). The Language of Programming: A Cognitive Perspective. *Trends in Cognitive Sciences*, 23(7), 525–528. <https://doi.org/10.1016/j.tics.2019.04.010>
- Fedorenko, E., & Kanwisher, N. (2009). Neuroimaging of Language: Why Hasn't a Clearer Picture Emerged? *Language and Linguistics Compass*, 3(4), 839–865. <https://doi.org/10.1111/j.1749-818X.2009.00143.x>
- Fedorenko, E., McDermott, J. H., Norman-Haignere, S., & Kanwisher, N. (2012). Sensitivity to musical structure in the human brain. *Journal of Neurophysiology*, 108(12), 3289–3300. <https://doi.org/10.1152/jn.00209.2012>
- Fedorenko, E., Nieto-Castañón, A., & Kanwisher, N. (2012). Lexical and syntactic representations in the brain: An fMRI investigation with multi-voxel pattern analyses. *Neuropsychologia*, 50(4), 499–513. <https://doi.org/10.1016/j.neuropsychologia.2011.09.014>
- Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in Cognitive Sciences*, 18(3), 120–126. <https://doi.org/10.1016/j.tics.2013.12.006>
- Fedorenko, E., & Varley, R. A. (2016). Language and thought are not the same thing: Evidence from neuroimaging and neurological patients: Language versus thought. *Annals of the New York Academy of Sciences*, 1369(1), 132–153. <https://doi.org/10.1111/nyas.13046>
- Feng, W., Wu, Y., Jan, C., Yu, H., Jiang, X., & Zhou, X. (2017). Effects of contextual relevance on pragmatic inference during conversation: An fMRI study. *Brain and Language*, 171, 52–61. <https://doi.org/10.1016/j.bandl.2017.04.005>
- Feng, W., Yu, H., & Zhou, X. (2021). Understanding particularized and generalized conversational implicatures: Is theory-of-mind necessary? *Brain and Language*, 212, 104878. <https://doi.org/10.1016/j.bandl.2020.104878>
- Ferguson, B., & Waxman, S. (2017). Linking language and categorization in infancy. *Journal of Child Language*, 44(3), 527–552. <https://doi.org/10.1017/S0305000916000568>
- Ferstl, E. C., Neumann, J., Bogler, C., & von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping*, 29(5), 581–593. <https://doi.org/10.1002/hbm.20422>
- Ferstl, E. C., & von Cramon, D. Y. (2002). What Does the Frontomedian Cortex Contribute to Language Processing: Coherence or Theory of Mind? *NeuroImage*, 17(3), 1599–1612. <https://doi.org/10.1006/nimg.2002.1247>
- Fillmore, C. J. (1968). Lexical entries for verbs. *Foundations of Language*, 373–393.
- Fillmore, C. J. (2002). *Form and Meaning in Language: Volume I, Papers on Semantic Roles* (74 edition). Center for the Study of Language and Information.
- Fillmore, C. J. (2006). Frame semantics. In *Cognitive linguistics: Basic readings* (Vol. 34, pp. 373–400). Mouton de Gruyter New York.
- Fischer, J., Mikhael, J. G., Tenenbaum, J. B., & Kanwisher, N. (2016). Functional neuroanatomy of intuitive physical inference. *Proceedings of the National Academy of Sciences*, 113(34), E5072–E5081. <https://doi.org/10.1073/pnas.1610344113>
- Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, 97(2), 179–210. <https://doi.org/10.1016/j.cognition.2005.02.005>
- Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, 1316, 87–104. <https://doi.org/10.1111/nyas.12406>

- Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition*, 57(2), 109–128. [https://doi.org/10.1016/0010-0277\(95\)00692-R](https://doi.org/10.1016/0010-0277(95)00692-R)
- Floyd, B., Santander, T., & Weimer, W. (2017). Decoding the Representation of Code in the Brain: An fMRI Study of Code Review and Expertise. 2017 IEEE/ACM 39th International Conference on Software Engineering (ICSE), 175–186. <https://doi.org/10.1109/ICSE.2017.24>
- Fodor, J. A. (1975). *The language of thought*. Harvard University Press.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35(1), 116–124. <https://doi.org/10.3758/BF03195503>
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27), 9673–9678. <https://doi.org/10.1073/pnas.0504136102>
- Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, 108(3), 819–824. <https://doi.org/10.1016/j.cognition.2008.04.007>
- Frankland, S. M., & Greene, J. D. (2015). An architecture for encoding sentence meaning in left mid-superior temporal cortex. *Proceedings of the National Academy of Sciences*, 112(37), 11732–11737. <https://doi.org/10.1073/pnas.1421236112>
- Frankland, S. M., & Greene, J. D. (2020). Two Ways to Build a Thought: Distinct Forms of Compositional Semantic Representation across Brain Regions. *Cerebral Cortex*. <https://doi.org/10.1093/cercor/bhaa001>
- Franklin, A., Drivonikou, G. V., Clifford, A., Kay, P., Regier, T., & Davies, I. R. L. (2008). Lateralization of categorical perception of color changes with color term acquisition. *Proceedings of the National Academy of Sciences*, 105(47), 18221–18225. <https://doi.org/10.1073/pnas.0809952105>
- Franklin, S., Turner, J., & Ellis, A. (1992). The ADA auditory comprehension battery. York, UK: University of York.
- Frässle, S., Sommer, J., Jansen, A., Naber, M., & Einhäuser, W. (2014). Binocular Rivalry: Frontal Activity Relates to Introspection and Action But Not to Perception. *Journal of Neuroscience*, 34(5), 1738–1747. <https://doi.org/10.1523/JNEUROSCI.4403-13.2014>
- Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science (New York, N.Y.)*, 291(5502), 312–316. <https://doi.org/10.1126/science.291.5502.312>
- Friederici, A. D. (2020). Hierarchy processing in human neurobiology: How specific is it? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1789), 20180391. <https://doi.org/10.1098/rstb.2018.0391>
- Friederici, A. D., Bahlmann, J., Friedrich, R., & Makuuchi, M. (2011). The Neural Basis of Recursion and Complex Syntactic Hierarchy. *BIOLINGUISTICS*, 5(1–2), 087–104.
- Friston, Karl. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping*, 3(3), 165–189. <https://doi.org/10.1002/hbm.460030303>
- Frost, M. A., & Goebel, R. (2012). Measuring structural-functional correspondence: Spatial variability of specialised brain regions after macro-anatomical alignment. *NeuroImage*, 59(2), 1369–1381. <https://doi.org/10.1016/j.neuroimage.2011.08.035>
- Fuster, J. M. (1991). Up and down the frontal hierarchies; whither Broca's area? *Behavioral and Brain Sciences*, 14(4), 558–558. <https://doi.org/10.1017/S0140525X00071302>
- Gainotti, G. (2011). The organization and dissolution of semantic-conceptual knowledge: Is the "amodal hub" the only plausible model? *Brain and Cognition*, 75(3), 299–309. <https://doi.org/10.1016/j.bandc.2010.12.001>

- Gainotti, G. (2015). Is the difference between right and left ATLs due to the distinction between general and social cognition or between verbal and non-verbal representations? *Neuroscience & Biobehavioral Reviews*, 51, 296–312. <https://doi.org/10.1016/j.neubiorev.2015.02.004>
- Gainotti, G., D'Erme, P., Villa, G., & Caltagirone, C. (1986). Focal brain lesions and intelligence: A study with a new version of Raven's Colored Matrices. *Journal of Clinical and Experimental Neuropsychology*, 8(1), 37–50. <https://doi.org/10.1080/01688638608401295>
- Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: An fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia*, 38(1), 11–21. [https://doi.org/10.1016/s0028-3932\(99\)00053-6](https://doi.org/10.1016/s0028-3932(99)00053-6)
- Gauker, C. (2011). *Words and images: An essay on the origin of ideas*. OUP Oxford.
- Gershkoff-Stowe, L., Thal, D. J., Smith, L. B., & Namy, L. L. (1997). Categorization and Its Developmental Relation to Early Language. *Child Development*, 68(5), 843–859. <https://doi.org/10.2307/1132037>
- Geschwind, N. (1970). The Organization of Language and the Brain: Language disorders after brain damage help in elucidating the neural basis of verbal behavior. *Science*, 170(3961), 940–944. <https://doi.org/10.1126/science.170.3961.940>
- Giglio, L., Ostarek, M., Weber, K., & Hagoort, P. (2021). Commonalities and Asymmetries in the Neurobiological Infrastructure for Language Production and Comprehension. *Cerebral Cortex, bhab287*. <https://doi.org/10.1093/cercor/bhab287>
- Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences of the United States of America*, 103(2), 489–494. <https://doi.org/10.1073/pnas.0509868103>
- Gläscher, J., Rudrauf, D., Colom, R., Paul, L. K., Tranel, D., Damasio, H., & Adolphs, R. (2010). Distributed neural system for general intelligence revealed by lesion mapping. *Proceedings of the National Academy of Sciences*, 107(10), 4705–4709. <https://doi.org/10.1073/pnas.0910397107>
- Giliozzi, V., Mayor, J., Hu, J.-F., & Plunkett, K. (2009). Labels as Features (Not Names) for Infant Categorization: A Neurocomputational Approach. *Cognitive Science*, 33(4), 709–738. <https://doi.org/10.1111/j.1551-6709.2009.01026.x>
- Goel, V. (2007). Anatomy of deductive reasoning. *Trends in Cognitive Sciences*, 11(10), 435–441. <https://doi.org/10.1016/j.tics.2007.09.003>
- Goel, V., & Dolan, R. J. (2001). Functional neuroanatomy of three-term relational reasoning. *Neuropsychologia*, 39(9), 901–909. [https://doi.org/10.1016/S0028-3932\(01\)00024-0](https://doi.org/10.1016/S0028-3932(01)00024-0)
- Goldstein, I., & Papert, S. (1977). Artificial intelligence, language, and the study of knowledge. *Cognitive Science*, 1(1), 84–123. [https://doi.org/10.1016/S0364-0213\(77\)80006-2](https://doi.org/10.1016/S0364-0213(77)80006-2)
- Gorno-Tempini, M. L., Rankin, K. P., Woolley, J. D., Rosen, H. J., Phengrasamy, L., & Miller, B. L. (2004). Cognitive and Behavioral Profile in a Case of Right Anterior Temporal Lobe Neurodegeneration. *Cortex*, 40(4), 631–644. [https://doi.org/10.1016/S0010-9452\(08\)70159-X](https://doi.org/10.1016/S0010-9452(08)70159-X)
- Grand, G., Blank, I. A., Pereira, F., & Fedorenko, E. (2022). Semantic projection recovers rich human knowledge of multiple object features from word embeddings. *Nature Human Behaviour*, 1–13. <https://doi.org/10.1038/s41562-022-01316-8>
- Greene, M. R., & Fei-Fei, L. (2014). Visual categorization is automatic and obligatory: Evidence from Stroop-like paradigm. *Journal of Vision*, 14(1). <https://doi.org/10.1167/14.1.14>
- Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences*, 100(1), 253–258. <https://doi.org/10.1073/pnas.0135058100>
- Groen, I. I. A., Silson, E. H., Pitcher, D., & Baker, C. I. (2021). Theta-burst TMS of lateral occipital cortex reduces BOLD responses across category-selective areas in ventral temporal cortex. *NeuroImage*, 230, 117790. <https://doi.org/10.1016/j.neuroimage.2021.117790>

