



HHS Public Access

Author manuscript

Curr Dir Psychol Sci. Author manuscript; available in PMC 2022 March 15.

Published in final edited form as:

Curr Dir Psychol Sci. 2021 December ; 30(6): 526–534. doi:10.1177/09637214211046955.

Similarity of computations across domains does not imply shared implementation: The case of language comprehension

Evelina Fedorenko, Cory Shain

Brain & Cognitive Sciences Department and McGovern Institute for Brain Research, MIT

Abstract

Understanding language requires applying cognitive operations (e.g., memory retrieval, prediction, structure building)—relevant across many cognitive domains—to specialized knowledge structures (a particular language’s phonology, lexicon, and syntax). Are these computations carried out by domain-general circuits or by circuits that store domain-specific representations? Recent work has characterized the roles in language comprehension of the language-selective network and the multiple demand (MD) network, which has been implicated in executive functions and linked to fluid intelligence, making it a prime candidate for implementing computations that support information processing across domains. The language network responds robustly to diverse aspects of comprehension, but the MD network shows no sensitivity to linguistic variables. We therefore argue that the MD network does not play a core role in language comprehension, and that past claims to the contrary are likely due to methodological artifacts. Although future studies may discover some aspects of language that require the MD network, evidence to date suggests that those will not be related to core linguistic processes like lexical access or composition. The finding that the circuits that store linguistic knowledge carry out computations on those representations aligns with general arguments against the separation between memory and computation in the mind and brain.

Keywords

language; domain specificity; executive functions; working memory; cognitive control; prediction

Introduction

Incremental language comprehension likely relies on general cognitive operations like retrieval of representations from memory, predictive processing, attentional selection, and hierarchical structure building (e.g., Tanenhaus et al., 1995; Gibson, 2000). For example, in any sentence containing a non-local dependency between words, the first dependent has to be retrieved from memory when the second dependent is encountered. These kinds of operations are also invoked in other domains of perception and cognition, including object recognition, numerical and spatial reasoning, music perception, social cognition, and task planning (e.g., Dehaene et al., 2003; Botvinick, 2007). The apparent similarity of these kinds of mental operations across domains has led to arguments that the brain contains

domain-general circuits that carry out these operations, and that language draws on these circuits (e.g., Patel, 2003; Novick et al., 2005; Koechlin & Jubault, 2006; Fitch & Martins, 2014; Figure 1A).

Indeed, a network of frontal and parietal brain regions—the “multiple demand” (MD) system (also known as the executive/cognitive control network; Figure 2B)—has been shown to respond during diverse cognitive tasks and linked to constructs like working memory, inhibition, attention, prediction, structure building, and fluid intelligence (e.g., Duncan et al., 2020), making it a perfect candidate for carrying out hypothesized domain-general operations. However, many domains—including language—rely on domain-specific knowledge representations stored in specialized brain areas/networks. For example, language recruits a network of frontal and temporal brain regions that respond in a highly selective manner during language comprehension (Fedorenko et al., 2011; Figure 2A) and—when damaged in adulthood—lead to selectively linguistic deficits (e.g., Fedorenko & Varley, 2016).

During language comprehension, language-selective and MD brain regions may work together, with the MD system carrying out general operations on domain-specific knowledge representations, as hypothesized above. However, it is also possible that domain-specific networks locally implement general types of computations (e.g., retrieval of information from memory, predictive processing, and structure building), as has previously been argued for language comprehension (Martin et al., 1994; R. L. Lewis, 1996; Caplan & Waters, 1999). In other words, these kinds of computations may be involved across domains without drawing on shared circuits (e.g., Hasson et al., 2015; Dasgupta & Gershman, 2021; Figure 1B), possibly as a way of minimizing wiring lengths (Chklovskii & Koulakov, 2004). Under this view, the MD network may be a general ‘fallback’ system for domains or tasks for which the brain lacks specialized circuitry.

