REVIEW

The neurobiology of language beyond single-word processing

Peter Hagoort*

In this Review, I propose a multiple-network view for the neurobiological basis of distinctly human language skills. A much more complex picture of interacting brain areas emerges than in the classical neurobiological model of language. This is because using language is more than single-word processing, and much goes on beyond the information given in the acoustic or orthographic tokens that enter primary sensory cortices. This requires the involvement of multiple networks with functionally nonoverlapping contributions.

The capacity for language is a central feature of the human condition. It allows us to communicate with our fellow citizens, to accumulate knowledge, to create cultural practices, and to support our thought processes. Language is a complex biocultural hybrid. To understand its intricate organization and neurobiological underpinnings, we have to decompose language skills into the basic building blocks and core operations. Basic building blocks include the knowledge that has been acquired during development about the sound patterns of the one or more languages a speaker commands, the meaning of its lexical items, their syntactic features (such as noun, verb, and grammatical gender), the orthographic patterns (in reading), or the signs in the languages of the deaf. Next to these elementary linguistic units (ELUs), there are elementary linguistic operations (ELOs) that enable the retrieval of ELUs from memory (for example, in word recognition), or the generation of larger structures from these elementary building blocks [as in morphological (de)composition in compounding or verb inflection and as in the construction of sentence level meaning]. In addition, the proposition created by combining ELUs and ELOs has to be linked to the actual or imagined situation in which it is embedded in order to establish its truth value (1, 2). For example, if I say "The editor of the journal loved the paper," I have produced an impeccable sentence, but one can only know whether what I said is true or false if the nouns "editor," "journal," and "paper" can be linked to specific tokens. As we will see below, all this is only part of the complex story of language.

Although evolutionary precursors of human language can be identified (3), in its full capacity it is distinctly human. Despite its complexity, most children acquire the core of this capacity in the first few years of life, without formal instruction and well before they have learned to lace their shoes or perform simple arithmetic operations. This suggests that the infrastructure of the human brain provides the child with a certain language-readiness. One functional feature that has been argued to enable human singularity in this regard is our ability to infer tree structures from sequential data (4–6). Tree structures, as a representational format introduced by Wilhelm Wundt (7), are intricately linked to the notion of hierarchy. This is exemplified in, among other things, the morphological make-up of words and the hierarchical interpretation of phrases (Fig. 1). This propensity to compute hierarchical structures is not limited to language but
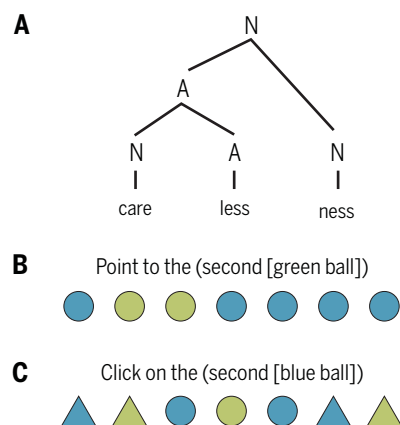

Fig. 1. Tree-like structures that are characteristic for word formation and for the interpretation of phrases and sentences. (**A**) The word "carelessness" is made up of the morphemes "care" (N, noun), "less" (A, marker for adjective), and "ness" (N, marker for noun). The final morpheme determines its status as a noun. (**B**) The phrase "the second green ball" is interpreted by 30 out of 31 participants as referring to the third circle and not the second. This is the result of a hierarchical interpretation of the phrase as referring to the second of the green balls. The representation in brackets is formally equivalent to a tree structure. Participants saw the instructions without brackets. (**C**) In the case of an array with balls and triangles, the phrase "the second blue ball" is again interpreted as the second of the blue balls (the fifth shape) instead of as the second blue item, which is a ball (the third shape) by about 99% of the participants.

generalizes to other domains of cognition, such as planning and music.