- Gruber, J. S. (1965). *Studies in lexical relations*. [PhD Thesis]. Massachusetts Institute of Technology.
- Guell, X., Gabrieli, J. D. E., & Schmahmann, J. D. (2018). Triple representation of language, working memory, social and emotion processing in the cerebellum: Convergent evidence from task and seed-based resting-state fMRI analyses in a single large cohort. *NeuroImage*, 172, 437–449. <https://doi.org/10.1016/j.neuroimage.2018.01.082>
- Guell, X., Schmahmann, J. D., Gabrieli, J. D., & Ghosh, S. S. (2018). Functional gradients of the cerebellum. *eLife*, 7, e36652. <https://doi.org/10.7554/eLife.36652>
- Gurevich, O., Johnson, M. A., & Goldberg, A. E. (2010). *Incidental verbatim memory for language*. 2(1), 45–78. <https://doi.org/10.1515/langcog.2010.003>
- Guzdial, M. (2008). EducationPaving the way for computational thinking. *Communications of the ACM*, 51(8), 25–27. <https://doi.org/10.1145/1378704.1378713>
- Habas, C. (2021). Functional Connectivity of the Cognitive Cerebellum. *Frontiers in Systems Neuroscience*, 15. <https://www.frontiersin.org/articles/10.3389/fnsys.2021.642225>
- Häberling, I. S., Corballis, P. M., & Corballis, M. C. (2016). Language, gesture, and handedness: Evidence for independent lateralized networks. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 82, 72–85. <https://doi.org/10.1016/j.cortex.2016.06.003>
- Hafri, A., Trueswell, J. C., & Strickland, B. (2018). Encoding of event roles from visual scenes is rapid, spontaneous, and interacts with higher-level visual processing. *Cognition*, 175, 36–52. <https://doi.org/10.1016/j.cognition.2018.02.011>
- Hagoort, P., & Levinson, S. C. (2014). Neuropragmatics. In *The cognitive neurosciences*, 5th ed (pp. 667–674). MIT Press.
- Hagoort, P., & van Berkum, J. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481), 801–811. <https://doi.org/10.1098/rstb.2007.2089>
- Handjaras, G., Leo, A., Cecchetti, L., Papale, P., Lenci, A., Marotta, G., Pietrini, P., & Ricciardi, E. (2017). Modality-independent encoding of individual concepts in the left parietal cortex. *Neuropsychologia*, 105, 39–49. <https://doi.org/10.1016/j.neuropsychologia.2017.05.001>
- Hassabis, D., & Maguire, E. A. (2009). The construction system of the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1263–1271. <https://doi.org/10.1098/rstb.2008.0296>
- Hassenfeld, Z. R., & Bers, M. U. (2020). Debugging the Writing Process: Lessons From a Comparison of Students' Coding and Writing Practices. *International Literacy Association*, 1–12. <https://doi.org/10.1002/trtr.1885>
- Hassenfeld, Z. R., Govind, M., Ruiter, L. E. D., & Bers, M. U. (2020). If You Can Program, You Can Write: Learning Introductory Programming Across Literacy Levels. *Journal of Information Technology Education: Research*, 19, 065–085.
- Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: Memory as an integral component of information processing. *Trends in Cognitive Sciences*, 19(6), 304–313. <https://doi.org/10.1016/j.tics.2015.04.006>
- Hauptman, M., Blank, I., & Fedorenko, E. (2022). *Non-literal language processing is jointly supported by the language and Theory of Mind networks: Evidence from a novel meta-analytic fMRI approach* (p. 2022.03.08.481056). bioRxiv. <https://doi.org/10.1101/2022.03.08.481056>
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298(5598), 1569–1579. <https://doi.org/10.1126/science.298.5598.1569>
- Heavey, C. L., & Hurlburt, R. T. (2008). The phenomena of inner experience. *Consciousness and Cognition*, 17(3), 798–810. <https://doi.org/10.1016/j.concog.2007.12.006>
- Hermans, F., & Aldewereld, M. (2017). Programming is Writing is Programming. *Companion to the First International Conference on the Art, Science and Engineering of Programming*, 33:1–33:8. <https://doi.org/10.1145/3079368.3079413>
- Highnam, C. L., & Bleile, K. M. (2011). Language in the Cerebellum. *American Journal of Speech-Language Pathology*, 20(4), 337–347. [https://doi.org/10.1044/1058-0360\(2011/10-0096\)](https://doi.org/10.1044/1058-0360(2011/10-0096))

- Hinzen, W. (2013). Narrow syntax and the language of thought. *Philosophical Psychology*, 26(1), 1–23. <https://doi.org/10.1080/09515089.2011.627537>
- Holmes, A. P., & Friston, K. J. (1998). Generalisability, Random Effects & Population Inference. *NeuroImage*. [https://doi.org/10.1016/S1053-8119\(18\)31587-8](https://doi.org/10.1016/S1053-8119(18)31587-8)
- Holmes, K. J., & Wolff, P. (2012). Does categorical perception in the left hemisphere depend on language? *Journal of Experimental Psychology. General*, 141(3), 439–443. <https://doi.org/10.1037/a0027289>
- Hough, M. S. (1993). Categorization in aphasia: Access and organization of goal-derived and common categories. *Aphasiology*, 7(4), 335–357. <https://doi.org/10.1080/02687039308249515>
- Howard, D., & Patterson, K. (1992). *The pyramids and palm trees test: A test of semantic access from words and pictures*. Pearson Assessment.
- Hu, J., Small, H., Kean, H., Takahashi, A., Zekelman, L., Kleinman, D., Ryan, E., Ferreira, V., & Fedorenko, E. (2021). *The language network supports both lexical access and sentence generation during language production* (p. 2021.09.10.459596). <https://doi.org/10.1101/2021.09.10.459596>
- Hu, Z., Yang, H., Yang, Y., Nishida, S., Madden-Lombardi, C., Ventre-Dominey, J., Dominey, P. F., & Ogawa, K. (2019). Common Neural System for Sentence and Picture Comprehension Across Languages: A Chinese–Japanese Bilingual Study. *Frontiers in Human Neuroscience*, 13. <https://doi.org/10.3389/fnhum.2019.00380>
- Huang, Y., Liu, X., Krueger, R., Santander, T., Hu, X., Leach, K., & Weimer, W. (2019). Distilling Neural Representations of Data Structure Manipulation using fMRI and fNIRS. *2019 IEEE/ACM 41st International Conference on Software Engineering (ICSE)*, 396–407. <https://doi.org/10.1109/ICSE.2019.00053>
- Hugdahl, K. (2011). Hemispheric asymmetry: Contributions from brain imaging. *Wiley Interdisciplinary Reviews. Cognitive Science*, 2(5), 461–478. <https://doi.org/10.1002/wcs.122>
- Hugdahl, K., Raichle, M. E., Mitra, A., & Specht, K. (2015). On the existence of a generalized non-specific task-dependent network. *Frontiers in Human Neuroscience*, 9. <https://doi.org/10.3389/fnhum.2015.00430>
- Humphreys, G. F., Halai, A. D., Branzi, F. M., & Ralph, M. A. L. (2022). *The angular gyrus is engaged by autobiographical recall not object-semantics, or event-semantics: Evidence from contrastive propositional speech production* (p. 2022.04.04.487000). bioRxiv. <https://doi.org/10.1101/2022.04.04.487000>
- Humphreys, G. F., Jung, J., & Lambon Ralph, M. A. (2022). The convergence and divergence of episodic and semantic functions across lateral parietal cortex. *Cerebral Cortex*, bhac044. <https://doi.org/10.1093/cercor/bhac044>
- Humphreys, G. F., Lambon Ralph, M. A., & Simons, J. S. (2021). A Unifying Account of Angular Gyrus Contributions to Episodic and Semantic Cognition. *Trends in Neurosciences*, 44(6), 452–463. <https://doi.org/10.1016/j.tins.2021.01.006>
- Humphreys, G. F., & Tibon, R. (2022). Dual-axes of functional organisation across lateral parietal cortex: The angular gyrus forms part of a multi-modal buffering system. *Brain Structure and Function*. <https://doi.org/10.1007/s00429-022-02510-0>
- Hurlburt, R. T., Alderson-Day, B., Kühn, S., & Fernyhough, C. (2016). Exploring the Ecological Validity of Thinking on Demand: Neural Correlates of Elicited vs. Spontaneously Occurring Inner Speech. *PLOS ONE*, 11(2), e0147932. <https://doi.org/10.1371/journal.pone.0147932>
- Hurlburt, R. T., Heavey, C. L., & Kelsey, J. M. (2013). Toward a phenomenology of inner speaking. *Consciousness and Cognition*, 22(4), 1477–1494. <https://doi.org/10.1016/j.concog.2013.10.003>
- Huth, A. G., Heer, W. A. de, Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, 532(7600), 453–458. <https://doi.org/10.1038/nature17637>

- Ikutani, Y., Kubo, T., Nishida, S., Hata, H., Matsumoto, K., Ikeda, K., & Nishimoto, S. (2020). Expert programmers have fine-tuned cortical representations of source code. *BioRxiv*, 2020.01.28.923953. <https://doi.org/10.1101/2020.01.28.923953>
- Ikutani, Y., & Uwano, H. (2014). Brain activity measurement during program comprehension with NIRS. *15th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, 1–6. <https://doi.org/10.1109/SNPD.2014.6888727>
- Isik, L., Koldewyn, K., Beeler, D., & Kanwisher, N. (2017). Perceiving social interactions in the posterior superior temporal sulcus. *Proceedings of the National Academy of Sciences*, 114(43), E9145–E9152. <https://doi.org/10.1073/pnas.1714471114>
- Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9(4), 304–313. <https://doi.org/10.1038/nrn2332>
- Ivanova, A. A., & Hofer, M. (2020). Linguistic Overhypotheses in Category Learning: Explaining the Label Advantage Effect. *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*, 723–729.
- Ivanova, A. A., Mineroff, Z., Zimmerer, V., Kanwisher, N., Varley, R., & Fedorenko, E. (2019). The language network is recruited but not required for non-verbal semantic processing. *BioRxiv*, 696484. <https://doi.org/10.1101/696484>
- Ivanova, A. A., Mineroff, Z., Zimmerer, V., Kanwisher, N., Varley, R., & Fedorenko, E. (2021). The Language Network is Recruited but Not Required for Nonverbal Event Semantics. *Neurobiology of Language*, 1–26. [https://doi.org/10.1162/nol\\_a\\_00030](https://doi.org/10.1162/nol_a_00030)
- Ivanova, A. A., Srikant, S., Sueoka, Y., Kean, H. H., Dhamala, R., O'Reilly, U.-M., Bers, M. U., & Fedorenko, E. (2020). Comprehension of computer code relies primarily on domain-general executive resources. *BioRxiv*, 2020.04.16.045732. <https://doi.org/10.1101/2020.04.16.045732>
- Jackendoff, R. (1987). The status of thematic relations in linguistic theory. *Linguistic Inquiry*, 18(3), 369–411.
- Jackendoff, R. (1996). How language helps us think. *Pragmatics & Cognition*, 4(1), 1–34. <https://doi.org/10.1075/pc.4.1.03jac>
- Jackendoff, R. (2009). Parallels and Nonparallels between Language and Music. *Music Perception*, 26(3), 195–204. <https://doi.org/10.1525/mp.2009.26.3.195>
- Jackson, R. L., Hoffman, P., Pobric, G., & Lambon Ralph, M. A. (2015). The Nature and Neural Correlates of Semantic Association versus Conceptual Similarity. *Cerebral Cortex*, 25(11), 4319–4333. <https://doi.org/10.1093/cercor/bhv003>
- Jacobi, H., Faber, J., Timmann, D., & Klockgether, T. (2021). Update cerebellum and cognition. *Journal of Neurology*, 268(10), 3921–3925. <https://doi.org/10.1007/s00415-021-10486-w>
- Jacoby, N., Bruneau, E., Koster-Hale, J., & Saxe, R. (2016). Localizing Pain Matrix and Theory of Mind networks with both verbal and non-verbal stimuli. *NeuroImage*, 126, 39–48. <https://doi.org/10.1016/j.neuroimage.2015.11.025>
- Jacoby, N., & Fedorenko, E. (2020). Discourse-level comprehension engages medial frontal Theory of Mind brain regions even for expository texts. *Language, Cognition and Neuroscience*, 35(6), 780–796. <https://doi.org/10.1080/23273798.2018.1525494>
- Jain, S., & Huth, A. G. (2018). Incorporating Context into Language Encoding Models for fMRI. *BioRxiv*, 327601. <https://doi.org/10.1101/327601>
- Jain, S., Vo, V. A., Mahto, S., LeBel, A., Turek, J. S., & Huth, A. G. (2020). Interpretable multi-timescale models for predicting fMRI responses to continuous natural speech. *BioRxiv*, 2020.10.02.324392. <https://doi.org/10.1101/2020.10.02.324392>
- Jang, G., Yoon, S., Lee, S.-E., Park, H., Kim, J., Ko, J. H., & Park, H.-J. (2013). Everyday conversation requires cognitive inference: Neural bases of comprehending implicated meanings in conversations. *NeuroImage*, 81, 61–72. <https://doi.org/10.1016/j.neuroimage.2013.05.027>