We review a recent body of work that investigated various aspects of human sentence comprehension using fMRI techniques that reliably distinguish the language-selective network from the domain-general MD network (see Fedorenko & Blank, 2020 for review), so their functional response properties can be probed independently. This approach is complementary to—but more direct than—dual-task paradigms and patient investigations that have been used in the past to probe the role of domain-general resources in language comprehension (e.g., Martin et al., 1994; R. L. Lewis, 1996; Caplan & Waters, 1999). Results consistently show (1) strong sensitivity in the language network, (2) little response in the MD network, and (3) significantly stronger responses in the language network than the MD network to every investigated component of natural language comprehension, including word predictability, working memory retrieval, and generalized measures of language comprehension difficulty. Together, these findings support the existence of a self-sufficient specialized language system that carries out the bulk of language-related processing demands.

1. The MD network does not closely ‘track’ the linguistic signal

Activity in a brain region or network that supports linguistic computations should be modulated by the properties of the linguistic stimulus. One method for estimating the degree of stimulus-linked activity (or stimulus ‘tracking’) was developed by Hasson and colleagues (e.g., Hasson et al., 2010) based on the correlations across individuals during the processing of naturalistic stimuli. The logic is as follows: if a brain region/network processes features of a stimulus, different individuals should show similar patterns of increases and decreases in neural response over time. Importantly, this method makes no assumptions about *what features* in the stimulus are important, providing a theory-neutral way to estimate the degree of stimulus-linked activity. Across three experiments, Blank & Fedorenko (2017) investigated synchrony across brains during naturalistic language processing, and found the (expected) strongly stimulus-linked responses in the language-selective network. Critically, however, the MD network exhibited substantially lower levels of stimulus-linked activity (Figure 2C). To rule out the possibility that the MD network tracks linguistic stimuli closely, but in a more variable way (than the language system) across individuals, Blank & Fedorenko examined within-participant correlations to multiple presentations of the same stimulus. Such correlations were similarly low, indicating that the MD network’s activity is less strongly modulated by changes in the linguistic signal than that of the language network.

2. The MD network does not show a core functional signature of language processing

Natural language sentences exhibit rich patterns of syntactic (e.g., Chomsky, 1957) and semantic (e.g., Montague, 1973) structure that are not present in perceptually matched stimuli, such as lists of unconnected words or non-words. Processing syntactic and semantic dependencies is widely thought to impose a computational burden (e.g., S. Lewis & Phillips, 2015), and thus an expected signature of language processing is an increased response to sentences over control conditions that lack structure. The language network robustly bears out this prediction (e.g., Fedorenko et al., 2011). In a recent large-scale study, Diachek, Blank, Siegelman et al., (2020) investigated whether the same is true of the MD network. Their sample consisted of fMRI responses from 481 participants each of whom completed one or more of 30 language comprehension experiments varying in linguistic materials. Some experiments included sentences, others – lists of unconnected words, and others both of these conditions. Results replicated past findings of stronger responses in the language network during the processing of sentences compared to word-lists, but showed systematically greater MD engagement during the processing of word-lists than during the processing of sentences, plausibly reflecting the greater difficulty of encoding unstructured stimuli (Figure 2D). This pattern is inconsistent with generalized MD involvement in sentence comprehension.

3. The MD network does not show effects of word predictability

Word predictability effects are robustly attested in behavioral (e.g., Ehrlich & Rayner, 1981) and electrophysiological (e.g., Kutas & Hillyard, 1984) measures of human language

processing, and prior work has argued that frontal and parietal cortical areas—likely within the MD network—encode expectancies across domains (e.g., Corbetta & Shulman, 2002), including language (Strijkers et al., 2019). Thus, one possible role for the MD network in language processing is to encode incremental prediction error. Shain, Blank, et al. (2020) investigated this possibility by analyzing measures of word-by-word *surprisal* (e.g., Levy, 2008) in fMRI responses to naturalistic audio stories.¹ They examined effects of surprisal estimates based on both word sequences (5-gram surprisal models that predict the next word based on the preceding four words) and syntactic structures (probabilistic context-free grammar models that predict the next word based on an incomplete syntactic analysis of the unfolding sentence). These effects were significant (and separable) in the language-selective network, but neither was significant in the MD network (Figure 2E, first two sets of bars). Thus, whereas the results support the existence of a rich predictive architecture that exploits both word co-occurrences and syntactic patterns, this architecture appears to be housed in language-specific cortical circuits, rather than relying on a domain-general predictive coding mechanism that may reside in the MD network.