The classical view and its shortcomings

For a long time, insights into the neurobiological basis of language were derived from the views of neurologists in the late 19th and early 20th century [such as Broca, Wernicke, and Lichtheim (8)], as interpreted and extended in (9). According to this classic model, the human language faculty was situated in the left perisylvian cortex, with a strict division of labor between the frontal and temporal regions. Wernicke's area in the left temporal cortex was assumed to subserve the comprehension of speech, whereas Broca's area in the left inferior frontal cortex (LIFC) was claimed to support language production. The arcuate fasciculus connected these two areas. Although still influential, this model proved to have severe limitations and to be largely wrong (10). These are the key issues: (i) Broca's and Wernicke's area are ill-defined and do not form natural neuroanatomical elements. Moreover, they are further parcellated into multiple areas with different cytoarchitectonic profiles and receptorarchitectonic fingerprints (11–13). (ii) Functional magnetic resonance imaging (fMRI) and lesion studies have shown that the language-relevant cortex is much more extended than assumed, including large parts of the temporal cortex, part of the parietal cortex, and areas in the LIFC beyond Brodmann areas 44 and 45. Moreover, language is less strictly left-lateralized than once thought (Fig. 2). (iii) Both frontal and temporal regions are involved in language comprehension as well as in language production (14, 15). (iv) The connectivity of the language-relevant cortex is much more extended than the classical model assumed and certainly not restricted to the arcuate fasciculus (16). And (v) the cerebellum and subcortical structures such as the thalamus and basal ganglia play an important role as well (17), especially for the fine-tuning of timing and sequencing in speaking. It is relevant to realize that the classic view was mainly based on single-word processing. The idea that language has combinatorial machinery beyond single words was lacking. This might in part explain why the classic model substantially underrepresents the brain regions and fiber tracts that are important for language.

The ability to combine words in often new ways is a hallmark of human language. This combinatoriality is realized by dynamic interactions between Broca's area and the adjacent cortex in the LIFC with areas in the temporal and parietal cortex. The interplay between these areas guarantees that lexical information retrieved from memory is unified into coherent multiword sequences with an overarching syntactic structure and semantic interpretation (18).

Max Planck Institute for Psycholinguistics, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands.
*Corresponding author. Email: peter.hagoort@donders.ru.nl

Neuroplasticity

An additional insight is that language-relevant cortex is somewhat variable. Recent studies have shown that under certain conditions, cortex outside the classical perisylvian areas can be involved. Language processing occurs in the occipital cortex of congenitally blind individuals (19, 20). Although the exact computational contributions still need to be determined, there is evidence that recruitment of the visual cortex is related to behavioral performance in a verbal memory task (21). Moreover, the visual cortex of congenitally blind individuals responds to sentences with syntactic movement but not to the difficulty of math equations (22). This result supports the hypothesis that this area can contribute to the computation of sentence-level syntactic structure. Overall, these findings suggest that the cyto-architectonic constraints for specifications of cognitive function leaves certain degrees of freedom. If input patterns to a particular brain area change, this area might be recruited for quite different functionality. On the basis of these and similar findings, Bedny concludes that "human cortices are cognitively pluripotent, that is, capable of assuming a wide range of cognitive functions. Specialization is driven by input during development, which is constrained by connectivity and experience" (23). The mapping relation between elementary building blocks and elementary operations for language on the one hand and the neural architecture of the human brain on the other is still far from fully understood. The degree of individual variation in the neurobiological infrastructure for language is also still largely unchartered territory (24).

The immediacy principle

A much clearer picture has emerged for key features of the processing dynamics of language, partly as a consequence of the high temporal resolution of neurophysiological measurements. Language processing occurs at an amazing speed. We easily produce and understand two to five words per second. The consequence of the need for speed is that in both production and comprehension, the retrieval of the linguistic building blocks and the combinatorial operations happen incrementally. Furthermore, language processing is characterized by the "immediacy principle" (25). In comprehension, linguistic and extra-linguistic information is used immediately upon becoming available. That is, knowledge about the context and the world, concomitant information from other modalities (such as co-speech gestures) (26, 27), and knowledge about the speaker (28) are brought to bear immediately on the same fast-acting brain system that combines the meanings of individual words. In other words, all available relevant information will be used without delay to codetermine the interpretation of the speaker's message. In all cases, the LIFC in dynamic interaction with temporoparietal areas plays an important role in unifying the different sources of information that determine the interpretation of an utterance (29).