- Jefferies, E. (2013). The neural basis of semantic cognition: Converging evidence from neuropsychology, neuroimaging and TMS. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 49(3), 611–625. <https://doi.org/10.1016/j.cortex.2012.10.008>
- Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: A case-series comparison. *Brain*, 129(8), 2132–2147. <https://doi.org/10.1093/brain/awl153>
- Jefferies, E., Thompson, H., Cornelissen, P., & Smallwood, J. (2020). The neurocognitive basis of knowledge about object identity and events: Dissociations reflect opposing effects of semantic coherence and control. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1791). <https://doi.org/10.1098/rstb.2019.0300>
- Jouen, A.-L., Cazin, N., Hidot, S., Madden-Lombardi, C., Ventre-Dominey, J., & Dominey, P. F. (2019). Beyond the word and image: III. Neurodynamic properties of the semantic network. *BioRxiv*, 767384. <https://doi.org/10.1101/767384>
- Jouen, A.-L., Ellmore, T. M., Madden, C. J., Pallier, C., Dominey, P. F., & Ventre-Dominey, J. (2015). Beyond the word and image: Characteristics of a common meaning system for language and vision revealed by functional and structural imaging. *NeuroImage*, 106, 72–85. <https://doi.org/10.1016/j.neuroimage.2014.11.024>
- Jouravlev, O., Zheng, D., Balewski, Z., Pongos, A. L. A., Levan, Z., Goldin-Meadow, S., & Fedorenko, E. (2019). Speech-accompanying gestures are not processed by the language-processing mechanisms. *Neuropsychologia*, 132, 107132. <https://doi.org/10.1016/j.neuropsychologia.2019.107132>
- Julian, J. B., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *NeuroImage*, 60(4), 2357–2364. <https://doi.org/10.1016/j.neuroimage.2012.02.055>
- Jung, J., Rice, G., & Lambon Ralph, M. (2021). The neural bases of resilient semantic system: Evidence of variable neuro-displacement in cognitive systems. *Brain Structure and Function*. <https://doi.org/10.17863/CAM.66668>
- Kahn, H. J., & Whitaker, H. A. (1991). Acalculia: An historical review of localization. *Brain and Cognition*, 17(2), 102–115. [https://doi.org/10.1016/0278-2626\(91\)90071-F](https://doi.org/10.1016/0278-2626(91)90071-F)
- Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences*, 107(25), 11163–11170. <https://doi.org/10.1073/pnas.1005062107>
- Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic Assessments of Language Processing in Aphasia* (Hove, UK: Lawrence Earlbaum Associates).
- Keller, T. A., Carpenter, P. A., & Just, M. A. (2001). The Neural Bases of Sentence Comprehension: A fMRI Examination of Syntactic and Lexical Processing. *Cerebral Cortex*, 11(3), 223–237. <https://doi.org/10.1093/cercor/11.3.223>
- Kim, H. S. (2002). We talk, therefore we think? A cultural analysis of the effect of talking on thinking. *Journal of Personality and Social Psychology*, 83(4), 828–842. <https://doi.org/10.1037/0022-3514.83.4.828>
- Klare, G. R. (1963). *Measurement of readability*. Iowa State University Press.
- Klessinger, N., Szczerbinski, M., & Varley, R. A. (2007). Algebra in a man with severe aphasia. *Neuropsychologia*, 45(8), 1642–1648. <https://doi.org/10.1016/j.neuropsychologia.2007.01.005>
- Klinger, E., & Cox, W. M. (1987). Dimensions of Thought Flow in Everyday Life. *Imagination, Cognition and Personality*, 7(2), 105–128. <https://doi.org/10.2190/7K24-G343-MTQW-115V>
- Koechlin, E., & Jubault, T. (2006). Broca's Area and the Hierarchical Organization of Human Behavior. *Neuron*, 50(6), 963–974. <https://doi.org/10.1016/j.neuron.2006.05.017>
- Koameda-Lutz, M., Cohen, R., & Meier, E. (1987). Organization of and access to semantic memory in aphasia. *Brain and Language*, 30(2), 321–337. [https://doi.org/10.1016/0093-934X\(87\)90106-4](https://doi.org/10.1016/0093-934X(87)90106-4)

- Krieger-Redwood, K., Teige, C., Davey, J., Hymers, M., & Jefferies, E. (2015). Conceptual control across modalities: Graded specialisation for pictures and words in inferior frontal and posterior temporal cortex. *Neuropsychologia*, 76, 92–107. <https://doi.org/10.1016/j.neuropsychologia.2015.02.030>
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience*, 12(5), 535–540. <https://doi.org/10.1038/nn.2303>
- Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, Mind, and Brain. *Annual Review of Linguistics*, 1, 377–394. <https://doi.org/10.1146/annurev-linguist-030514-124937>
- Krueger, R., Huang, Y., Liu, X., Santander, T., Weimer, W., & Leach, K. (2020). Neurological Divide: An fMRI Study of Prose and Code Writing. *2020 IEEE/ACM 42nd International Conference on Software Engineering (ICSE)*, 13.
- Kuhnke, P., Chapman, C. A., Cheung, V. K. M., Turker, S., Graessner, A., Martin, S., Williams, K. A., & Hartwigsen, G. (2022). The role of the angular gyrus in semantic cognition: A synthesis of five functional neuroimaging studies. *Brain Structure & Function*. <https://doi.org/10.1007/s00429-022-02493-y>
- Kuperberg, G. R., Holcomb, P. J., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct Patterns of Neural Modulation during the Processing of Conceptual and Syntactic Anomalies. *Journal of Cognitive Neuroscience*, 15(2), 272–293. <https://doi.org/10.1162/089892903321208204>
- Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, 62(1), 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203–205.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(1), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lambon Ralph, M. A., Cipolotti, L., Manes, F., & Patterson, K. (2010). Taking both sides: Do unilateral anterior temporal lobe lesions disrupt semantic memory? *Brain*, 133(11), 3243–3255. <https://doi.org/10.1093/brain/awq264>
- Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42–55. <https://doi.org/10.1038/nrn.2016.150>
- Lanzoni, L., Ravasio, D., Thompson, H., Vatansever, D., Margulies, D., Smallwood, J., & Jefferies, E. (2020). The role of default mode network in semantic cue integration. *NeuroImage*, 219, 117019. <https://doi.org/10.1016/j.neuroimage.2020.117019>
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933. <https://doi.org/10.1038/nrn2532>
- Lau, E. F., Weber, K., Gramfort, A., Härmäläinen, M. S., & Kuperberg, G. R. (2016). Spatiotemporal Signatures of Lexical-Semantic Prediction. *Cerebral Cortex (New York, N.Y.: 1991)*, 26(4), 1377–1387. <https://doi.org/10.1093/cercor/bhu219>
- Le Dorze, G., & Nespolous, J. L. (1989). Anomia in moderate aphasia: Problems in accessing the lexical representation. *Brain and Language*, 37(3), 381–400. [https://doi.org/10.1016/0093-934x\(89\)90026-6](https://doi.org/10.1016/0093-934x(89)90026-6)
- LeBel, A., Jain, S., & Huth, A. G. (2021). Voxelwise Encoding Models Show That Cerebellar Language Representations Are Highly Conceptual. *Journal of Neuroscience*, 41(50), 10341–10355. <https://doi.org/10.1523/JNEUROSCI.0118-21.2021>
- Lecours, A. R., & Joanette, Y. (1980). Linguistic and other psychological aspects of paroxysmal aphasia. *Brain and Language*, 10(1), 1–23. [https://doi.org/10.1016/0093-934x\(80\)90034-6](https://doi.org/10.1016/0093-934x(80)90034-6)

- Lee Masson, H., & Isik, L. (2021). Functional selectivity for social interaction perception in the human superior temporal sulcus during natural viewing. *NeuroImage*, 245, 118741. <https://doi.org/10.1016/j.neuroimage.2021.118741>
- Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia*, 41(14), 1942–1958. [https://doi.org/10.1016/S0028-3932\(03\)00123-4](https://doi.org/10.1016/S0028-3932(03)00123-4)
- Lenartowicz, A., Kalar, D. J., Congdon, E., & Poldrack, R. A. (2010). Towards an Ontology of Cognitive Control. *Topics in Cognitive Science*, 2(4), 678–692. <https://doi.org/10.1111/j.1756-8765.2010.01100.x>
- Lerdahl, F., & Jackendoff, R. S. (1996). *A Generative Theory of Tonal Music*. MIT Press.
- Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic Mapping of a Hierarchy of Temporal Receptive Windows Using a Narrated Story. *The Journal of Neuroscience*, 31(8), 2906–2915. <https://doi.org/10.1523/JNEUROSCI.3684-10.2011>
- Lewis, G. A., Poeppel, D., & Murphy, G. L. (2015). The neural bases of taxonomic and thematic conceptual relations: An MEG study. *Neuropsychologia*. <https://doi.org/10.1016/j.neuropsychologia.2015.01.011>
- Li, P., Abarbanell, L., Gleitman, L., & Papafragou, A. (2011). Spatial Reasoning in Tenejapan Mayans. *Cognition*, 120(1), 33–53. <https://doi.org/10.1016/j.cognition.2011.02.012>
- Li, P., & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition*, 83(3), 265–294. [https://doi.org/10.1016/S0010-0277\(02\)00009-4](https://doi.org/10.1016/S0010-0277(02)00009-4)
- Lichtheim, L. (1885). On aphasia. *Brain*, 7, 433–484.
- Lipkin, B., Tuckute, G., Affourtit, J., Small, H., Mineroff, Z., Kean, H., Jouravlev, O., Rakocevic, L., Pritchett, B., Siegelman, M., Hoeflin, C., Pongos, A., Blank, I. A., Struhl, M. K., Ivanova, A., Shannon, S., Sathe, A., Hoffmann, M., Nieto-Castañón, A., & Fedorenko, E. (2022). *LanA (Language Atlas): A probabilistic atlas for the language network based on fMRI data from >800 individuals* (p. 2022.03.06.483177). bioRxiv. <https://doi.org/10.1101/2022.03.06.483177>
- Liu, Y., Kim, J., Wilson, C., & Bedny, M. (2020). Computer code comprehension shares neural resources with formal logical inference in the fronto-parietal network. *BioRxiv*, 2020.05.24.096180. <https://doi.org/10.1101/2020.05.24.096180>
- Lombardi, L., & Potter, M. C. (1992). The regeneration of syntax in short term memory. *Journal of Memory and Language*, 31(6), 713–733. [https://doi.org/10.1016/0749-596X\(92\)90036-W](https://doi.org/10.1016/0749-596X(92)90036-W)
- Lubbock, T. (2010). A memoir of living with a brain tumour. *The Observer*, 7, 2010.
- Lubbock, T. (2012). *Until Further Notice, I Am Alive*. Granta Publications.
- Lucy, L., & Gauthier, J. (2017). Are distributional representations ready for the real world? Evaluating word vectors for grounded perceptual meaning. *ArXiv:1705.11168 [Cs]*. <http://arxiv.org/abs/1705.11168>
- Luo, X., Sexton, N. J., & Love, B. C. (2021). A deep learning account of how language affects thought. *Language, Cognition and Neuroscience*, 1–10. <https://doi.org/10.1080/23273798.2021.2001023>
- Lupyan, G. (2009). Extracommunicative functions of language: Verbal interference causes selective categorization impairments. *Psychonomic Bulletin & Review*, 16(4), 711–718. <https://doi.org/10.3758/PBR.16.4.711>
- Lupyan, G. (2012). Linguistically Modulated Perception and Cognition: The Label-Feedback Hypothesis. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00054>
- Lupyan, G., & Casasanto, D. (2015). Meaningless words promote meaningful categorization. *Language and Cognition: An Interdisciplinary Journal of Language and Cognitive Science*, 7(2), 167–193. <https://doi.org/10.1017/langcog.2014.21>
- Lupyan, G., & Mirman, D. (2013). Linking language and categorization: Evidence from aphasia. *Cortex*, 49(5), 1187–1194. <https://doi.org/10.1016/j.cortex.2012.06.006>