4. The MD network does not show effects of syntactic integration

Influential theories of human sentence comprehension posit a critical role for working memory retrieval in integrating words into an incomplete parse of the unfolding sentence (e.g., Gibson, 2000). Given that working memory is thought to be one of the core functions supported by the MD network (Duncan et al., 2020), one plausible role for this network in language comprehension is as a working memory resource for syntactic structure building. Shain et al. (in prep.) investigated this possibility by exploring the contribution of multiple theory-derived measures of working memory cost to explaining variance in the language and MD networks' responses to naturalistic linguistic stimuli. The language network showed a systematic and generalizable (to an unseen data portion) response to variants of *integration cost* as proposed by Gibson's (2000) Dependency Locality Theory. Gibson posits that constructing syntactic dependencies incurs a retrieval cost proportional to the number of intervening elements that compete referentially with the retrieval target. This pattern did not hold in the MD network (Figure 2E, third set of bars), where activity did not reliably increase with measures of integration cost (or other types of working memory demand explored in the study). Thus, whereas these results support a role for working memory retrieval in naturalistic language processing, the working memory resources that support such computations reside in language-specific circuits, with little role for working memory resources housed in the MD network.

5. The MD network does not show effects of comprehension difficulty

The foregoing review challenges a role for the MD network in either of the two of the core classes of computation posited by current theorizing in human sentence processing research: prediction (e.g., Levy, 2008) and integration (e.g., Gibson, 2000). However, it is infeasible to enumerate and test the many other possible computations involved in human

¹Given a probability model p , surprisal I is the negative log probability of a word given its preceding context: $I(w_j) = -\log[p(w_j | w_0 \dots w_{j-1})]$

language processing—including those not covered by existing theory—in which MD may play a role. Wehbe et al. (2021) bypassed this limitation by leveraging independent measures of reading latency to predict fMRI responses to naturalistic stories. Reading latencies are widely regarded in psycholinguistics as reliable, theory-neutral proxies for language comprehension difficulty, and are commonly used as dependent variables to test hypotheses about the determinants of comprehension difficulty (Rayner, 1998). This design enabled Wehbe et al. (2021) to test whether comprehension difficulty in general registers in the MD network, without pre-commitment to a particular theory of sentence processing. The language, but not the MD, network showed a strong effect of comprehension difficulty (Figure 2F). Thus, the MD network is unlikely to play a critical role in the computations that govern incremental (word-by-word) language comprehension difficulty, regardless of how this difficulty is explained theoretically.

6. The MD network does not respond during comprehension in the absence of extraneous task demands

This lack of evidence for the MD network's engagement during language processing appears to contradict many prior reports of activity in what appear to be MD regions during language processing (e.g., Novick et al., 2005). Critically, such results are almost always obtained in paradigms where word/sentence comprehension is accompanied by extraneous tasks (e.g. multitasking), which may engage the MD network given its robust sensitivity to task demands (Duncan et al., 2020). Diachek, Blank, Siegelman et al. (2020) investigated this possibility by contrasting the MD network's engagement in language experiments that involved passive comprehension (visual or auditory) vs. those that involved an additional task, such as a memory probe task, semantic association judgments, or comprehension questions. Whereas the language network was equally engaged both in the presence and absence of a task, the MD network was only engaged in the presence of a task (Figure 2G). In other words, passive language comprehension is sufficient to engage language-selective regions, but not MD regions, suggesting that MD engagement during language comprehension is primarily induced by non-linguistic task demands (see Discussion).

Discussion

The evidence presented here challenges the hypothesis that domain-general executive resources support core computations of incremental language processing. The multiple demand (MD) network (Duncan et al., 2020), where such resources are likely housed, does not closely track linguistic stimuli, responds more robustly to less language-like materials (e.g., lists of unconnected words relative to sentences), does not show evidence of engagement in predictive linguistic processing, retrieval of previously encountered linguistic elements from working memory, or any other linguistic operation that leads to comprehension difficulty during language processing, and, unlike the core language network, is not engaged by passive comprehension, instead requiring a secondary task (e.g., memory probe or sentence judgments, Figure 2C–G). These findings greatly constrain the space of plausible language-related computations that the MD network might support and align with the architecture outlined in Figure 1B. In particular, it appears that the network

that stores linguistic knowledge representations is the also the network that performs all the relevant computations on these representations in the course of incremental comprehension, in spite of the fact that many of these computations may be similar to, or the same as, computations used in other domains.