Although there is no firm evidence that prediction is necessary for comprehension, it has become clear that prediction contributes to language processing and might be relevant to meet the demand for speed (30–32). Very likely, lexical, semantic, and syntactic cues conspire to predict characteristics of the next anticipated word, including its syntactic and semantic make-up. A mismatch between contextual prediction and the output of bottom-up analysis results in an immediate brain response that recruits additional processing resources for the sake of salvaging the on-line interpretation process (18).
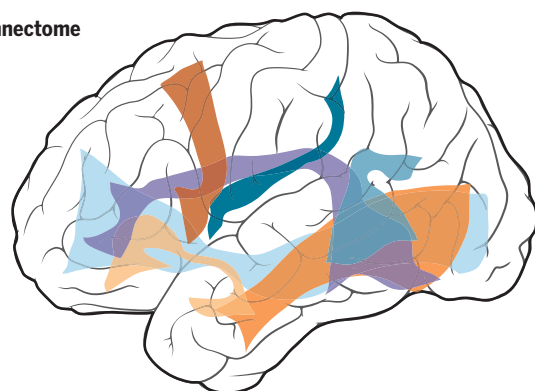
Information structure

The speed of language processing is also helped by the fact that utterances are most often part of a longer interaction between speaker and listener. In such an interaction scheme, some information is already shared between the conversational partners. This is often referred to as "topic." The new information is what is added and, hence, the key component of the message. This is usually referred to as the "comment." Information structure refers to the way in which the topic and comment are packaged in the organization of a sentence (1). Speakers often take care to mark the focus constituent that contains the new information. The way this is done differs between languages. There is no linguistic universal for signaling information structure. In some languages, syntactic locations are used for marking the focus constituent; other languages use focus-marking particles or prosodic features such as phrasing and accentuation. For example, in English question-answer pairs, the new or relevant information in the answer will typically be pitch accented. After a question such as "What did Mary buy at the market?" the answer might be "Mary bought VEGETABLES" (accented word in capitals). In this case, "vegetables" is the focus constituent, providing the new information.

Information structure was found to modulate language-relevant activations in fMRI (33, 34) and in electrophysiological readouts (35–37). For example, unexpected semantic and illegal syntactic continuations resulted in sizeable N400 and P600 effects, respectively, if they were part of the focus constituent (35, 36). These effects were, however, strongly reduced or even absent if the continuations were part of the nonfocus constituent (Fig. 3). New information marked by pitch accent activated the domain-general attention network (33). To gain-modulate the depth of language processing, linguistic devices, such as pitch accent, that mark the relevant information in an expression (such as the comment) seem to trigger the attention network into operation.

A The language-relevant connectome

- Arcuate fasc. (long)
- Arcuate fasc. (ant)
- Arcuate fasc. (post)
- Inf. frontal-occipital fasc.
- frontal aslant tract
- inf. logitudinal fasc.
- uncinate fasc.
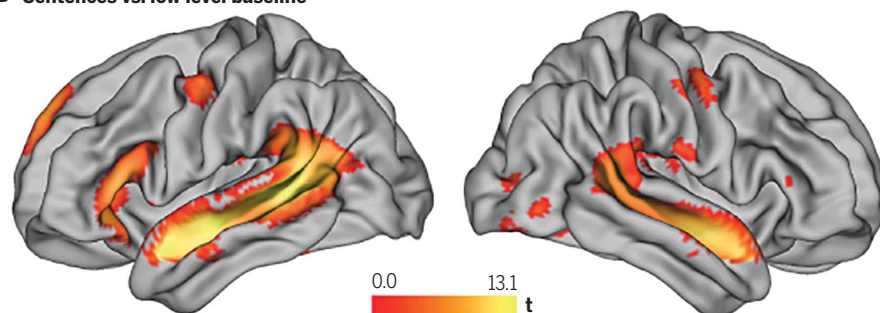
B Sentences vs. low level baseline

0.0 13.1 t

Fig. 2. Components of the neural infrastructure for language. (**A**) Fiber tracts connecting language-relevant brain regions (16). (**B**) Common activations for listening to sentences and sentence reading in an fMRI study with 204 participants, compared with a low-level baseline (59).

F3

This account ties in with the idea that for listeners, a full reconstruction of the linguistic input is in many cases not achieved, but for most purposes, a superficial and incomplete analysis is good enough (*38*). According to this "good-enough" processing account, some semantic and grammatical details might be ignored. More recently, Ferreira and Lowder (*39*) have made the connection between prediction, information structure, and good-enough language processing. They propose that topic information is processed in a good-enough manner, whereas the listener's prediction effort is devoted to the comment (the new or focused information) because "all that is needed for comprehension and interpretation is an identification of the comment—the topic being already available" (*1*).