- Lupyan, G., Mirman, D., Hamilton, R., & Thompson-Schill, S. L. (2012). Categorization is modulated by transcranial direct current stimulation over left prefrontal cortex. *Cognition*, 124(1), 36–49. <https://doi.org/10.1016/j.cognition.2012.04.002>
- Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not Just for Talking: Redundant Labels Facilitate Learning of Novel Categories. *Psychological Science*, 18(12), 1077–1083. <https://doi.org/10.1111/j.1467-9280.2007.02028.x>
- Luria, A. R., Tsvetkova, L. S., & Futer, D. S. (1965). Aphasia in a composer (V. G. Shebalin). *Journal of the Neurological Sciences*, 2(3), 288–292. [https://doi.org/10.1016/0022-510x\(65\)90113-9](https://doi.org/10.1016/0022-510x(65)90113-9)
- MacGregor, L. J., Gilbert, R. A., Balewski, Z., Mitchell, D. J., Erzinclioglu, S. W., Rodd, J. M., Duncan, J., Fedorenko, E., & Davis, M. H. (2022). *Causal contributions of the domain-general (Multiple Demand) and the language-selective brain networks to perceptual and semantic challenges in speech comprehension* (p. 2022.04.12.487989). bioRxiv. <https://doi.org/10.1101/2022.04.12.487989>
- MacSweeney, M., Woll, B., Campbell, R., McGuire, P. K., David, A. S., Williams, S. C. R., Suckling, J., Calvert, G. A., & Brammer, M. J. (2002). Neural systems underlying British Sign Language and audio-visual English processing in native users. *Brain*, 125(7), 1583–1593. <https://doi.org/10.1093/brain/awf153>
- Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: Studies with artificial lexicons. *Journal of Experimental Psychology. General*, 132(2), 202–227.
- Mahon, B. Z. (2015). What is embodied about cognition? *Language, Cognition and Neuroscience*, 30(4), 420–429. <https://doi.org/10.1080/23273798.2014.987791>
- Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, 102(1), 59–70. <https://doi.org/10.1016/j.jphysparis.2008.03.004>
- Mahowald, K., & Fedorenko, E. (2016). Reliable individual-level neural markers of high-level language processing: A necessary precursor for relating neural variability to behavioral and genetic variability. *NeuroImage*, 139, 74–93. <https://doi.org/10.1016/j.neuroimage.2016.05.073>
- Malik-Moraleda, S., Ayyash, D., Gallée, J., Affourtit, J., Hoffmann, M., Mineroff, Z., Jouravlev, O., & Fedorenko, E. (2022). *The universal language network: A cross-linguistic investigation spanning 45 languages and 12 language families* (p. 2021.07.28.454040). bioRxiv. <https://doi.org/10.1101/2021.07.28.454040>
- Mareschal, D., & Quinn, P. C. (2001). Categorization in infancy. *Trends in Cognitive Sciences*, 5(10), 443–450. [https://doi.org/10.1016/S1364-6613\(00\)01752-6](https://doi.org/10.1016/S1364-6613(00)01752-6)
- Mariën, P., & Borgatti, R. (2018). Language and the cerebellum. *Handbook of Clinical Neurology*, 154, 181–202. <https://doi.org/10.1016/B978-0-444-63956-1.00011-4>
- Marshall, J., Pring, T., & Chiat, S. (1993). Sentence processing therapy: Working at the level of the event. *Aphasiology*, 7(2), 177–199. <https://doi.org/10.1080/02687039308249505>
- Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology*, 11(2), 194–201. [https://doi.org/10.1016/S0959-4388\(00\)00196-3](https://doi.org/10.1016/S0959-4388(00)00196-3)
- Martinovic, J., Paramei, G. V., & MacInnes, W. J. (2020). Russian blues reveal the limits of language influencing colour discrimination. *Cognition*, 201, 104281. <https://doi.org/10.1016/j.cognition.2020.104281>
- Matchin, W. (2018). A neuronal retuning hypothesis of sentence-specificity in Broca's area. *Psychonomic Bulletin & Review*, 25(5), 1682–1694. <https://doi.org/10.3758/s13423-017-1377-6>
- Mather, M., Cacioppo, J. T., & Kanwisher, N. (2013). How fMRI Can Inform Cognitive Theories. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 8(1), 108–113. <https://doi.org/10.1177/1745691612469037>

- Matsumoto, A., Iidaka, T., Haneda, K., Okada, T., & Sadato, N. (2005). Linking semantic priming effect in functional MRI and event-related potentials. *NeuroImage*, 24(3), 624–634. <https://doi.org/10.1016/j.neuroimage.2004.09.008>
- McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, 4(4), 310–322. <https://doi.org/10.1038/nrn1076>
- McNamara, W. J. (1967). The Selection of Computer Personnel: Past, Present, Future. *Proceedings of the Fifth SIGCPR Conference on Computer Personnel Research*, 52–56. <https://doi.org/10.1145/1142662.1142667>
- Menenti, L., Gierhan, S. M. E., Segaert, K., & Hagoort, P. (2011). Shared language: Overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychological Science*, 22(9), 1173–1182. <https://doi.org/10.1177/0956797611418347>
- Mervis, C. B., & Rosch, E. (1981). Categorization of Natural Objects. *Annual Review of Psychology*, 32(1), 89–115. <https://doi.org/10.1146/annurev.ps.32.020181.000513>
- Meyer, A. S., Belke, E., Telling, A. L., & Humphreys, G. W. (2007). Early activation of object names in visual search. *Psychonomic Bulletin & Review*, 14(4), 710–716.
- Michalka, S. W., Kong, L., Rosen, M. L., Shinn-Cunningham, B. G., & Somers, D. C. (2015). Short-Term Memory for Space and Time Flexibly Recruit Complementary Sensory-Biased Frontal Lobe Attention Networks. *Neuron*, 87(4), 882–892. <https://doi.org/10.1016/j.neuron.2015.07.028>
- Micheloyannis, S., Sakkalis, V., Vourkas, M., Stam, C. J., & Simos, P. G. (2005). Neural networks involved in mathematical thinking: Evidence from linear and non-linear analysis of electroencephalographic activity. *Neuroscience Letters*, 373(3), 212–217. <https://doi.org/10.1016/j.neulet.2004.10.005>
- Milberg, W., & Blumstein, S. E. (1981). Lexical decision and aphasia: Evidence for semantic processing. *Brain and Language*, 14(2), 371–385. [https://doi.org/10.1016/0093-934X\(81\)90086-9](https://doi.org/10.1016/0093-934X(81)90086-9)
- Miller, E. K., & Cohen, J. D. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience*, 24(1), 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>
- Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural Mechanisms of Visual Working Memory in Prefrontal Cortex of the Macaque. *Journal of Neuroscience*, 16(16), 5154–5167. <https://doi.org/10.1523/JNEUROSCI.16-16-05154.1996>
- Mineroff, Z., Blank, I. A., Mahowald, K., & Fedorenko, E. (2018). A robust dissociation among the language, multiple demand, and default mode networks: Evidence from inter-region correlations in effect size. *Neuropsychologia*, 119, 501–511. <https://doi.org/10.1016/j.neuropsychologia.2018.09.011>
- Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y. T., Fryer, T. D., Williams, G. B., Hodges, J. R., & Nestor, P. J. (2010). What the left and right anterior fusiform gyri tell us about semantic memory. *Brain*, 133(11), 3256–3268. <https://doi.org/10.1093/brain/awq272>
- Mirman, D., Landigan, J.-F., & Britt, A. E. (2017). Taxonomic and Thematic Semantic Systems. *Psychological Bulletin*, 143(5), 499–520. <https://doi.org/10.1037/bul0000092>
- Mitchell, D. J., Bell, A. H., Buckley, M. J., Mitchell, A. S., Sallet, J., & Duncan, J. (2016). A Putative Multiple-Demand System in the Macaque Brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 36(33), 8574–8585. <https://doi.org/10.1523/JNEUROSCI.0810-16.2016>
- Mitchell, T. M., Shirkareva, S. V., Carlson, A., Chang, K.-M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *Science (New York, N.Y.)*, 320(5880), 1191–1195. <https://doi.org/10.1126/science.1152876>
- Moeller, M. P., & Schick, B. (2006). Relations Between Maternal Input and Theory of Mind Understanding in Deaf Children. *Child Development*, 77(3), 751–766. <https://doi.org/10.1111/j.1467-8624.2006.00901.x>