Why have prior studies reached different conclusions about the reliance of language processing on domain-general resources? The answer is likely two-fold. *First*, for many years, researchers have not clearly differentiated between the language-selective and the domain-general circuits that co-habit the left frontal lobe but are robustly and unambiguously distinct (see Fedorenko & Blank, 2020 for discussion). This failure to separate the two networks is due to a combination of a) traditional group-averaging analyses, which blur nearby functionally distinct regions, especially in the association cortex where the precise locations of such regions differ across individuals (e.g., Frost & Goebel, 2012), and b) frequent reverse inference from coarse anatomical locations (i.e., concluding that a cognitive function was involved because anatomical brain areas previously associated with that function were active, e.g., Poldrack, 2006). And *second*, many prior studies have used paradigms where word/sentence comprehension is accompanied by a secondary task, and/or where highly artificially constructed or manipulated linguistic materials are used. Such paradigms may indeed recruit the MD network, which is robustly sensitive to task demands, but this recruitment does not speak to the role of this network in core linguistic operations like lexical access, or syntactic/semantic structure building. For these reasons, we have restricted our review to studies that a) relied on well-validated functional localizers (Fedorenko et al., 2011) to identify the language and the MD networks within each individual brain, and b) used naturalistic comprehension tasks (Hasson et al., 2018). Such studies converged on a clear answer: the domain-general MD network does not support core linguistic computations.

The fact that the *language* system appears to locally implement general computations like memory retrieval, prediction, and structure building suggests that local computation may systematically accompany functional specialization. This conjecture aligns with prior arguments for a tight integration between memory and computation at the neuronal level (Hasson et al., 2015; Dasgupta & Gershman, 2021). If a particular stimulus (be it faces, spatial layouts, high-pitched sounds, or linguistic input) is encountered with sufficient frequency to support specialization in particular circuits, then it may be advantageous for those circuits to carry out as much processing as possible in that domain, given that local computation may reduce processing latencies that would result from interactive communication with other systems (Chklovskii & Koulakov, 2004). Nevertheless, novel cognitive demands regularly arise, and it is infeasible to dedicate cortical ‘real estate’ to each of them. A general-purpose cognitive system like the MD network is therefore indispensable to robust and flexible cognition (Duncan et al., 2020), including the critical ability to solve novel problems. Indeed, recent computational modeling work has shown that an artificial neural network trained on multiple tasks will spontaneously develop functionally specialized sub-networks for different tasks; however, if new tasks are continually introduced, a subset of the network will remain flexible and not show a preference for any known task (Yang et al., 2019).

Although the studies summarized here rule out a large set of possibilities for the role of the MD network in language processing, more work is needed to evaluate the role of this network in language production, in more diverse linguistic phenomena (e.g., pragmatic inference, including during conversational exchanges), and in recovery from damage to the language network (e.g., Hartwigsen, 2018). Furthermore, the contributions of (possibly domain-general) subcortical and cerebellar circuits to language comprehension and cognitive processing require additional investigation; future work may show overlap between linguistic and non-linguistic functions in such circuits.

Conclusion

Despite the apparent similarity between the mental operations required for language comprehension and those required by other cognitive domains, the evidence we reviewed here challenges the hypothesis that domain-general executive circuits (housed within the multiple-demand, or MD, network) play a core role in language comprehension. We conjecture that the same holds for other domains that rely on domain-specific representations, and that the core contribution of the MD network to human cognition lies in supporting flexible behavior and the ability to solve new problems.