Pragmatic inferencing

Although areas in the perisylvian cortex, especially in the left hemisphere, are important for encoding and decoding propositional content, this is not the whole story. The meaning of an utterance is often strongly dependent on contextual information that is not actually coded in what is said (*40*–*42*). What is coded in words and linguistic constructions underdetermines what is meant or intended. Extracting the intended meaning requires that inferences are made on the basis of assumptions about the beliefs and intentions of the interacting agents and about a common understanding of appropriate language use. For example, in the right context the utterance "It is hot here" will not only be interpreted as a statement about the state of affairs but first and foremost as an implicit request to do something about it (such as open the window or turn down the heating). The socially binding and communicative roles of language depend critically on the capacity to make the right pragmatic inferences (*1*, *43*).

So far, relatively few studies have investigated the neural infrastructure for effective communication beyond the core linguistic machinery for word retrieval, sentential syntax, and semantics. Nevertheless, the picture that emerges is fairly consistent (*44*–*49*). Pragmatic inferences depend on the contribution of core areas in the theory of mind (ToM) network. These include the right temporoparietal junction and medial prefrontal cortex, areas typically involved in tasks that require mental state reasoning—that is, thinking about other people's beliefs, emotions, and desires (*50*–*52*).

An example is a study in which the authors presented participants with sentences in the presence of a picture (*48*). In one condition, the sentence in combination with the picture
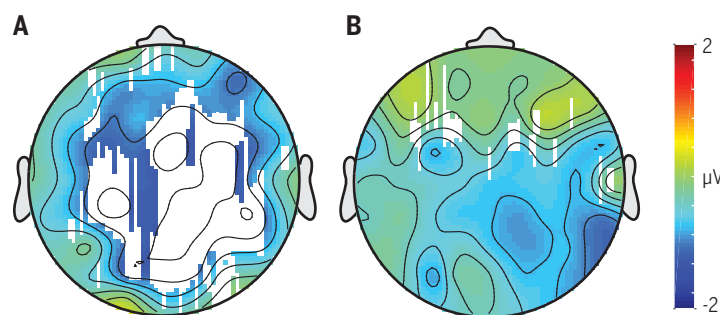

Fig. 3. Topographic distribution of N400 effects triggered by a semantically unexpected compared with an expected word. (**A**) In the focus constituent. (**B**) The same words in a nonfocus constituent (*35*).

could be interpreted as an indirect request for action. One of their items combined the utterance "It is hot here" with a picture of a door. Participants interpreted this as a request to open the door. However, the same utterance combined with the picture of a desert was interpreted as a statement. Sentences in the indirect-request condition activated the ToM network much more strongly than the same sentences without the possibility to interpret them as an indirect request. The recognition of a speech act induced by an utterance in combination with its context seems to require the contribution of the mentalizing machinery instantiated in the ToM network. Further insights will depend on a more precise account than is currently available of the computational contributions of the different nodes in the ToM network (*53*).

The multiple network view

One approach to the neurobiology of language is to start from identifying the essence of language. Once the few core features of language that make it stand out in the animal kingdom have been established, the question can be addressed regarding which aspects of human brain organization enable the essence of language. If, for example, recursion is key (*54*), one might identify Broca's area as the neural equivalent of the push-down stack subserving this distinctly human computational capacity (*5*, *55*).

The account that I have sketched here takes a different, nonessentialistic stance. It is based on the conviction that accounting for the full picture of human language skills is not helped by a distinction between essential and nonessential aspects of speech and language. In a first step, we need to decompose complex language skills such as speaking and listening into ELUs and ELOs as the key building block for encoding and decoding propositional content. These are mainly supported by the LIFC and substantial parts of the temporal cortex and parietal cortex, with a left hemisphere bias. Interaction with the attentional control–multiple demand system is needed to up-regulate the processing of the utterance

constituent that is marked as comment or new information. In addition, integrating the utterance content into a situation model that spans multiple connected utterances is required. This process appears to involve the right inferior frontal gyrus and the right angular gyrus (*56*). Last, in order to extract the intended message from the coded meaning provided by the linguistic utterance, the listener has to integrate linguistic and nonlinguistic information. She also has to draw the required pragmatic inferences, which requires a contribution of areas crucial for mentalizing.