- Mollica, F., Siegelman, M., Diachek, E., Piantadosi, S. T., Mineroff, Z., Futrell, R., Kean, H., Qian, P., & Fedorenko, E. (2020). Composition is the Core Driver of the Language-selective Network. *Neurobiology of Language*, 1–30. [https://doi.org/10.1162/nol\\_a\\_00005](https://doi.org/10.1162/nol_a_00005)
- Monti, M. M., Osherson, D. N., Martinez, M. J., & Parsons, L. M. (2007). Functional neuroanatomy of deductive inference: A language-independent distributed network. *NeuroImage*, 37(3), 1005–1016. <https://doi.org/10.1016/j.neuroimage.2007.04.069>
- Monti, M. M., Parsons, L. M., & Osherson, D. N. (2009). The boundaries of language and thought in deductive inference. *Proceedings of the National Academy of Sciences*, 106(30), 12554–12559. <https://doi.org/10.1073/pnas.0902422106>
- Monti, M. M., Parsons, L. M., & Osherson, D. N. (2012). Thought beyond language: Neural dissociation of algebra and natural language. *Psychological Science*, 23(8), 914–922. <https://doi.org/10.1177/0956797612437427>
- Mukamel, R., & Fried, I. (2012). Human Intracranial Recordings and Cognitive Neuroscience. *Annual Review of Psychology*, 63(1), 511–537. <https://doi.org/10.1146/annurev-psych-120709-145401>
- Murdoch, B. E. (2010). The cerebellum and language: Historical perspective and review. *Cortex*, 46(7), 858–868. <https://doi.org/10.1016/j.cortex.2009.07.018>
- Murnane, J. S. (1993). The psychology of computer languages for introductory programming courses. *New Ideas in Psychology*, 11(2), 213–228. [https://doi.org/10.1016/0732-118X\(93\)90035-C](https://doi.org/10.1016/0732-118X(93)90035-C)
- Murphy, G. (2002). *The big book of concepts*. MIT Press.
- Nakagawa, T., Kamei, Y., Uwano, H., Monden, A., Matsumoto, K., & German, D. M. (2014). Quantifying Programmers' Mental Workload During Program Comprehension Based on Cerebral Blood Flow Measurement: A Controlled Experiment. *Companion Proceedings of the 36th International Conference on Software Engineering*, 448–451. <https://doi.org/10.1145/2591062.2591098>
- Nakai, T., Yamaguchi, H. Q., & Nishimoto, S. (2020). Convergence of modality invariance and attention selectivity in the cortical semantic circuit. *BioRxiv*, 2020.06.19.160960. <https://doi.org/10.1101/2020.06.19.160960>
- Nakamura, M., Monden, A., Itoh, T., Matsumoto, K., Kanzaki, Y., & Satoh, H. (2003). Queue-based cost evaluation of mental simulation process in program comprehension. *Proceedings. 5th International Workshop on Enterprise Networking and Computing in Healthcare Industry (IEEE Cat. No.03EX717)*, 351–360. <https://doi.org/10.1109/METRIC.2003.1232480>
- Nieto-Castañon, A. (2020). *Handbook of functional connectivity Magnetic Resonance Imaging methods in CONN*. Hilbert Press.
- Nieto-Castañon, A., & Fedorenko, E. (2012). Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *NeuroImage*, 63(3), 1646–1669. <https://doi.org/10.1016/j.neuroimage.2012.06.065>
- Nieuwland, M. S., Martin, A. E., & Carreiras, M. (2012). Brain regions that process case: Evidence from Basque. *Human Brain Mapping*, 33(11), 2509–2520.
- Niu, M., & Palomero-Gallagher, N. (2022). Architecture and connectivity of the human angular gyrus and of its homolog region in the macaque brain. *Brain Structure and Function*. <https://doi.org/10.1007/s00429-022-02509-7>
- Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition. *Neuron*, 88(6), 1281–1296. <https://doi.org/10.1016/j.neuron.2015.11.035>
- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's Area and Language Processing: Evidence for the Cognitive Control Connection. *Language and Linguistics Compass*, 4(10), 906–924. <https://doi.org/10.1111/j.1749-818X.2010.00244.x>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)

- Oller Jr., J. W. (1981). Language as Intelligence? *Language Learning*, 31(2), 465–492. <https://doi.org/10.1111/j.1467-1770.1981.tb01394.x>
- Ormerod, T. (1990). *Human Cognition and Programming*. <https://doi.org/10.1016/B978-0-12-350772-3.50009-4>
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 45(3), 255–287. <https://doi.org/10.1037/h0084295>
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108(1), 155–184. <https://doi.org/10.1016/j.cognition.2008.02.007>
- Papert, S. A. (1972). Teaching Children to be Mathematicians Versus Teaching About Mathematics. *International Journal of Mathematical Education in Science and Technology*, 3(3), 249–262. <https://doi.org/10.1080/0020739700030306>
- Papert, S. A. (1993). *Mindstorms: Children, Computers, And Powerful Ideas* (2 edition). Basic Books.
- Parvizi, J., & Kastner, S. (2018). Promises and limitations of human intracranial electroencephalography. *Nature Neuroscience*, 21(4), 474–483. <https://doi.org/10.1038/s41593-018-0108-2>
- Patel, A., Bhattacharya, S., & Goyal, N. (2021). Are NLP Models really able to Solve Simple Math Word Problems? *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2080–2094. <https://doi.org/10.18653/v1/2021.nacl-main.168>
- Patel, A. D., Iversen, J. R., Wassenaar, M., & Hagoort, P. (2008). Musical syntactic processing in agrammatic Broca's aphasia. *Aphasiology*, 22(7–8), 776–789. <https://doi.org/10.1080/02687030701803804>
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), 976–987. <https://doi.org/10.1038/nrn2277>
- Paunov, A. M., Blank, I. A., & Fedorenko, E. (2019). Functionally distinct language and Theory of Mind networks are synchronized at rest and during language comprehension. *Journal of Neurophysiology*, 121(4), 1244–1265. <https://doi.org/10.1152/jn.00619.2018>
- Paunov, A. M., Blank, I. A., Jouravlev, O., Mineroff, Z., Gallée, J., & Fedorenko, E. (2022). Differential Tracking of Linguistic vs. Mental State Content in Naturalistic Stimuli by Language and Theory of Mind (ToM) Brain Networks. *Neurobiology of Language*, 1–29. [https://doi.org/10.1162/nol\\_a\\_00071](https://doi.org/10.1162/nol_a_00071)
- Pea, R. D., & Kurland, D. M. (1984). On the cognitive effects of learning computer programming. [https://doi.org/10.1016/0732-118X\(84\)90018-7](https://doi.org/10.1016/0732-118X(84)90018-7)
- Pearce, J. M. (1994). Discrimination and Categorization. In N. J. Mackintosh (Ed.), *Animal Learning and Cognition* (pp. 109–134). Academic Press. <https://doi.org/10.1016/B978-0-08-057169-0.50011-5>
- Pennington, N., & Grabowski, B. (1990). The tasks of programming. *Psychology of Programming*, 45–62.
- Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., Botvinick, M., & Fedorenko, E. (2018). Toward a universal decoder of linguistic meaning from brain activation. *Nature Communications*, 9(1), 1–13. <https://doi.org/10.1038/s41467-018-03068-4>
- Perkins, D. N., & Simmons, R. (1988). Patterns of Misunderstanding: An Integrative Model for Science, Math, and Programming. *Review of Educational Research*, 58(3), 303–326. <https://doi.org/10.3102/00346543058003303>
- Peterson, C. C. (2016). Empathy and Theory of Mind in Deaf and Hearing Children. *The Journal of Deaf Studies and Deaf Education*, 21(2), 141–147. <https://doi.org/10.1093/deafed/env058>

- Peterson, C. C., & Siegal, M. (1995). Deafness, Conversation and Theory of Mind. *Journal of Child Psychology and Psychiatry*, 36(3), 459–474. <https://doi.org/10.1111/j.1469-7610.1995.tb01303.x>
- Peterson, C. C., & Slaughter, V. P. (2006). Telling the story of theory of mind: Deaf and hearing children's narratives and mental state understanding. *British Journal of Developmental Psychology*, 24(1), 151–179. <https://doi.org/10.1348/026151005X60022>
- Petroni, F., Rocktäschel, T., Lewis, P., Bakhtin, A., Wu, Y., Miller, A. H., & Riedel, S. (2019). Language Models as Knowledge Bases? *ArXiv:1909.01066 [Cs]*. <http://arxiv.org/abs/1909.01066>
- Pinel, P., & Dehaene, S. (2009). Beyond Hemispheric Dominance: Brain Regions Underlying the Joint Lateralization of Language and Arithmetic to the Left Hemisphere. *Journal of Cognitive Neuroscience*, 22(1), 48–66. <https://doi.org/10.1162/jocn.2009.21184>
- Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, 41(6), 983–993. [https://doi.org/10.1016/s0896-6273\(04\)00107-2](https://doi.org/10.1016/s0896-6273(04)00107-2)
- Pinker, & Levin, B. (1991). *Lexical and Conceptual Semantics*. Cambridge, MA: MIT Press.
- Pitcher, D., Duchaine, B., Walsh, V., Yovel, G., & Kanwisher, N. (2011). The role of lateral occipital face and object areas in the face inversion effect. *Neuropsychologia*, 49(12), 3448–3453. <https://doi.org/10.1016/j.neuropsychologia.2011.08.020>
- Pitt, B., Gibson, E., & Piantadosi, S. T. (2022). Exact Number Concepts Are Limited to the Verbal Count Range. *Psychological Science*, 33(3), 371–381. <https://doi.org/10.1177/09567976211034502>
- Plunkett, K., Hu, J.-F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. *Cognition*, 106(2), 665–681. <https://doi.org/10.1016/j.cognition.2007.04.003>
- Pobric, G., Jefferies, E., & Ralph, M. A. L. (2010). Amodal semantic representations depend on both anterior temporal lobes: Evidence from repetitive transcranial magnetic stimulation. *Neuropsychologia*, 48(5), 1336–1342. <https://doi.org/10.1016/j.neuropsychologia.2009.12.036>
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10(2), 59–63. <https://doi.org/10.1016/j.tics.2005.12.004>
- Poldrack, R. A. (2008). The role of fMRI in cognitive neuroscience: Where do we stand? *Current Opinion in Neurobiology*, 18(2), 223–227. <https://doi.org/10.1016/j.conb.2008.07.006>
- Poldrack, R. A. (2011). Inferring mental states from neuroimaging data: From reverse inference to large-scale decoding. *Neuron*, 72(5), 692–697. <https://doi.org/10.1016/j.neuron.2011.11.001>
- Poldrack, R. A., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., Parker, D., Sabb, F., & Bilder, R. (2011). The Cognitive Atlas: Toward a Knowledge Foundation for Cognitive Neuroscience. *Frontiers in Neuroinformatics*, 5. <https://www.frontiersin.org/articles/10.3389/fninf.2011.00017>
- Pontillo, D. F., Salverda, A. P., & Tanenhaus, M. K. (2015). Flexible Use of Phonological and Visual Memory in Language-mediated Visual Search. *Proceedings of the 37th Meeting of the Cognitive Science Society*.
- Popham, S. F., Huth, A. G., Bilenko, N. Y., Deniz, F., Gao, J. S., Nunez-Elizalde, A. O., & Gallant, J. L. (2021). Visual and linguistic semantic representations are aligned at the border of human visual cortex. *Nature Neuroscience*, 24(11), 1628–1636. <https://doi.org/10.1038/s41593-021-00921-6>
- Potter, M. C., & Faulconer, B. A. (1975). Time to understand pictures and words. *Nature*, 253(5491), 437–438. <https://doi.org/10.1038/253437a0>
- Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: Understanding without words. *Journal of Experimental Psychology. General*, 115(3), 281–294.
- Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language*, 29(6), 633–654. [https://doi.org/10.1016/0749-596X\(90\)90042-X](https://doi.org/10.1016/0749-596X(90)90042-X)