Acknowledgements:

We thank a) former and current EvLab and TedLab members, especially Idan Blank, for helpful comments and discussions over the last few years; b) Yev Diachek and Leila Wehbe for help with organizing the data for Figure 2; c) Hannah Small for creating the figures; and d) Matt Davis and John Duncan for comments on the earlier draft of the manuscript. The work summarized here and EF were supported by the R00 award HD057522, R01 awards DC016607 and DC016950, a grant from the Simons Foundation to the Simons Center for the Social Brain at MIT, and funds from the Brain and Cognitive Sciences department and the McGovern Institute for Brain Research. We apologize to researchers whose relevant papers we do not cite: this is due to the strict limit on the number of references allowed by the journal; we provide a more complete list of relevant references at <https://osf.io/dx4ah/>.

Bibliography²

- Blank I, & Fedorenko E (2017). Domain-general brain regions do not track linguistic input as closely as language-selective regions. *Journal of Neuroscience*, 3616–3642.
- Botvinick M (2007). Multilevel structure in behavior and in the brain: A computational model of Fuster's hierarchy. *Philosophical Transactions of the Royal Society, Series B: Biological Sciences*, 362, 1615–1626.
- Caplan D, & Waters GS (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22(1), 77–94.
- Chklovskii DB, & Koulakov AA (2004). Maps in the brain: what can we learn from them? *Annu. Rev. Neurosci*, 27, 369–392. [PubMed: 15217337]
- Chomsky N (1957). *Syntactic Structures*. Mouton.
- Corbetta M, & Shulman GL (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3), 201. [PubMed: 11994752]
- Dasgupta I, & Gershman SJ (2021). Memory as a Computational Resource. *Trends in Cognitive Sciences*.
- Dehaene S, Piazza M, Pinel P, & Cohen L (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, 20(3–6), 487–506. [PubMed: 20957581]

²For an expanded list of references, see <https://osf.io/dx4ah/>.

- Diachek E, Blank I, Siegelman M, & 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. 10.1523/JNEUROSCI.2036-19.2020 [PubMed: 32317387]
- Duncan J, Assem M, & Shashidhara S (2020). Integrated intelligence from distributed brain activity. *Trends in Cognitive Sciences*.
- Ehrlich SF, & Rayner K (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, 20(6), 641–655.
- Fedorenko E, Behr MK, & Kanwisher N (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*.
- Fedorenko E, & Blank I (2020). Broca's Area Is Not a Natural Kind. *Trends in Cognitive Sciences*.
- 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*, 201315235.
- Fedorenko E, & Varley R (2016). Language and thought are not the same thing: evidence from neuroimaging and neurological patients. *Annals of the New York Academy of Sciences*, 1369(1), 132. [PubMed: 27096882]
- Fitch WT, & Martins MD (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, 1316(1), 87–104. [PubMed: 24697242]
- Frost MA, & Goebel R (2012). Measuring structural–functional correspondence: spatial variability of specialised brain regions after macro-anatomical alignment. *Neuroimage*, 59(2), 1369–1381. [PubMed: 21875671]
- Gibson E (2000). The dependency locality theory: A distance-based theory of linguistic complexity. *Image, Language, Brain: Papers from the First Mind Articulation Project Symposium*, 95–126.
- Hartwigsen G (2018). Flexible redistribution in cognitive networks. *Trends in Cognitive Sciences*, 22(8), 687–698. [PubMed: 29914734]
- Hasson U, Chen J, & Honey CJ (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in Cognitive Sciences*, 19(6), 304–313. [PubMed: 25980649]
- Hasson U, Egidi G, Marelli M, & Willems RM (2018). Grounding the neurobiology of language in first principles: The necessity of non-language-centric explanations for language comprehension. *Cognition*, 180, 135–157. [PubMed: 30053570]
- Hasson U, Malach R, & Heeger DJ (2010). Reliability of cortical activity during natural stimulation. *Trends in Cognitive Sciences*, 14(1), 40–48. [PubMed: 20004608]
- Koechlin E, & Jubault T (2006). Broca's Area and the Hierarchical Organization of Human Behavior. *Neuron*, 50(6), 963–974. [PubMed: 16772176]
- Kutas M, & Hillyard SA (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161–163. [PubMed: 6690995]
- Levy R (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177. [PubMed: 17662975]
- Lewis RL (1996). Interference in short-term memory: The magical number two (or three) in sentence processing. *The Journal of Psycholinguistic Research*, 25, 93–115. [PubMed: 8789368]
- Lewis S, & Phillips C (2015). Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research*, 44(1), 27–46. [PubMed: 25408514]
- Martin RC, Shelton JR, & Yaffee LS (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language*, 33(1), 83–111.
- Montague R (1973). The proper treatment of quantification in ordinary {E}nglish. In Hintikka J, Moravcsik JME, & Suppes P (Eds.), *Approaches to Natural Language* (pp. 221–242). D. Riedel.
- Novick JM, Trueswell JC, & Thompson-Schill SL (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, 5(3), 263–281.
- Patel AD (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6(7), 674–681. [PubMed: 12830158]