Next to core areas for retrieving lexical information from memory and the unification of the lexical building blocks in producing and understanding multiword utterances, other brain networks are needed to realize language-driven interactions to their full extent. Multiple brain network contributions will not be language-specific but shared with other cognitive functions. ELUs will almost certainly be domain-specific (*57*), but some key aspects of the ELOs (such as unification) might well be shared with other domains, such as music and arithmetic. A much more complex picture of interacting brain areas emerges than in the classic model. The reason is that understanding language is more than single-word processing, and much goes on beyond the information given in the acoustic or orthographic tokens that enter primary sensory cortices. This multiple-brain-network view does not take away the need for further detailed specifications of the computational contributions of the different brain areas involved in language processing.

REFERENCES AND NOTES

1. P. A. M. Seuren, *Language in Cognition: Language From Within* (Oxford Univ. Press, 2009), vol. 1.
2. P. A. M. Seuren, *Saussure and Sechehaye: Myth and Genius. A Study in the History of Linguistics and the Foundations of Language* (Brill, 2018).
3. S. E. Fisher, *Curr. Biol.* **29**, R65–R67 (2019).
4. A. E. Martin, *Front. Psychol.* **7**, 120 (2016).
5. W. T. Fitch, *Phys. Life Rev.* **11**, 329–364 (2014).
6. S. Dehaene, F. Meyniel, C. Wacongne, L. Wang, C. Pallier, *Neuron* **88**, 2–19 (2015).
7. W. Wundt, *Logik* (Enke, 1880), vol. II.
8. W. J. M. Levelt, *A History of Psycholinguistics: The Pre-Chomskyan Era.* (Oxford Univ. Press, 2013).
9. N. Geschwind, *Science* **170**, 940–944 (1970).
10. P. Tremblay, A. S. Dick, *Brain Lang.* **162**, 60–71 (2016).
11. K. Amunts, K. Zilles, *Trends Cogn. Sci.* **16**, 418–426 (2012).
12. K. Amunts, K. Zilles, *Neuron* **88**, 1086–1107 (2015).
13. K. Zilles, K. Amunts, *Curr. Opin. Behav. Sci.* **21**, 93–105 (2018).
14. L. Menenti, S. M. E. Gierhan, K. Segaert, P. Hagoort, *Psychol. Sci.* **22**, 1173–1182 (2011).
15. K. Segaert, L. Menenti, K. Weber, K. M. Petersson, P. Hagoort, *Cereb. Cortex* **22**, 1662–1670 (2012).
16. K. Amunts, M. Catani, in *The Cognitive Neurosciences*, M. S. Gazzaniga, G. R. Mangun, Eds. (MIT Press, 2015), pp. 619–628.
17. H. Barbas, M. A. García-Cabezas, B. Zikopoulos, *Brain Lang.* **126**, 49–61 (2013).
18. P. Hagoort, *Neurosci. Biobehav. Rev.* **81** (Pt B), 194–204 (2017).
19. M. Bedny, A. Pascual-Leone, D. Dodell-Feder, E. Fedorenko, R. Saxe, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 4429–4434 (2011).