- Potter, M. C., & Lombardi, L. (1998). Syntactic Priming in Immediate Recall of Sentences. *Journal of Memory and Language*, 38(3), 265–282. <https://doi.org/10.1006/jmla.1997.2546>
- Prabhakaran, V., Smith, J. A. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1997). Neural Substrates of Fluid Reasoning: An fMRI Study of Neocortical Activation during Performance of the Raven's Progressive Matrices Test. *Cognitive Psychology*, 33(1), 43–63. <https://doi.org/10.1006/cogp.1997.0659>
- Prat, C. S., Madhyastha, T. M., Mottarella, M. J., & Kuo, C.-H. (2020). Relating Natural Language Aptitude to Individual Differences in Learning Programming Languages. *Scientific Reports*, 10(1), 1–10. <https://doi.org/10.1038/s41598-020-60661-8>
- Price, A. R., Bonner, M. F., Peelle, J. E., & Grossman, M. (2015). Converging Evidence for the Neuroanatomic Basis of Combinatorial Semantics in the Angular Gyrus. *Journal of Neuroscience*, 35(7), 3276–3284. <https://doi.org/10.1523/JNEUROSCI.3446-14.2015>
- Price, C. J., & Friston, K. J. (2005). Functional ontologies for cognition: The systematic definition of structure and function. *Cognitive Neuropsychology*, 22(3–4), 262–275. <https://doi.org/10.1080/02643290442000095>
- Pritchett, B. L., Hoeflin, C., Koldewyn, K., Dechter, E., & Fedorenko, E. (2018). High-level language processing regions are not engaged in action observation or imitation. *Journal of Neurophysiology*, 120(5), 2555–2570. <https://doi.org/10.1152/jn.00222.2018>
- Proverbio, A. M., & Riva, F. (2009). RP and N400 ERP components reflect semantic violations in visual processing of human actions. *Neuroscience Letters*, 459(3), 142–146. <https://doi.org/10.1016/j.neulet.2009.05.012>
- Pujol, J., Deus, J., Losilla, J. M., & Capdevila, A. (1999). Cerebral lateralization of language in normal left-handed people studied by functional MRI. *Neurology*, 52(5), 1038–1043. <https://doi.org/10.1212/wnl.52.5.1038>
- Pulvermuller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, 22(2), 253–279. <https://doi.org/10.1017/S0140525X9900182X>
- Rabe, M. M., Vasishth, S., Hohenstein, S., Kliegl, R., & Schad, D. J. (2020). hypr: An R package for hypothesis-driven contrast coding. *Journal of Open Source Software*, 5(48), 2134. <https://doi.org/10.21105/joss.02134>
- Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, 38, 433–447. <https://doi.org/10.1146/annurev-neuro-071013-014030>
- Rapoport, M., van Reekum, R., & Mayberg, H. (2000). The Role of the Cerebellum in Cognition and Behavior. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 12(2), 193–198. <https://doi.org/10.1176/jnp.12.2.193>
- Raven, J., & Raven, J. (2003). Raven Progressive Matrices. In R. S. McCallum (Ed.), *Handbook of Nonverbal Assessment* (pp. 223–237). Springer US. [https://doi.org/10.1007/978-1-4615-0153-4\\_11](https://doi.org/10.1007/978-1-4615-0153-4_11)
- Regev, M., Honey, C. J., Simony, E., & Hasson, U. (2013). Selective and Invariant Neural Responses to Spoken and Written Narratives. *Journal of Neuroscience*, 33(40), 15978–15988. <https://doi.org/10.1523/JNEUROSCI.1580-13.2013>
- Regev, T. I., Affourtit, J., Chen, X., Schipper, A. E., Bergen, L., Mahowald, K., & Fedorenko, E. (2021). *High-level language brain regions are sensitive to sub-lexical regularities* (p. 2021.06.11.447786). bioRxiv. <https://doi.org/10.1101/2021.06.11.447786>
- Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences*, 13(10), 439–446. <https://doi.org/10.1016/j.tics.2009.07.001>
- Rehrig, G., Hayes, T. R., Henderson, J. M., & Ferreira, F. (2020). When scenes speak louder than words: Verbal encoding does not mediate the relationship between scene meaning and visual attention. *Memory & Cognition*. <https://doi.org/10.3758/s13421-020-01050-4>
- Reverberi, C., Shallice, T., D'Agostini, S., Skrap, M., & Bonatti, L. L. (2009). Cortical bases of elementary deductive reasoning: Inference, memory, and metadecision. *Neuropsychologia*, 47(4), 1107–1116. <https://doi.org/10.1016/j.neuropsychologia.2009.01.004>

- Richardson, H., Koster-Hale, J., Caselli, N., Magid, R., Benedict, R., Olson, H., Pyers, J., & Saxe, R. (2020). Reduced neural selectivity for mental states in deaf children with delayed exposure to sign language. *Nature Communications*, 11(1), 3246. <https://doi.org/10.1038/s41467-020-17004-y>
- Rissman, L., & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin & Review*, 26(6), 1850–1869. <https://doi.org/10.3758/s13423-019-01634-5>
- Roebuck, H., & Lupyan, G. (2020). The Internal Representations Questionnaire: Measuring modes of thinking. *Behavior Research Methods*, 52(5), 2053–2070. <https://doi.org/10.3758/s13428-020-01354-y>
- Rogalsky, C., Rong, F., Saberi, K., & Hickok, G. (2011). Functional anatomy of language and music perception: Temporal and structural factors investigated using functional magnetic resonance imaging. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 31(10), 3843–3852. <https://doi.org/10.1523/JNEUROSCI.4515-10.2011>
- Rosselli, M., & Ardila, A. (1989). Calculation deficits in patients with right and left hemisphere damage. *Neuropsychologia*, 27(5), 607–617. [https://doi.org/10.1016/0028-3932\(89\)90107-3](https://doi.org/10.1016/0028-3932(89)90107-3)
- Rossoni, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's Object Pictorial Set: The Role of Surface Detail in Basic-Level Object Recognition. *Perception*, 33(2), 217–236. <https://doi.org/10.1080/p5117>
- Sanner, M. F. (1999). Python: A programming language for software integration and development. *Journal of Molecular Graphics & Modelling*, 17(1), 57–61.
- Sapir, E. (1921). *Language: An introduction to the study of speech*. Harcourt, Brace.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, 16(2), 235–239. <https://doi.org/10.1016/j.conb.2006.03.001>
- Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: A defense of functional localizers. *NeuroImage*, 30(4), 1088–1096; discussion 1097-1099. <https://doi.org/10.1016/j.neuroimage.2005.12.062>
- Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in “theory of mind.” *NeuroImage*, 19(4), 1835–1842. [https://doi.org/10.1016/s1053-8119\(03\)00230-1](https://doi.org/10.1016/s1053-8119(03)00230-1)
- Saxe, R., Moran, J. M., Scholz, J., & Gabrieli, J. (2006). Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. *Social Cognitive and Affective Neuroscience*, 1(3), 229–234. <https://doi.org/10.1093/scan/nsl034>
- Saxe, R., & Powell, L. J. (2006). It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychological Science*, 17(8), 692–699. <https://doi.org/10.1111/j.1467-9280.2006.01768.x>
- Saygin, A. P., Wilson, S. M., Dronkers, N. F., & Bates, E. (2004). Action comprehension in aphasia: Linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia*, 42(13), 1788–1804. <https://doi.org/10.1016/j.neuropsychologia.2004.04.016>
- Saygin, A. P., Dick, F., W. Wilson, S., F. Dronkers, N., & Bates, E. (2003). Neural resources for processing language and environmental soundsEvidence from aphasia. *Brain*, 126(4), 928–945. <https://doi.org/10.1093/brain/awg082>
- Schapiro, A. C., McClelland, J. L., Welbourne, S. R., Rogers, T. T., & Lambon Ralph, M. A. (2013). Why Bilateral Damage Is Worse than Unilateral Damage to the Brain. *Journal of Cognitive Neuroscience*, 25(12), 2107–2123. [https://doi.org/10.1162/jocn\\_a\\_00441](https://doi.org/10.1162/jocn_a_00441)
- Schmahmann, J. D. (2019). The cerebellum and cognition. *Neuroscience Letters*, 688, 62–75. <https://doi.org/10.1016/j.neulet.2018.07.005>
- Schmahmann, J. D., Guell, X., Stoodley, C. J., & Halko, M. A. (2019). The Theory and Neuroscience of Cerebellar Cognition. *Annual Review of Neuroscience*, 42, 337–364. <https://doi.org/10.1146/annurev-neuro-070918-050258>
- Schrimpf, M., Blank, I. A., Tuckute, G., Kauf, C., Hosseini, E. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2021). The neural architecture of language: Integrative modeling

- converges on predictive processing. *Proceedings of the National Academy of Sciences*, 118(45). <https://doi.org/10.1073/pnas.2105646118>
- Schuster, S., & Linzen, T. (2022). When a sentence does not introduce a discourse entity, Transformer-based models still sometimes refer to it (arXiv:2205.03472). arXiv. <https://doi.org/10.48550/arXiv.2205.03472>
- Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Brecher, A., Faseyitan, O. K., Dell, G. S., Mirman, D., & Coslett, H. B. (2011). Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 108(20), 8520–8524. <https://doi.org/10.1073/pnas.1014935108>
- Scott, T. L., Gallée, J., & Fedorenko, E. (2017). A new fun and robust version of an fMRI localizer for the frontotemporal language system. *Cognitive Neuroscience*, 8(3), 167–176. <https://doi.org/10.1080/17588928.2016.1201466>
- Seed, A., & Tomasello, M. (2010). Primate Cognition. *Topics in Cognitive Science*, 2(3), 407–419. <https://doi.org/10.1111/j.1756-8765.2010.01099.x>
- Seghier, M. L. (2013). The Angular Gyrus. *The Neuroscientist*, 19(1), 43–61. <https://doi.org/10.1177/1073858412440596>
- Sevostianov, A., Horwitz, B., Nechaev, V., Williams, R., Fromm, S., & Braun, A. R. (2002). FMRI study comparing names versus pictures of objects. *Human Brain Mapping*, 16(3), 168–175. <https://doi.org/10.1002/hbm.10037>
- Shain, C., Blank, I. A., van Schijndel, M., Schuler, W., & Fedorenko, E. (2020). FMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia*, 138, 107307. <https://doi.org/10.1016/j.neuropsychologia.2019.107307>
- Shain, C., Paunov, A. M., Chen, X., Lipkin, B., & Fedorenko, E. (2022). No evidence of theory of mind reasoning in the human language network (p. 2022.07.18.500516). bioRxiv. <https://doi.org/10.1101/2022.07.18.500516>
- Shashidhara, S., Mitchell, D. J., Erez, Y., & Duncan, J. (2019). Progressive Recruitment of the Frontoparietal Multiple-demand System with Increased Task Complexity, Time Pressure, and Reward. *Journal of Cognitive Neuroscience*, 31(11), 1617–1630. [https://doi.org/10.1162/jocn\\_a\\_01440](https://doi.org/10.1162/jocn_a_01440)
- Shashidhara, S., Spronkers, F. S., & Erez, Y. (2019). Individual-subject functional localization increases univariate activation but not multivariate pattern discriminability in the ‘multiple-demand’ frontoparietal network. *BioRxiv*, 661934. <https://doi.org/10.1101/661934>
- Shashidhara, S., Spronkers, F. S., & Erez, Y. (2020). Individual-subject Functional Localization Increases Univariate Activation but Not Multivariate Pattern Discriminability in the “Multiple-demand” Frontoparietal Network. *Journal of Cognitive Neuroscience*, 32(7), 1348–1368. [https://doi.org/10.1162/jocn\\_a\\_01554](https://doi.org/10.1162/jocn_a_01554)
- Sheremata, S. L., Bettencourt, K. C., & Somers, D. C. (2010). Hemispheric Asymmetry in Visuotopic Posterior Parietal Cortex Emerges with Visual Short-Term Memory Load. *Journal of Neuroscience*, 30(38), 12581–12588. <https://doi.org/10.1523/JNEUROSCI.2689-10.2010>
- Shinkareva, S. V., Malave, V. L., Mason, R. A., Mitchell, T. M., & Just, M. A. (2011). Commonality of neural representations of words and pictures. *NeuroImage*, 54(3), 2418–2425. <https://doi.org/10.1016/j.neuroimage.2010.10.042>
- Shuman, M., & Kanwisher, N. (2004). Numerical Magnitude in the Human Parietal Lobe: Tests of Representational Generality and Domain Specificity. *Neuron*, 44(3), 557–569. <https://doi.org/10.1016/j.neuron.2004.10.008>
- Siegal, M., & Varley, R. (2006). Aphasia, language, and theory of mind. *Social Neuroscience*, 1(3–4), 167–174. <https://doi.org/10.1080/17470910600985597>
- Siegmund, J., Kästner, C., Apel, S., Parnin, C., Bethmann, A., Leich, T., Saake, G., & Brechmann, A. (2014). Understanding Understanding Source Code with Functional Magnetic

- Resonance Imaging. *Proceedings of the 36th International Conference on Software Engineering*, 378–389. <https://doi.org/10.1145/2568225.2568252>
- Siegmund, J., Peitek, N., Parnin, C., Apel, S., Hofmeister, J., Kästner, C., Begel, A., Bethmann, A., & Brechmann, A. (2017). Measuring Neural Efficiency of Program Comprehension. *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering*, 140–150. <https://doi.org/10.1145/3106237.3106268>
- Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Sciences*, 111(43), E4687–E4696. <https://doi.org/10.1073/pnas.1323812111>
- Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., & Hasson, U. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nature Communications*, 7(1), 1–13. <https://doi.org/10.1038/ncomms12141>
- Sitnikova, T., Holcomb, P. J., Kiyonaga, K. A., & Kuperberg, G. R. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience*, 20(11), 2037–2057. <https://doi.org/10.1162/jocn.2008.20143>
- Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology. General*, 133(2), 166–188. <https://doi.org/10.1037/0096-3445.133.2.166>
- Smet, H. J. D., Baillieux, H., Deyn, P. P. D., Mariën, P., & Paquier, P. (2007). The Cerebellum and Language: The Story So Far. *Folia Phoniatrica et Logopaedica*, 59(4), 165–170. <https://doi.org/10.1159/000102927>
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Harvard University Press.
- Smith, L. B., & Heise, D. (1992). Perceptual Similarity and Conceptual Structure. In B. Burns (Ed.), *Advances in Psychology* (Vol. 93, pp. 233–272). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)61009-2](https://doi.org/10.1016/S0166-4115(08)61009-2)
- Snowden, J. S., Harris, J. M., Thompson, J. C., Kobylecki, C., Jones, M., Richardson, A. M., & Neary, D. (2018). Semantic dementia and the left and right temporal lobes. *Cortex*, 107, 188–203. <https://doi.org/10.1016/j.cortex.2017.08.024>
- Sokolov, A. (1972). *Inner Speech and Thought*. Springer US. <https://doi.org/10.1007/978-1-4684-1914-6>
- Spelke, E. S. (1976). Infants' intermodal perception of events. *Cognitive Psychology*, 8(4), 553–560. [https://doi.org/10.1016/0010-0285\(76\)90018-9](https://doi.org/10.1016/0010-0285(76)90018-9)
- Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, 10(1), 89–96. <https://doi.org/10.1111/j.1467-7687.2007.00569.x>
- Spotorno, N., Koun, E., Prado, J., Van Der Henst, J.-B., & Noveck, I. A. (2012). Neural evidence that utterance-processing entails mentalizing: The case of irony. *NeuroImage*, 63(1), 25–39. <https://doi.org/10.1016/j.neuroimage.2012.06.046>
- Starowicz-Filip, A., Chrobak, A. A., Moskała, M., Krzyżewski, R. M., Kwinta, B., Kwiatkowski, S., Milczarek, O., Rajtar-Zembaty, A., & Przewoźnik, D. (2017). The role of the cerebellum in the regulation of language functions. *Psychiatria Polska*, 51(4), 661–671. <https://doi.org/10.12740/PP/68547>
- Strick, P. L., Dum, R. P., Fiez, J. A., & others. (2009). Cerebellum and nonmotor function. *Annual Review of Neuroscience*, 32(1), 413–434.
- Strickland, B. (2017). Language Reflects "Core" Cognition: A New Theory About the Origin of Cross-Linguistic Regularities. *Cognitive Science*, 41(1), 70–101. <https://doi.org/10.1111/cogs.12332>
- Sueoka, Y., Paunov, A., Ivanova, A., Blank, I. A., & Fedorenko, E. (2022). *The language network reliably 'tracks' naturalistic meaningful non-verbal stimuli* (p. 2022.04.24.489316). bioRxiv. <https://doi.org/10.1101/2022.04.24.489316>

- Sullivan, A., & Bers, M. U. (2019). Computer Science Education in Early Childhood: The Case of ScratchJr. *Journal of Information Technology Education: Innovations in Practice*, 18, 113–138.
- Swain, J. P. (1995). The Concept of Musical Syntax. *The Musical Quarterly*, 79(2), 281–308.
- Tahmasebi, A. M., Davis, M. H., Wild, C. J., Rodd, J. M., Hakyemez, H., Abolmaesumi, P., & Johnsrude, I. S. (2012). Is the Link between Anatomical Structure and Function Equally Strong at All Cognitive Levels of Processing? *Cerebral Cortex*, 22(7), 1593–1603. <https://doi.org/10.1093/cercor/bhr205>
- Takayama, Y., Sugishita, M., Akiguchi, I., & Kimura, J. (1994). Isolated Acalculia due to Left Parietal Lesion. *Archives of Neurology*, 51(3), 286–291. <https://doi.org/10.1001/archneur.1994.00540150084021>
- Talmor, A., Elazar, Y., Goldberg, Y., & Berant, J. (2020). oLMpics—On what Language Model Pre-training Captures (arXiv:1912.13283). arXiv. <https://doi.org/10.48550/arXiv.1912.13283>
- Talmy, L. (2000). *Toward a cognitive semantics* (Vol. 2). MIT press.
- Tarhan, L., & Konkle, T. (2020). Sociality and interaction envelope organize visual action representations. *Nature Communications*, 11(1), 3002. <https://doi.org/10.1038/s41467-020-16846-w>
- Taylor, J. B. (2006). *My stroke of insight*. Hachette UK.
- Taylor, J. B. (Director). (2008). *Jill Bolte Taylor: My stroke of insight* | TED Talk. [https://www.ted.com/talks/jill\\_bolte\\_taylor\\_my\\_stroke\\_of\\_insight](https://www.ted.com/talks/jill_bolte_taylor_my_stroke_of_insight)
- Teige, C., Cornelissen, P. L., Mollo, G., Gonzalez Alam, T. R. del J., McCarty, K., Smallwood, J., & Jefferies, E. (2019). Dissociations in semantic cognition: Oscillatory evidence for opposing effects of semantic control and type of semantic relation in anterior and posterior temporal cortex. *Cortex*, 120, 308–325. <https://doi.org/10.1016/j.cortex.2019.07.002>
- Tettamanti, M., & Weniger, D. (2006). Broca's Area: A Supramodal Hierarchical Processor? *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 42, 491–494. [https://doi.org/10.1016/S0010-9452\(08\)70384-8](https://doi.org/10.1016/S0010-9452(08)70384-8)
- Thierry, G., & Price, C. J. (2006). Dissociating verbal and nonverbal conceptual processing in the human brain. *Journal of Cognitive Neuroscience*, 18(6), 1018–1028. <https://doi.org/10.1162/jocn.2006.18.6.1018>
- Thompson, S. A., Patterson, K., & Hodges, J. R. (2003). Left/right asymmetry of atrophy in semantic dementia: Behavioral-cognitive implications. *Neurology*, 61(9), 1196–1203. <https://doi.org/10.1212/01.wnl.0000091868.28557.b8>
- Thompson-Schill, S. L., Swick, D., Farah, M. J., D'Esposito, M., Kan, I. P., & Knight, R. T. (1998). Verb generation in patients with focal frontal lesions: A neuropsychological test of neuroimaging findings. *Proceedings of the National Academy of Sciences of the United States of America*, 95(26), 15855–15860.
- Toneva, M., Stretcu, O., Poczos, B., Wehbe, L., & Mitchell, T. M. (2020). Modeling Task Effects on Meaning Representation in the Brain via Zero-Shot MEG Prediction. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), *Advances in Neural Information Processing Systems* (Vol. 33, pp. 5284–5295). Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2020/file/38a8e18d75e95ca619af8df0da1417f2-Paper.pdf>
- Tremblay, P., & Dick, A. S. (2016). Broca and Wernicke are dead, or moving past the classic model of language neurobiology. *Brain and Language*, 162, 60–71. <https://doi.org/10.1016/j.bandl.2016.08.004>
- Trueswell, J. C., & Papafragou, A. (2010). Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language*, 63(1), 64–82. <https://doi.org/10.1016/j.jml.2010.02.006>
- Tshitoyan, V., Dagdelen, J., Weston, L., Dunn, A., Rong, Z., Kononova, O., Persson, K. A., Ceder, G., & Jain, A. (2019). Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature*, 571(7763), 95–98. <https://doi.org/10.1038/s41586-019-1335-8>

- Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. F. (2015). No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. *Trends in Cognitive Sciences*, 19(12), 757–770. <https://doi.org/10.1016/j.tics.2015.10.002>
- Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(October), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Tyler, L. K., & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences*, 5(6), 244–252. [https://doi.org/10.1016/s1364-6613\(00\)01651-x](https://doi.org/10.1016/s1364-6613(00)01651-x)
- Uddin, L. Q., Supekhar, K., Amin, H., Rykhlevskaia, E., Nguyen, D. A., Greicius, M. D., & Menon, V. (2010). Dissociable Connectivity within Human Angular Gyrus and Intraparietal Sulcus: Evidence from Functional and Structural Connectivity. *Cerebral Cortex*, 20(11), 2636–2646. <https://doi.org/10.1093/cercor/bhq011>
- Unger, L., & Fisher, A. V. (2021). The emergence of richly organized semantic knowledge from simple statistics: A synthetic review. *Developmental Review*, 60, 100949. <https://doi.org/10.1016/j.dr.2021.100949>
- van Ackeren, M. J., Casasanto, D., Bekkering, H., Hagoort, P., & Rueschemeyer, S.-A. (2012). Pragmatics in action: Indirect requests engage theory of mind areas and the cortical motor network. *Journal of Cognitive Neuroscience*, 24(11), 2237–2247. [https://doi.org/10.1162/jocn\\_a\\_00274](https://doi.org/10.1162/jocn_a_00274)
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., & WU-Minn HCP Consortium. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage*, 80, 62–79. <https://doi.org/10.1016/j.neuroimage.2013.05.041>
- Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, 383(6597), 254–256. <https://doi.org/10.1038/383254a0>
- Varley, R. A. (1998). Aphasic language, aphasic thought: An investigation of propositional thinking in an a-propositional aphasic. In P. Carruthers & J. Boucher (Eds.), *Language and Thought: Interdisciplinary Themes* (pp. 128–145). Cambridge University Press. <https://doi.org/10.1017/CBO9780511597909.009>
- Varley, R. A., Klessinger, N. J. C., Romanowski, C. A. J., & Siegal, M. (2005). Agrammatic but numerate. *Proceedings of the National Academy of Sciences of the United States of America*, 102(9), 3519–3524. <https://doi.org/10.1073/pnas.0407470102>
- Varley, R. A., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and “theory of mind” in an agrammatic aphasic patient. *Current Biology: CB*, 10(12), 723–726. [https://doi.org/10.1016/s0960-9822\(00\)00538-8](https://doi.org/10.1016/s0960-9822(00)00538-8)
- Varley, R. A., Siegal, M., & Want, S. C. (2001). Severe impairment in grammar does not preclude theory of mind. *Neurocase*, 7(6), 489–493. <https://doi.org/10.1093/neucas/7.6.489>
- Vázquez-Rodríguez, B., Suárez, L. E., Markello, R. D., Shafiei, G., Paquola, C., Hagmann, P., Heuvel, M. P. van den, Bernhardt, B. C., Spreng, R. N., & Misic, B. (2019). Gradients of structure–function tethering across neocortex. *Proceedings of the National Academy of Sciences*, 116(42), 21219–21227. <https://doi.org/10.1073/pnas.1903403116>
- Visser, M., Jefferies, E., Embleton, K. V., & Lambon Ralph, M. A. (2012). Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: Distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *Journal of Cognitive Neuroscience*, 24(8), 1766–1778. [https://doi.org/10.1162/jocn\\_a\\_00244](https://doi.org/10.1162/jocn_a_00244)
- Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic processing in the anterior temporal lobes: A meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, 22(6), 1083–1094. <https://doi.org/10.1162/jocn.2009.21309>
- Võ, M. L.-H., & Wolfe, J. M. (2013). Differential ERP Signatures Elicited by Semantic and Syntactic Processing in Scenes. *Psychological Science*, 24(9), 1816–1823. <https://doi.org/10.1177/0956797613476955>