- Poldrack RA (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10(2), 59–63. [PubMed: 16406760]
- Rayner K (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, 124(3), 372–422. [PubMed: 9849112]
- Shain C, Blank I, van Schijndel M, Schuler W, & Fedorenko E (2020). fMRI reveals language-specific predictive coding during naturalistic sentence comprehension. *Neuropsychologia*, 138. 10.1016/j.neuropsychologia.2019.107307
- Strijkers K, Chanoine V, Munding D, Dubarry A-S, Trébuchon A, Badier J-M, & Alario F-X (2019). Grammatical class modulates the (left) inferior frontal gyrus within 100 milliseconds when syntactic context is predictive. *Scientific Reports*, 9(1), 4830. [PubMed: 30886251]
- Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, & Sedivy JCE (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632–1634. [PubMed: 7777863]
- Wehbe L, Blank IA, Shain C, Futrell R, Levy R, von der Malsburg T, Smith N, Gibson E, & Fedorenko E (2020). Incremental language comprehension difficulty predicts activity in the language network but not the multiple demand network. *BioRxiv*.
- Yang GR, Joglekar MR, Song HF, Newsome WT, & Wang X-J (2019). Task representations in neural networks trained to perform many cognitive tasks. *Nature Neuroscience*, 22(2), 297–306. [PubMed: 30643294]

Recommended readings:

- 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. [PubMed: 20171926] A review of evidence for the existence of a broad, domain-general fronto-parietal multiple-demand brain system that supports executive functions and fluid intelligence.
- 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. A review of evidence for functional specialization in the human brain, with emphasis on the visual system and discussion of general implications for cognitive architecture and research methods.
- Fedorenko E, & Blank I (2020). Broca's Area Is Not a Natural Kind. *Trends in Cognitive Sciences*. A review arguing that Broca's area contains functionally distinct subregions that belong to the language-selective and multiple demand networks; these subregions have been conflated in much prior work leading to substantial confusion in the field.
- Dasgupta I, & Gershman SJ (2021). Memory as a Computational Resource. *Trends in Cognitive Sciences*. A review arguing that memory, in the form of *memoization* (storing computational outputs for future use), may be a ubiquitous component of neural information processing, rather than the domain of a designated resource.