20. M. Bedny, M. MacSweeney, in *Human Language: From Genes and Brains to Behavior*, P. Hagoort, Ed. (MIT Press, 2019).
21. A. Amedi, N. Raz, P. Pianka, R. Malach, E. Zohary, *Nat. Neurosci.* **6**, 758–766 (2003).
22. C. Lane, S. Kanjlia, A. Omaki, M. Bedny, *J. Neurosci.* **35**, 12859–12868 (2015).
23. M. Bedny, *Trends Cogn. Sci.* **21**, 637–648 (2017).
24. K. Mahowald, E. Fedorenko, *Neuroimage* **139**, 74–93 (2016).
25. P. Hagoort, J. van Berkum, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **362**, 801–811 (2007).
26. L. Drijvers, A. Özyürek, *Brain Lang.* **177-178**, 7–17 (2018).
27. A. Ozyürek, R. M. Willems, S. Kita, P. Hagoort, *J. Cogn. Neurosci.* **19**, 605–616 (2007).
28. J. J. A. Van Berkum, D. van den Brink, C. M. Tesink, M. Kos, P. Hagoort, *J. Cogn. Neurosci.* **20**, 580–591 (2008).
29. P. Hagoort, *Curr. Opin. Neurobiol.* **28**, 136–141 (2014).
30. T. Brothers, T. Y. Swaab, M. J. Traxler, *Cognition* **136**, 135–149 (2015).
31. G. R. Kuperberg, T. F. Jaeger, *Lang. Cogn. Neurosci.* **31**, 32–59 (2016).
32. J. Rommers, A. S. Meyer, P. Praamstra, F. Huettig, *Neuropsychologia* **51**, 437–447 (2013).
33. L. B. Kristensen, L. Wang, K. M. Petersson, P. Hagoort, *Cereb. Cortex* **23**, 1836–1848 (2013).
34. T. M. van Leeuwen *et al.*, *Neuropsychologia* **58**, 64–74 (2014).
35. L. Wang, M. Bastiaansen, Y. Yang, P. Hagoort, *Neuropsychologia* **49**, 813–820 (2011).
36. L. Wang, M. Bastiaansen, Y. Yang, P. Hagoort, *PLOS ONE* **7**, e47917 (2012).
37. D. La Rocca, V. Masia, E. Maiorana, E. Lombardi Vallauri, P. Campisi, *Neurocomputing* **199**, 1–15 (2016).
38. F. Ferreira, N. D. Patson, *Lang. Linguist. Compass* **1**, 71–83 (2007).
39. F. Ferreira, M. W. Lowder, in *Psychology of Learning and Motivation, Vol 65,* B. H. Ross, Ed. (Elsevier Academic Press Inc, San Diego, 2016), vol. 65, pp. 217-247.
40. S. C. Levinson, *Pragmatics* (Cambridge Univ. Press, 1983).
41. D. Sperber, D. Wilson, *Behav. Brain Sci.* **10**, 697–710 (1987).
42. P. Hagoort, S. C. Levinson, in *The Cognitive Neurosciences*, M. S. Gazzaniga, G. R. Mangun, Eds. (MIT Press, 2015).
43. During a lecture, the famous physicist Paul Dirac was told by someone in the audience, "I don't understand that formula." Dirac continued his lecture without addressing the issue. The chairman then interrupted and asked Dirac whether he could reply to the question. Dirac was astonished, replying: "Question? What question? My colleague has made an assertion." As on other occasions, Dirac failed to grasp the intended message behind the assertion (*58*).
44. J. Bašnáková, K. Weber, K. M. Petersson, J. van Berkum, P. Hagoort, *Cereb. Cortex* **24**, 2572–2578 (2014).
45. J. Bašnáková, J. van Berkum, K. Weber, P. Hagoort, *Neuropsychologia* **76**, 79–91 (2015).
46. A. M. Paunov, I. A. Blank, E. Fedorenko, *J. Neurophysiol.* **121**, 1244–1265 (2019).
47. N. Spotorno, E. Koun, J. Prado, J. B. Van Der Henst, I. A. Noveck, *Neuroimage* **63**, 25–39 (2012).
48. M. J. van Ackeren, D. Casasanto, H. Bekkering, P. Hagoort, S. A. Rueschemeyer, *J. Cogn. Neurosci.* **24**, 2237–2247 (2012).
49. F. Vanlangendonck, R. M. Willems, P. Hagoort, *PLOS ONE* **13**, e0202943 (2018).
50. D. M. Amodio, C. D. Frith, *Nat. Rev. Neurosci.* **7**, 268–277 (2006).
51. J. Koster-Hale *et al.*, *Neuroimage* **161**, 9–18 (2017).
52. R. Saxe, in *Handbook of Theory of Mind*, A. Leslie, T. German, Eds. (Psychology Press, Philadelphia, PA, 2010).
53. E. Gibson, L. Bergen, S. T. Piantadosi, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8051–8056 (2013).
54. M. D. Hauser, N. Chomsky, W. T. Fitch, *Science* **298**, 1569–1579 (2002).
55. A. D. Friederici, *Language in Our Brain. The Origins of a Uniquely Human Capacity.* (MIT Press, 2017).
56. L. Menenti, K. M. Petersson, R. Scheeringa, P. Hagoort, *J. Cogn. Neurosci.* **21**, 2358–2368 (2009).
57. E. Fedorenko, M. K. Behr, N. Kanwisher, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 16428–16433 (2011).
58. C. Rovelli, *Reality Is Not What It Seems: The Journey to Quantum Gravity* (Penguin Books, 2017).
59. J. M. Schoffelen *et al.*, *Sci. Data* **6**, 17 (2019).