- Vul, E., & Kanwisher, N. (2010). Begging the question: The non-independence error in fMRI data analysis. *Foundational Issues for Human Brain Mapping*, 71–91.
- Vygotsky, L. S. (1934). *Thought and language*. MIT press (translation published in 2012).
- Wagner, A. D., Desmond, J. E., Demb, J. B., Glover, G. H., & Gabrieli, J. D. (1997). Semantic repetition priming for verbal and pictorial knowledge: A functional MRI study of left inferior prefrontal cortex. *Journal of Cognitive Neuroscience*, 9(6), 714–726. <https://doi.org/10.1162/jocn.1997.9.6.714>
- Wagner, L., & Lakusta, L. (2009). Using Language to Navigate the Infant Mind. *Perspectives on Psychological Science*, 4(2), 177–184. <https://doi.org/10.1111/j.1745-6924.2009.01117.x>
- Walbrin, J., Downing, P., & Koldewyn, K. (2018). Neural responses to visually observed social interactions. *Neuropsychologia*, 112, 31–39. <https://doi.org/10.1016/j.neuropsychologia.2018.02.023>
- Wang, J., Cherkassky, V. L., Yang, Y., Chang, K.-M. K., Vargas, R., Diana, N., & Just, M. A. (2016). Identifying thematic roles from neural representations measured by functional magnetic resonance imaging. *Cognitive Neuropsychology*, 33(3–4), 257–264. <https://doi.org/10.1080/02643294.2016.1182480>
- Wang, X., Gao, Z., Smallwood, J., & Jefferies, E. (2021). Both default and multiple-demand regions represent semantic goal information. *Journal of Neuroscience*. <https://doi.org/10.1523/JNEUROSCI.1782-20.2021>
- Wang, X., Wu, W., Ling, Z., Xu, Y., Fang, Y., Wang, X., Binder, J. R., Men, W., Gao, J.-H., & Bi, Y. (2018). Organizational Principles of Abstract Words in the Human Brain. *Cerebral Cortex (New York, N.Y.: 1991)*, 28(12), 4305–4318. <https://doi.org/10.1093/cercor/bhx283>
- Wasserman, E., Kiedinger, R. E., & Bhatt, R. (1988). *Conceptual behavior in pigeons: Categories, subcategories, and pseudocategories*. <https://doi.org/10.1037/0097-7403.14.3.235>
- Watson, J. B. (1920). Is Thinking Merely Action of Language Mechanisms? *British Journal of Psychology. General Section*, 11(1), 87–104. <https://doi.org/10.1111/j.2044-8295.1920.tb00010.x>
- Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences*, 13(6), 258–263. <https://doi.org/10.1016/j.tics.2009.03.006>
- Wehbe, L., Murphy, B., Talukdar, P., Fyshe, A., Ramdas, A., & Mitchell, T. (2014). Simultaneously Uncovering the Patterns of Brain Regions Involved in Different Story Reading Subprocesses. *PLOS ONE*, 9(11), e112575. <https://doi.org/10.1371/journal.pone.0112575>
- Wen, T., Duncan, J., & Mitchell, D. J. (2019). Representation of task episodes in human cortical networks. *BioRxiv*, 582858. <https://doi.org/10.1101/582858>
- Wen, T., Mitchell, D. J., & Duncan, J. (2020). The Functional Convergence and Heterogeneity of Social, Episodic, and Self-Referential Thought in the Default Mode Network. *Cerebral Cortex*, 30(11), 5915–5929. <https://doi.org/10.1093/cercor/bhaa166>
- Wernicke, C. (1874). *Der aphasische Symptomencomplex: Eine psychologische Studie auf anatomischer Basis*. Cohn.
- West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research*, 13(3), 363–375. [https://doi.org/10.1016/S0926-6410\(01\)00129-X](https://doi.org/10.1016/S0926-6410(01)00129-X)
- Whitehouse, P., Caramazza, A., & Zurif, E. (1978). Naming in aphasia: Interacting effects of form and function. *Brain and Language*, 6(1), 63–74. [https://doi.org/10.1016/0093-934x\(78\)90044-5](https://doi.org/10.1016/0093-934x(78)90044-5)
- Whitney, C., Jefferies, E., & Kircher, T. (2011). Heterogeneity of the left temporal lobe in semantic representation and control: Priming multiple versus single meanings of ambiguous words. *Cerebral Cortex (New York, N.Y.: 1991)*, 21(4), 831–844. <https://doi.org/10.1093/cercor/bhq148>
- Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2012). Executive semantic processing is underpinned by a large-scale neural network: Revealing the

- contribution of left prefrontal, posterior temporal, and parietal cortex to controlled retrieval and selection using TMS. *Journal of Cognitive Neuroscience*, 24(1), 133–147. [https://doi.org/10.1162/jocn\\_a\\_00123](https://doi.org/10.1162/jocn_a_00123)
- Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful Listening: The Processing of Degraded Speech Depends Critically on Attention. *Journal of Neuroscience*, 32(40), 14010–14021. <https://doi.org/10.1523/JNEUROSCI.1528-12.2012>
- Willems, R. M., Benn, Y., Hagoort, P., Toni, I., & Varley, R. A. (2011). Communicating without a functioning language system: Implications for the role of language in mentalizing. *Neuropsychologia*, 49(11), 3130–3135. <https://doi.org/10.1016/j.neuropsychologia.2011.07.023>
- Willems, R. M., der Haegen, L. V., Fisher, S. E., & Francks, C. (2014). On the other hand: Including left-handers in cognitive neuroscience and neurogenetics. *Nature Reviews Neuroscience*, 15(3), 193–201. <https://doi.org/10.1038/nrn3679>
- Williams, A., Reddigari, S., & Pylkkänen, L. (2017). Early sensitivity of left perisylvian cortex to relationality in nouns and verbs. *Neuropsychologia*, 100, 131–143. <https://doi.org/10.1016/j.neuropsychologia.2017.04.029>
- Willits, J. A., Amato, M. S., & MacDonald, M. C. (2015). Language knowledge and event knowledge in language use. *Cognitive Psychology*, 78, 1–27. <https://doi.org/10.1016/j.cogpsych.2015.02.002>
- Wilson, S. M., Molnar-Szakacs, I., & Iacoboni, M. (2008). Beyond Superior Temporal Cortex: Intersubject Correlations in Narrative Speech Comprehension. *Cerebral Cortex*, 18(1), 230–242. <https://doi.org/10.1093/cercor/bhm049>
- Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, 104(19), 7780–7785. <https://doi.org/10.1073/pnas.0701644104>
- Wing, J. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35. <https://doi.org/10.1145/1118178.1118215>
- Wing, J. (2011). Research notebook: Computational thinking—What and why. *The Link Magazine*, 6.
- Winograd, T. (1976). *Artificial intelligence and language comprehension*. US Department of Health, Education, and Welfare, National Institute of Education.
- Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus* (Trans. Pears and McGuinness). Routledge (originally published in 1922).
- Woolgar, A., Duncan, J., Manes, F., & Fedorenko, E. (2018a). Fluid intelligence is supported by the multiple-demand system not the language system. *Nature Human Behaviour*, 2(3), 200–204. <https://doi.org/10.1038/s41562-017-0282-3>
- Woolgar, A., Duncan, J., Manes, F., & Fedorenko, E. (2018b). *Fluid intelligence is supported by the multiple-demand system not the language system*. <https://doi.org/10.17863/CAM.22222>
- Woolgar, A., Parr, A., Cusack, R., Thompson, R., Nimmo-Smith, I., Torralva, T., Roca, M., Antoun, N., Manes, F., & Duncan, J. (2010). Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, 107(33), 14899–14902. <https://doi.org/10.1073/pnas.1007928107>
- Woolgar, A., Thompson, R., Bor, D., & Duncan, J. (2011). Multi-voxel coding of stimuli, rules, and responses in human frontoparietal cortex. *NeuroImage*, 56(2), 744–752. <https://doi.org/10.1016/j.neuroimage.2010.04.035>
- Wu, D. H., Waller, S., & Chatterjee, A. (2007). The functional neuroanatomy of thematic role and locative relational knowledge. *Journal of Cognitive Neuroscience*, 19(9), 1542–1555. <https://doi.org/10.1162/jocn.2007.19.9.1542>
- Wurm, M. F., & Caramazza, A. (2019). Distinct roles of temporal and frontoparietal cortex in representing actions across vision and language. *Nature Communications*, 10(1), 1–10. <https://doi.org/10.1038/s41467-018-08084-y>

- Xu, Y., He, Y., & Bi, Y. (2017). A Tri-network Model of Human Semantic Processing. *Frontiers in Psychology*, 8. <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01538>
- Xu, Y., Lin, Q., Han, Z., He, Y., & Bi, Y. (2016). Intrinsic functional network architecture of human semantic processing: Modules and hubs. *NeuroImage*, 132, 542–555. <https://doi.org/10.1016/j.neuroimage.2016.03.004>
- Xu, Y., Wang, X., Wang, X., Men, W., Gao, J.-H., & Bi, Y. (2018). Doctor, Teacher, and Stethoscope: Neural Representation of Different Types of Semantic Relations. *The Journal of Neuroscience*, 38(13), 3303–3317. <https://doi.org/10.1523/JNEUROSCI.2562-17.2018>
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8(8), 665–670. <https://doi.org/10.1038/nmeth.1635>
- Yeshurun, Y., Nguyen, M., & Hasson, U. (2017). Amplification of local changes along the timescale processing hierarchy. *Proceedings of the National Academy of Sciences*, 114(35), 9475–9480. <https://doi.org/10.1073/pnas.1701652114>
- Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C. J., & Hasson, U. (2017). Same Story, Different Story: The Neural Representation of Interpretive Frameworks. *Psychological Science*, 28(3), 307–319. <https://doi.org/10.1177/0956797616682029>
- Zaccarella, E., & Friederici, A. D. (2017). The neurobiological nature of syntactic hierarchies. *Neuroscience & Biobehavioral Reviews*, 81, 205–212. <https://doi.org/10.1016/j.neubiorev.2016.07.038>
- Zettersten, M., & Lupyan, G. (2020). Finding categories through words: More nameable features improve category learning. *Cognition*, 196, 104135. <https://doi.org/10.1016/j.cognition.2019.104135>
- Zhang, M., Nathaniel, U., Savill, N., Smallwood, J., & Jefferies, E. (2022). Intrinsic connectivity of left ventrolateral prefrontal cortex predicts individual differences in controlled semantic retrieval. *NeuroImage*, 246, 118760. <https://doi.org/10.1016/j.neuroimage.2021.118760>
- Zhu, Z., Bastiaansen, M., Hakun, J. G., Petersson, K. M., Wang, S., & Hagoort, P. (2019). Semantic unification modulates N400 and BOLD signal change in the brain: A simultaneous EEG-fMRI study. *Journal of Neurolinguistics*, 52, 100855. <https://doi.org/10.1016/j.jneuroling.2019.100855>
- Zimmerer, V. C., Varley, R. A., Deamer, F., & Hinzen, W. (2019). Factive and counterfactive interpretation of embedded clauses in aphasia and its relationship with lexical, syntactic and general cognitive capacities. *Journal of Neurolinguistics*, 49, 29–44. <https://doi.org/10.1016/j.jneuroling.2018.08.002>