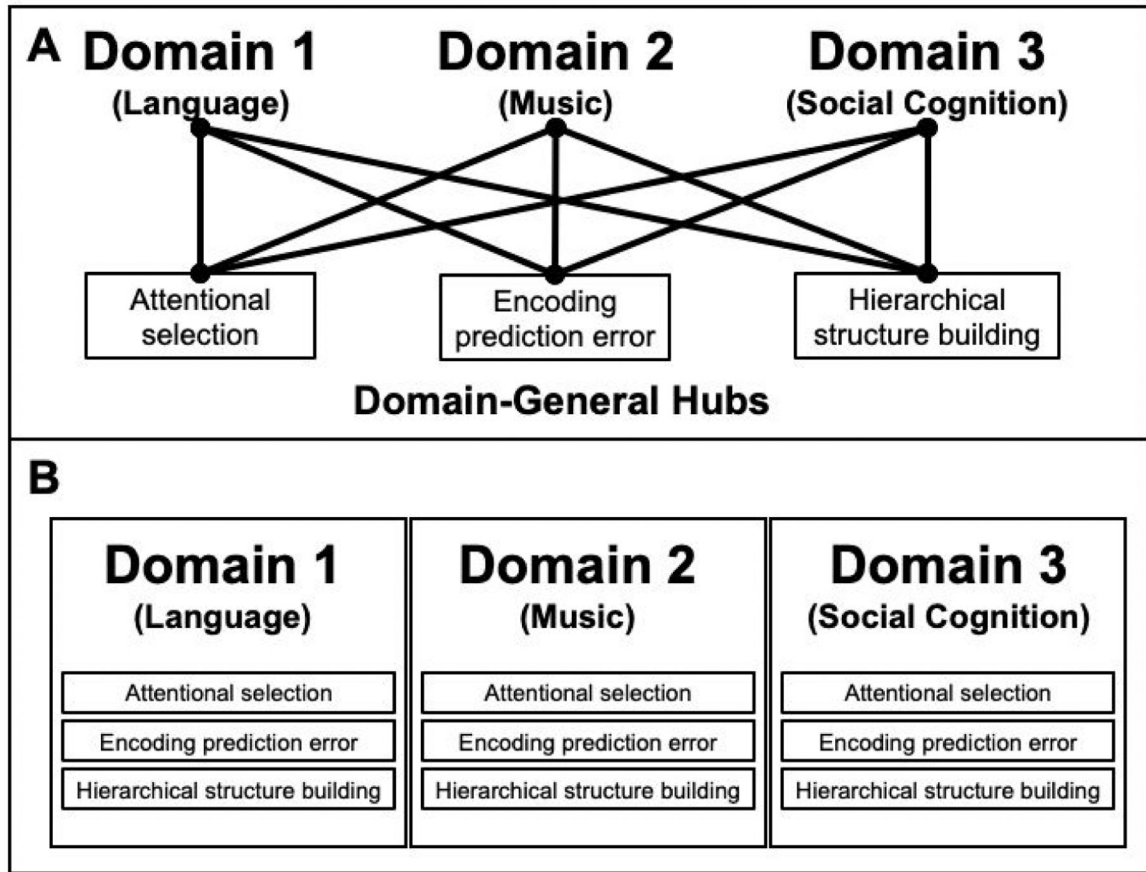


Figure 1. A schematic illustration of an architecture where computations that are used across domains are implemented in shared circuits (A) and an architecture where such general computations are implemented locally within each relevant set of domain-specific circuits (B). The architecture in A assumes separation between ‘memory’ circuits (which store domain-relevant knowledge representations) and ‘computation’ circuits; in the architecture in B, the circuits that store domain-specific knowledge representations also carry out computations on those representations (e.g., Hasson et al., 2015; Dasgupta & Gershman, 2021).

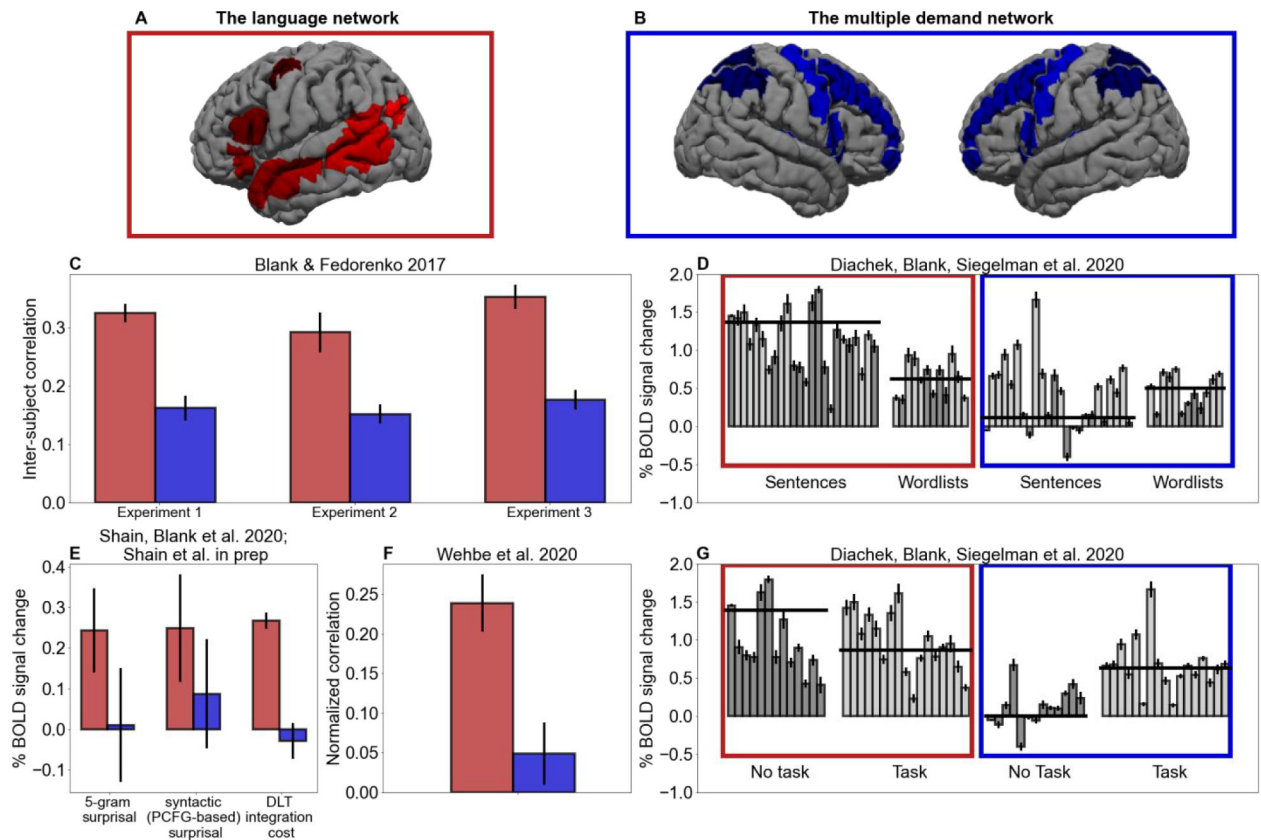


Figure 2.

A–B. The language (A) and the MD (B) networks shown as sets of ‘parcels’ derived from group-level probabilistic activation overlap maps for the *sentences>nonword-lists* contrast (Fedorenko et al., 2011) and a *hard>easy working memory task* contrast (Fedorenko et al., 2013), respectively. These parcels are used to constrain the definition of functional regions of interest in each individual participant in all the studies presented in C–G, as described in each relevant publication. **C–G.** A summary of recent fMRI findings showing sensitivity in the language, but not the MD, network to diverse aspects of language comprehension. For each study, the results are averaged across the regions within each network, but the patterns also hold for each region individually. **C.** The results from Blank & Fedorenko (2017) showing stronger inter-subject correlation over time during naturalistic story comprehension (suggesting stronger stimulus ‘tracking’) in the language vs. the MD network (error bars: standard errors of the mean (SEM) by participants). **D.** The results from Diacheck, Blank, Siegelman et al. (2020) showing a network by condition interaction: the language network shows a robust *sentences>word-lists* preference, whereas the MD network shows the opposite preference (here and in G, different bars correspond to different experiments, darker grey bars are passive reading/listening experiments, lighter bars are experiments where language processing is accompanied by a task, like a memory probe task, comprehension questions, or sentence judgments; for each experiment, error bars=SEM by participants; horizontal lines correspond to averages across participants for the relevant set of experiments). **E.** The results from Shain, Blank et al. (2020; first two sets of bars) and Shain et al. (in prep.; third set of bars) showing that the language, but

not the MD, network is sensitive to n -gram and syntactic surprisal, and integration cost (as operationalized in the Dependency Locality Theory, DLT; Gibson, 2000) during naturalistic story comprehension (error bars=SEM by functional ROIs; by-participant variation in effect size was not modeled due to evidence of poor generalization). **F.** The results from Wehbe et al. (2021) showing that the language, but not the MD, network is robustly sensitive to generalized comprehension difficulty, as measured with reading times in self-paced reading and eye-tracking during reading in independent groups of participants (error bars=SEM by participants). **G.** The results from Diachek, Blank, Siegelman et al. (2020) showing that the language network responds robustly during language comprehension regardless of the presence of an extraneous task, but the MD network responds in the presence of an extraneous task but not during passive reading/listening.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript