

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/262284156>

Neurolinguistics: Mind, brain, and language

Chapter · May 2014

CITATIONS

2

READS

17,977

1 author:



[David Kemmerer](#)

Purdue University

63 PUBLICATIONS 2,660 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Evolution of the language-ready brain [View project](#)

NEUROLINGUISTICS: MIND, BRAIN, AND LANGUAGE

David Kemmerer

*Department of Speech, Language, and Hearing Sciences; Department of Psychological Sciences
Purdue University, West Lafayette, IN 47906, USA
kemmerer@purdue.edu*

To appear as Chapter 19 in the following book:

Allan, K. (Ed.) (2014). *The Routledge Handbook of Linguistics*. New York: Routledge.

1. Introduction

The goal of neurolinguistics is to understand how the cognitive capacity for language is subserved by the biological tissue of the brain. This interdisciplinary field of research began in the mid 19th century, and by the early 20th century several basic aspects of the neural architecture of language had already been discovered, such as left-hemisphere dominance and the strong reliance of speech perception and production on certain regions in the temporal and frontal lobes. These initial insights came primarily from investigations of brain-damaged patients who displayed fairly consistent correlations between, on the one hand, particular patterns of impaired and preserved linguistic abilities, and on the other hand, particular patterns of lesioned and intact brain structures. But even though that first wave of neuropsychological exploration was of great historical importance, the value of the studies was limited by their lack of precision on both sides of the language-brain relationship. From the mid 20th century up to the present, however, the whole field evolved quite dramatically for a variety of reasons, including the maturation of modern linguistics, the cognitive revolution in psychology, the emergence of computer science and artificial intelligence, and the invention and progressive refinement of numerous brain mapping methods. Due to these developments, neurolinguistics is now a vibrant, rapidly growing field in which researchers with different backgrounds frequently collaborate to conduct empirical and theoretical studies on diverse topics.

The following synopsis begins by summarizing several kinds of cortical organization as well as the most commonly used brain mapping methods. It then provides a selective review of recent hypotheses and findings about the neural representation and processing of spoken word forms, printed word forms, word meanings, and sentences. Although some of the most well-supported contemporary perspectives are covered, space limitations only allow a small amount of material to be included, far less than is actually available in the relevant literature. Further information can be found in the references listed under “Recommended reading.”

2. Cortical organization

Although subcortical structures such as the thalamus, basal ganglia, and cerebellum certainly contribute to language, the most complex computations are carried out in the cerebral cortex, which is the convoluted outer mantle of the brain. The cortex contains approximately 30 billion neurons, and because each of them makes contact with at least 1,000 other cells, the whole system constitutes a massively interactive information processing matrix that is adaptively designed to support high-level mental functions.

For present purposes, the following types of cortical organization are noteworthy. First, each hemisphere is divided into five lobes, four of which are visible—namely, the frontal, parietal, temporal, and occipital lobes—and one which is hidden—namely, the insula, which faces laterally but lies at the bottom of the Sylvian fissure. Second, the cortex is essentially a sheet of gray matter that has been crumpled up so that a large amount of surface area can fit inside the cranium. The raised bulges are gyri, and the deep grooves are sulci ([Figure 1](#)). Third, crosscutting the gyral-sulcal topography is a mosaic of cytoarchitecturally distinct regions that resemble irregularly shaped patches on a quilt but actually reflect the presence/absence, packing density, and layering of various types of cells. Several parcellation schemes have been proposed, but the most widely used one involves Brodmann areas (BAs; [Figure 1](#)). For example, Broca's area, which is a major computational hub for language, occupies BAs 44 and 45, both of which reside in the posterior portion of the left inferior frontal gyrus. Finally, it is essential to realize that cortical regions do not operate in isolation; instead, they are intricately interconnected with each other, and most if not all of our complex cognitive processes, including those involving language, require the dynamic, cooperative interplay of signals among the widely distributed components of large-scale cortical circuits. These signals are transmitted along white matter pathways that run beneath the cortex like highways between cities. The fiber tracts that are most relevant to the left-hemisphere language system are still being elucidated, but at least seven major long-distance routes are currently under investigation.

3. Brain mapping methods

During the past few decades, neurolinguistics has benefited tremendously from advances in brain mapping methods. Most of the findings described in the sections below come from experimental studies that employed one or more of the following techniques.

There are two main hemodynamic (a.k.a. functional neuroimaging) approaches—positron emission tomography (PET) and functional magnetic resonance imaging (fMRI)—both of which use blood flow as a proxy for neural activity to identify which brain regions tend to be activated when people perform particular types of mental tasks. The primary strength of these methods is relatively fine-grained spatial resolution—about 10 mm for PET, and about 3 mm for fMRI (but sometimes as good as 0.5 mm). The primary weakness is relatively poor temporal resolution—30 seconds or so for PET, and often around 12 seconds for fMRI (but sometimes as good as 50 ms). Another limitation of hemodynamic approaches is that the associations they reveal between neural and mental processes are correlational rather than causal in nature; hence they cannot indicate which of the many brain regions that show up as being engaged during a task are in fact necessary for that task.

Traditionally, the chief method for inferring causal brain-behavior relationships has been the lesion method (a.k.a. neuropsychology), which, as mentioned above, was the source of the first insights about the neural underpinnings of language back in the mid 19th century. Basically, this approach takes advantage of otherwise unfortunate cases of brain damage (due to strokes, tumors, head injuries, and neurodegenerative diseases) by using them to make new discoveries about the design of the normal system. In the domain of language, neuropsychological studies usually pursue one or both of two general aims: first, to carve the language faculty at its joints,

so to speak, by determining which of its components can be selectively impaired; and second, to identify reliable links between specific linguistic deficits and specific lesion sites. Importantly, the tools for conducting such studies have undergone significant improvements in recent years.

Another technique for inferring causal brain-behavior relationships is transcranial magnetic stimulation (TMS). This non-invasive method involves inducing electrical currents in specific brain regions by means of brief magnetic pulses. The currents can either facilitate or disrupt the operation of the targeted area, depending on the parameters of the protocol. Moreover, the spatial resolution is in the order of millimeters, since each pulse is quite focal, and the temporal resolution is in the order of milliseconds, since each pulse is quite brief. Despite these merits, however, the method does have shortcomings: first, it is restricted to brain regions near the scalp; and second, the effects of stimulation can spread to remote regions, due to the massive interconnectivity of cortical areas.

Electrophysiological techniques can bring researchers even closer to the actual firing of neurons. In fact, one approach, which is only used in neurosurgical situations, involves directly stimulating parts of the exposed brain and observing the effects on cognition and behavior. Another approach involves recording the electrical signals of neurons as they unfold on a millisecond timescale during mental processes. This is done in either of two ways: intracranially by means of electrodes that are placed directly in the brain so as to record the firing of either single cells or relatively small populations of cells; or extracranially by means of electrodes that are placed on the scalp so as to record, through the skull, the simultaneous firing of many thousands or even millions of cells.

Other brain mapping methods include magnetoencephalography, near-infrared spectroscopy, transcranial direct current stimulation, and pharmacological manipulations. But they are not used as frequently as the ones described earlier.

The rest of this chapter reviews a number of functional-anatomical correspondences involving various aspects of the neural substrates of language ([Figure 2](#)). All of these correspondences have received empirical support from studies employing diverse brain mapping methods. At the same time, however, all of them remain tentative, some more so than others.

4. Spoken word forms

4.1. Speech perception

Speech perception is a deceptively simple cognitive capacity. Someone speaks, the sounds enter our ears, and we understand immediately. But in order for such seemingly effortless comprehension to occur, numerous computations must be carried out. Analog acoustic patterns must be converted to digital codes at multiple levels of language-specific structure, including distinctive features, phonemes, syllables, and words. Although the categorization of speech signals must be sensitive to fine-grained cues, it must also be flexible enough to accommodate talker variability. The boundaries between words must be identified even though there are rarely corresponding gaps in the acoustic waveform. And all of these operations,

together with many others, must be executed extremely quickly in order for comprehension to unfold at a normal pace.

Furthermore, speech input must be routed not only to the grammatical and semantic systems that analyze the forms and meanings of utterances, but also to the motor system that subserves articulation. This is mainly because we rely on auditory-motor transformations when we learn how to say new words that we hear, especially during the early phase of language acquisition. Such transformations also contribute, however, to the overt repetition of familiar words, and they are involved in covert auditory-verbal short-term memory as well, like when you silently rehearse a piece of important information, such as a phone number. In addition, abundant data indicate that the motor system contributes to ordinary, passive speech perception by constantly “resonating” to the speaker’s articulatory movements. As described below, however, the specific functional significance of this phenomenon is controversial.

During speech perception, acoustic signals are initially encoded in the cochlea, and they pass through three brainstem nuclei as well as the thalamus before finally reaching the cortex. Interestingly, although the auditory brainstem was once believed to function in a hardwired fashion, recent research has shown that it can be modified by linguistic experience. In particular, compared to speakers of non-tone languages (e.g., English), speakers of tone languages (e.g., Thai) exhibit enhanced processing of pitch contours in the brainstem.

At the cortical level, the early stages of speech perception involve spectrotemporal analysis—that is, the determination of how certain sound frequencies change over time. These computations operate not only on speech, but also on other kinds of environmental sounds, and they take place in several regions of the superior temporal cortex, particularly the primary auditory cortex (which occupies Heschl’s gyrus deep within the Sylvian fissure) and several adjacent auditory fields on the dorsal surface of the superior temporal gyrus (STG).

The outputs of these areas then flow into other portions of both the posterior STG and the posterior superior temporal sulcus (STS) that collectively implement a phonological network. Processing along this pathway is mostly hierarchical and integrative, since lower levels of neuronal populations close to the primary auditory cortex represent relatively simple aspects of speech sounds, whereas higher levels of neuronal populations extending across the lateral surface of the STG and into the STS detect increasingly complex featural patterns and sequential combinations of speech sounds, such as specific consonants and vowels, specific phoneme clusters, and specific word forms. The precise architecture of the phonological network is far from straightforward, however. For instance, the identification of a particular vowel, irrespective of talker, has been linked not with a single discrete neuronal population, but rather with several cortical patches distributed across the posterior STG/STS.

Although the left hemisphere is dominant for speech perception, the right hemisphere also contributes. In fact, either hemisphere by itself can match a spoken word like *bear* with a picture of a bear, instead of with a picture corresponding to a phonological distractor (e.g., a pear), a semantic distractor (e.g., a moose), or an unrelated distractor (e.g., grapes). The two hemispheres do, however, appear to support speech perception in somewhat different ways. According to one proposal, the left posterior STG/STS is better equipped than the right to handle

rapid auditory variation in the range of around 20-80 ms, which is ideal for registering fine-grained distinctions at the phonemic level, such as the contrast in voice-onset time between /k/ and /g/, or the contrast in linear order between *pets* and *pest*. Conversely, the right hemisphere is more sensitive than the left to longer-duration auditory patterns in the range of around 150-300 ms, which is optimal for extracting information at the syllabic level, like metrical stress.

After the sound structure of a perceived word has been recognized in the phonological network of the posterior STG/STS, there is a bifurcation of processing into two separate streams, one ventral and the other dorsal. The ventral stream has the function of mapping sound onto meaning. It does this by projecting first to the posterior middle temporal gyrus (MTG), and then to the anterior temporal lobe (ATL). Both of these regions contribute to semantic as well as morphosyntactic processing in ways that are elaborated further below. Although the ventral stream appears to be bilateral, it is more robust in the left than the right hemisphere.

The dorsal stream has the function of mapping sound onto action. It does this by projecting first to a region at the posterior tip of the Sylvian fissure that is sometimes referred to as area Spt (for Sylvian parietal-temporal), and then to a set of articulatory structures in the inferior frontal gyrus (IFG), precentral gyrus (PreG), and anterior insula. Area Spt serves as an interface for translating between the sound-based phonological network in the temporal lobe and the motor-based articulatory network in the frontal lobe. The dorsal stream is left-hemisphere dominant, and it supports auditory-verbal short-term memory by continually cycling spoken word forms back and forth between the posterior phonological network and the anterior articulatory network, thereby allowing them to be kept “in mind,” which is to say, in an activated state. The dorsal stream is also involved in basic speech perception, since the frontal motor programs for producing certain words are automatically engaged whenever those words are heard, and recognition can either be enhanced or reduced by using transcranial magnetic stimulation to modulate the operation of the relevant frontal regions. These modulatory effects are fairly small, however, and there is an ongoing debate over the degree to which “motor resonance” actually facilitates speech perception.

4.2. Speech production

The ability to produce spoken words is no less remarkable than the ability to perceive them. In ordinary conversational settings, English speakers generate about two to three words per second, which is roughly equivalent to three to six syllables consisting of ten to twelve phonemes. These words are retrieved from a mental lexicon that contains, for the average literate adult, between 50,000 and 100,000 entries, and articulating them requires the precise coordination of up to 100 muscles. Yet errors are only rarely made, occurring just once or twice every 1,000 words.

The first step in word production is to map the idea one wishes to express onto the meaning of a lexical item. Although the multifarious semantic features of individual words are widely distributed across the brain, there is growing evidence that the ATL plays an essential role in binding together and systematizing those features. This topic is discussed more fully in the section on word meanings, however, so in the current context it is sufficient to make the following points. To the extent that the ATL does subserve the integrated concepts that words

convey, it can be regarded (at least for the expository purposes required here) as not only near the endpoint of the ventral stream for speech perception, but also near the starting point of the pathway for speech production. In addition, it is noteworthy that many aspects of semantic processing in the ATL, such as the selection of certain lexical items over others, are regulated in part by cognitive control mechanisms in the IFG.

Once the meaning of a target word has been selected in the ATL, processing moves posteriorly along the lateral extent of the temporal cortex. There are some hints that the word's morphosyntactic features are accessed in the mid/posterior MTG, but other perisylvian cortical areas may also be involved. In contrast, a great deal of data suggests that the word's phonological form (defined here as just the sequence of segments that constitute its phonemic content) is accessed in the posterior STG/STS. In fact, many studies support the view that this region implements a sound-based phonological network that is recruited not only during speech perception, as described above, but also during speech production. However, the question of whether we operate with a single phonological lexicon, or with separate but anatomically adjacent ones for input and output processing, has been highly contentious throughout the history of neurolinguistics, and answering it once and for all will require new insights from future research.

From the phonological network in the temporal lobe, processing moves anteriorly through the dorsal stream to the articulatory network in the frontal lobe. Syllabification is generally thought to occur in the posterior IFG. During this relatively late process, the ordered phonemic segments of the target word are incrementally bundled into syllabic units that do not necessarily conform to morphemic units. For instance, the word *horses* is bimorphemic and bisyllabic, yet the final segment of the first morpheme is not treated as the final segment of the first syllable, but is instead treated as the initial segment of the second syllable: {[*hor*]-[*ses*]}. Some investigators have argued that the most frequently used syllables in a language gradually become stored in long-term motor memory as precompiled articulatory gestures, so that they can be efficiently activated as ready-made “chunks” rather than laboriously assembled again and again. No one would disagree, however, with the claim that complex articulatory orchestration is often required, and the available data suggest that this type of programming is handled not only by the posterior IFG, but also by the anterior insula—more specifically, by its superior sector, which is adjacent to the inferior sector of the posterior IFG.

After the appropriate high-level articulatory representations have been engaged in the posterior IFG and anterior insula, the corresponding sets of low-level motor commands are selected in the ventral PreG. The neuronal populations in this brain region are organized bilaterally and in a somatotopic manner that captures the layout of the various parts of the vocal tract—larynx, lips, jaw, tongue, and palate. Their main function is to “steer” the relevant muscles in a precisely coordinated, dynamic fashion during speech production. Their output signals, however, are relayed through several subcortical nuclei in the brainstem and spinal cord before finally reaching the appropriate parts of the motor periphery.

As with other kinds of bodily action, speech production is not entirely a feedforward process, but rather relies heavily on feedback mechanisms. When the goal is to utter a particular word, the sound-based representation of that word in the phonological network of the temporal

lobe serves as an "auditory target" that specifies what is expected to be heard. As articulation proceeds, the incoming acoustic signals of the resultant self-produced speech are immediately compared with that target representation, and if any discrepancies are detected, instructions for making the necessary corrections are sent to the frontal articulatory network. Recent research has shown that this feedback loop operates with remarkable speed and precision, allowing motor commands for speech to be adjusted "on the fly," often beneath the surface of awareness. Moreover, a parallel feedback loop in the somatosensory modality has been receiving increasing attention during the past few years. It recruits the supramarginal gyrus (SMG) to compare the predicted tactile and proprioceptive signals from the vocal tract with the actual signals. If errors are found, corrective messages are relayed to the frontal articulatory network, indicating how the motor program should be modified to generate the expected feelings in the vocal tract.

5. Printed word forms

5.1. Reading

For skilled readers, recognizing printed words seems easy. But the apparent simplicity of this process is an illusion. In reality, the reader's eyes make four or five saccades (i.e., jerky movements) every second, and with each fixation the amount of detail that is perceived drops off precipitously from the fovea (i.e., the small circular space of maximal visual acuity), so that it is only possible to register a few letters at a time. Furthermore, determining that a particular string of letters constitutes a familiar word is a formidable computational task. The essence of the problem is this: in order to recognize a written word, it is necessary to extract precisely those features that invariantly characterize that word across all of its possible manifestations, including changes in ^{position}, size, CASE, and *font*. To accomplish this feat, large differences in visual form must be ignored (e.g., between "a" and "A"), small ones must be noticed (e.g., between "e" and "c"), and alternative linear orders must be detected (e.g., between "dog" and "god"). Skilled readers, however, can effortlessly and accurately satisfy all these requirements within a time window of just a few hundred milliseconds.

The sight of a printed word triggers a cascade of transformations that extends from the retina to the thalamus, from there to the primary visual cortex at the back of the brain, and from there through a series of anteriorly directed ventral occipitotemporal way-stations that extract increasingly rich and informative combinations of orthographic features. From a representational perspective, this visual processing hierarchy starts with mere points and lines, but it leads progressively to case- and font-specific letter shapes, case- and font-invariant graphemes (i.e., abstract letter identities), short sequences of graphemes, and entire words. Many of the early stages of this hierarchy are bilateral, but there is growing evidence that the left hemisphere begins to dominate fairly quickly.

The hierarchy culminates in the Visual Word Form Area (VWFA), which is a cortical patch in the fusiform gyrus that has the following properties: it detects the identities of printed words regardless of their position, size, case, or font, and regardless of whether they are perceived consciously or unconsciously; it is more sensitive to real than unreal words; it is engaged equally by different types of familiar scripts (e.g., English, Arabic, Chinese, etc.), but it responds more strongly to familiar than unfamiliar scripts; and perhaps most important of all, it

prefers printed words to other kinds of visual objects. Now, because writing systems were not invented until very late in human history (about 5,400 years ago), the VWFA could not be innately designed for reading. It has been argued, however, that the reason why this particular region becomes relatively specialized for recognizing printed words when we learn to read is because it is inherently well-suited to handling complex combinations of spatially fine-grained shapes. Consistent with this view, and supporting the "meta-modal" nature of the VWFA, is the recent finding that the VWFA is the most significantly activated area not only when sighted people discriminate between real and unreal printed words, but also when congenitally blind people discriminate between real and unreal Braille words.

Once the form of a printed word has been recognized in the VWFA, how does it get mapped onto the associated phonological and semantic structures? These processes are enabled by multiple pathways—some sublexical, others lexical—but their precise neural underpinnings remain unclear. Still, some generalizations can be made. Access to the proper pronunciations of printed words seems to depend mainly on the perisylvian circuit for speech processing, whereas access to the concepts encoded by printed words seems to depend mainly on a more inferior set of structures that includes the ATL as well as several other temporal, parietal, and frontal areas. It is clear that printed words with regular spelling patterns, like the real word *desk* or the unreal word *blicket*, can be read aloud by mapping the graphemes directly onto the corresponding phonemes in rule-governed ways that bypass semantics. Some researchers have argued, however, that printed words with irregular spelling patterns, like *yacht*, can only be read aloud by first accessing their meaning, especially if they have low frequency. This is a controversial claim, though.

A final observation that leads naturally to the next topic is that, just as the auditory perception of spoken words automatically activates the oral motor programs for uttering them, so the visual perception of printed words automatically activates the manual motor programs for writing them. But while this is certainly an intriguing discovery, researchers have yet to determine how much such "motor resonance" enhances the efficiency of reading.

5.2. Writing

In neurolinguistics, writing has not received nearly as much attention as reading. Nevertheless, progress is being made in understanding how our brains allow us to produce printed words.

One of the earliest stages of writing involves retrieving the abstract spelling patterns of the intended lexical items (i.e., the appropriate grapheme strings, unspecified for size, case, and style). Several neuropsychological and functional neuroimaging studies suggest that these high-level representations are accessed in the VWFA. Needless to say, this is a very important finding, since it supports the hypothesis that the VWFA contains a single orthographic lexicon that is enlisted for both reading and writing.

After the abstract spelling pattern of a target word has been selected in the VWFA, it is kept "in mind" by the graphemic buffer. This is basically a short-term memory system that temporarily maintains in an activated state the identities and positions of the graphemes while the

word is being written. Whereas the graphemes themselves are most likely represented in the VWFA, the device that keeps them “alive” in a top-down, controlled manner appears to be implemented by the posterior IFG (i.e., Broca's area).

Finally, two low-level stages of written word production have been posited. The first is called allographic conversion, and it translates the abstract graphemes that are held in the graphemic buffer into concrete forms (e.g., upper or lower case, separate or cursively connected letters, etc.). The second is called graphomotor planning, and it provides even more precise instructions to the motor system for the hand, such as specifications for the size, direction, and sequence of strokes. It is widely believed that both of these processes are subserved mainly by hand-related dorsolateral frontoparietal regions.

Incidentally, when writing is performed with a keyboard instead of a pen or pencil, a distinct computational component devoted to graphomotor planning for the purpose of typing may take information directly from the graphemic buffer and use it to assemble a set of commands for consecutive button presses. So far, however, the operations that underlie typing have not been investigated as much as those that underlie handwriting.

6. Word meanings

How are the meanings of words represented and processed in the brain? Some people might be tempted to suppose that the cortical implementation of lexical knowledge includes, for every word, a nice, neat, neurally discrete dictionary definition that spells out all the relevant semantic information in an abstract symbolic code that might be called “mentalese.” Recent research suggests, however, that the real story is not only much more complicated than that, but also much more interesting.

There is mounting evidence that conceptual knowledge is, to some extent, grounded in modality-specific systems for perception and action, such that many forms of semantic processing involve unconscious simulations of fairly high-level sensory and motor states. Consider, for example, the meaning of the object noun *hammer*. Numerous studies suggest that this concept does not reside in any single place in the brain; instead, different fragments of it are scattered across different cortical regions according to the sensory or motor content of the type of information that is represented. Thus, visual-semantic specifications of how hammers look (i.e., the relevant shape patterns) appear to be stored in some of the same ventral temporal areas that are engaged when hammers are visually recognized; auditory-semantic specifications of how hammers sound (i.e., the relevant banging patterns) appear to be stored in some of the same superior/middle temporal areas that are engaged when hammers are auditorily recognized; motor-semantic specifications of how hammers are used (i.e., the relevant swinging patterns) appear to be stored in some of the same frontoparietal areas that are engaged when hammers are grasped and manipulated in customary ways; and so on.

It is important to note, however, that this field of inquiry is still quite young, and most of the key issues are hotly debated. Some of the questions that are currently being explored are as follows. To what degree do lexically based simulations really recruit neuronal populations that also contribute to perception and action? How much can such simulations be modulated by

contextual and strategic factors? And what is the relative weighting, or functional significance, of the multiple modality-specific components of lexical concepts?

As mentioned earlier, the meanings of words seem to depend not only on widely distributed modality-specific cortical regions, but also on the ATL. One influential hypothesis maintains that the ATL is a computational hub that plays a number of vital roles in semantic cognition. For one thing, it ensures that the multi-modal features of lexical concepts are properly integrated in long-term memory so that inferences across modalities can easily be made. For example, the word *duck* denotes a kind of bird with particular visual and auditory properties, and several studies suggest that certain sectors of the ATL capture these correlations, thereby compensating for the fact that the sight of ducks is not always accompanied by the sound of their quacking, and vice versa. In addition, there is growing evidence that the ATL allows conceptual processing to be driven by deep aspects of semantic structure, as opposed to being overly sensitive to superficial modality-specific similarities and differences. For example, the word *cat* most readily brings to mind small, furry, purring pets, like calicos; however, the relevant concept also includes atypical instances, like hairless cats, which certainly qualify as members of the cat category despite their lack of fur; and it excludes what might be called pseudotypical instances, like chihuahuas, which resemble cats but are actually dogs. Several studies suggest that the ATL is essential for making these kinds of judgments about which entities do and don't fall within the boundaries of certain concepts. Finally, a substantial body of data now supports the view that the ATL is crucially involved in integrating and organizing the meanings of not just object nouns, but also action verbs and various classes of abstract words.

In addition to modality-specific input/output systems and the ATL, many other cortical areas have recently been found to contribute to semantic cognition. Some of them are as follows. The anterior IFG interacts closely with the ATL and has been implicated in the resolution of conflicts between competing word meanings. The angular gyrus (AG) seems to have integrative hub-like semantic functions, possibly analogous to those of the ATL. And the posterior MTG appears to play an important role in the representation and/or processing of many types of lexical concepts, although the details have yet to be deciphered.

7. Sentences

7.1. Comprehension

The ability to understand complete sentences is underpinned by a large-scale, mostly left-lateralized neural circuit. This circuit consists of several widely distributed but tightly interconnected cortical areas that operate synergistically to transform incoming strings of words into compositionally unified messages. It is not yet known exactly how each component of the circuit contributes to the overall goal of comprehending complex multi-word expressions, but the rough outlines of the architecture are slowly beginning to emerge. In order to anchor our discussion of this intricate topic in some concrete examples, we will make frequent reference to the sentences in (1) and (2). They are instructive because they use partially different configurations of the very same words to describe partially different scenarios.

- (1) The reporter who attacked the senator admitted the error.
- (2) The reporter who the senator attacked admitted the error.

For both (1) and (2), as each word is encountered, not only must its form and meaning be rapidly retrieved, but its morphosyntactic features must be accessed as well. For example, *the* is a definite article, *reporter* and *senator* are both count nouns, *who* is a relative pronoun, *attacked* and *admitted* are both transitive verbs, and so on. Although the neural mechanisms that access these morphosyntactic features have not been precisely localized, most of the available data point to the posterior MTG. As mentioned above in the section on speech perception, during the receptive processing of spoken language, this region receives input directly from the phonological network in the posterior STG/STS. And more importantly, several studies suggest that the same region is crucially involved in identifying the grammatical categories of perceived words. It is also noteworthy that the posterior MTG operates in concert with the posterior IFG (i.e., Broca's area), especially when ambiguous expressions are encountered, like the phrase *flying planes*, in which *flying* could function as either a verb or an adjective. The basic idea is that in such situations the competing grammatical category assignments are represented in the posterior MTG, and the selection of the contextually appropriate one is executed in a top-down fashion by the posterior IFG.

We also observed in the section on speech perception that the posterior MTG projects forward to the ATL, and there is increasing evidence that the superior and middle sectors of this territory contribute to sentence comprehension in the following ways. Some of the neuronal populations here seem to be involved mainly in parsing, taking as input the morphosyntactic features and sequential orders of the incoming words, and yielding as output hierarchically organized phrases and clauses. In (1), for example, the subject of the sentence is the complex noun-phrase *the reporter who attacked the senator*, which consists of several smaller, nested constituents. Other neuronal populations in the superior/middle ATL appear to be devoted more to interpretation, specifically the compositional semantic process of incrementally building up the unified meanings of phrases and clauses. And still others have been implicated in both types of operations—that is, parsing as well as interpretation. In short, according to some lines of current thinking, the superior/middle ATL houses a combinatorial network that assembles and integrates progressively larger arrangements of grammatical and conceptual structures during online receptive sentence processing.

A crucial part of understanding a sentence is figuring out "who's doing what to whom," or, to put it somewhat more technically, determining the roles that the different participants play in the described situation. In both (1) and (2), for example, the structure of the main clause is the same, with the grammatical cues indicating that the reporter, not the senator, admitted the error. The two sentences vary, however, with regard to the structure of the relative clause, such that in (1) the grammatical cues indicate that the reporter is the agent of the attacking event and the senator is the patient, whereas in (2) the grammatical cues indicate that these roles are reversed. During the past few decades, a great deal research has focused on elucidating the neural mechanisms that carry out these sorts of role assignments during sentence comprehension. But even though many valuable insights have been made, the precise nature of the underlying machinery remains elusive. Still, there is increasing evidence that one of the key cortical areas is

the temporoparietal junction, which extends from the most posterior portion of the STS into the adjacent AG.

When sentences are heard rather than read, it is sometimes useful to keep the phonological forms of the words in an activated state until the comprehension process has been completed. This is accomplished by the auditory-verbal short-term memory (STM) system, which, as mentioned in the section on speech perception, has two components. First, the storage component represents activated word forms and is implemented by the phonological network in the posterior STG/STS; and second, the rehearsal component continually refreshes the contents of the storage component and is implemented by the articulatory network in the frontal lobe. This STM system is often called the phonological loop, and it is frequently employed to facilitate sentence comprehension. For instance, many studies have shown that, compared to sentences like (1), sentences like (2) are harder to understand and more likely to engage the phonological loop. These findings suggest that listeners tend to "re-play" sentences like (2) in their "mind's ear," thereby giving themselves more time to figure out exactly "who's doing what to whom." It is important to note, however, that the precise role of the phonological loop in sentence comprehension is controversial, largely because several neuropsychological studies have shown that a severe reduction of auditory-verbal STM capacity does not always disrupt the ability to understand long and convoluted sentences.

Finally, the most mysterious node in the circuit for sentence comprehension is the posterior IFG (i.e., Broca's area). Some researchers have proposed that it is involved in computing certain aspects of the hierarchical and sequential structures of sentences, like long-distance dependencies. In contrast, others have proposed that it aids comprehension by carrying out certain kinds of executive/supervisory operations, like using the rehearsal component of auditory-verbal STM to "re-play" unusually hard-to-understand sentences (as described above), or guiding the process of checking and, if necessary, revising the analysis of "who's doing what to whom." Although all of these hypotheses have some merits, none of them can account for the full range of data. For example, one of the most serious challenges for future work will be to explain the inconsistencies in the neuropsychological literature, since there is substantial evidence that damage to Broca's area sometimes does and sometimes doesn't lead to debilitating sentence comprehension deficits.

7.2. Production

The neural substrates of sentence production have not been explored in nearly as much depth as those of sentence comprehension. To some degree, this imbalance reflects the fact that, with regard to experimental design, it is much more challenging to control the relevant variables when people generate complex multi-word expressions than when they are given such expressions. Despite these limitations, however, progress is gradually being made in mapping the cortical architecture of sentence production.

During the 1980s and 1990s, most of the research on this topic focused on a disorder called agrammatism. Extensive cross-linguistic investigation showed that brain-damaged patients with this impairment tend to have five main problems with sentence production: a paucity of main verbs; syntactic simplification; omission of free-standing closed-class elements

(e.g., prepositions); substitution of bound closed-class elements (e.g., tense/aspect suffixes); and over-reliance on canonical word order. However, close scrutiny of individual patients revealed that all of these symptoms could dissociate from each other, leading some researchers to argue that agrammatism is not really a coherent disorder. Moreover, lesion analyses demonstrated that although patients classified as agrammatic usually have damage to left perisylvian frontal, parietal, and temporal regions, the specific sites vary tremendously. Taken together, these findings left it rather unclear how particular aspects of sentence production might be linked with particular brain areas.

During the past 15 years or so, a bit more has been learned about these issues, due in large part to several PET and fMRI studies that used sophisticated protocols to explore the neural correlates of sentence production in normal subjects. Interestingly, a few of these studies directly compared production tasks with comprehension tasks involving the same sentences and found that two main regions—the posterior MTG and posterior IFG (i.e., Broca’s area)—contribute to grammatical processing in both expressive and receptive modalities. While the functions of these regions remain unclear, it is possible that, as suggested earlier, the posterior MTG represents the morphosyntactic features of words while the posterior IFG is involved in hierarchical sequencing. Much more work will be necessary, though, to determine whether these hypotheses are on the right track.

8. Recommended reading

Binder, J.R., & Desai, R.H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, 15, 527-536.

Dehaene, S. (2009). *Reading in the brain: The science and evolution of a human invention*. New York: Viking.

DeWitt, I., & Rauschecker, J.P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences*, 109, E505-E514.

Dick, A.S., & Tremblay, P. (2012). Beyond the arcuate fasciculus: Consensus and controversy in the connective anatomy of language. *Brain*, 135, 3529-3550.

Friederici, A.D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, 16, 262-268.

Guenther, F.H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*, 25, 408-422.

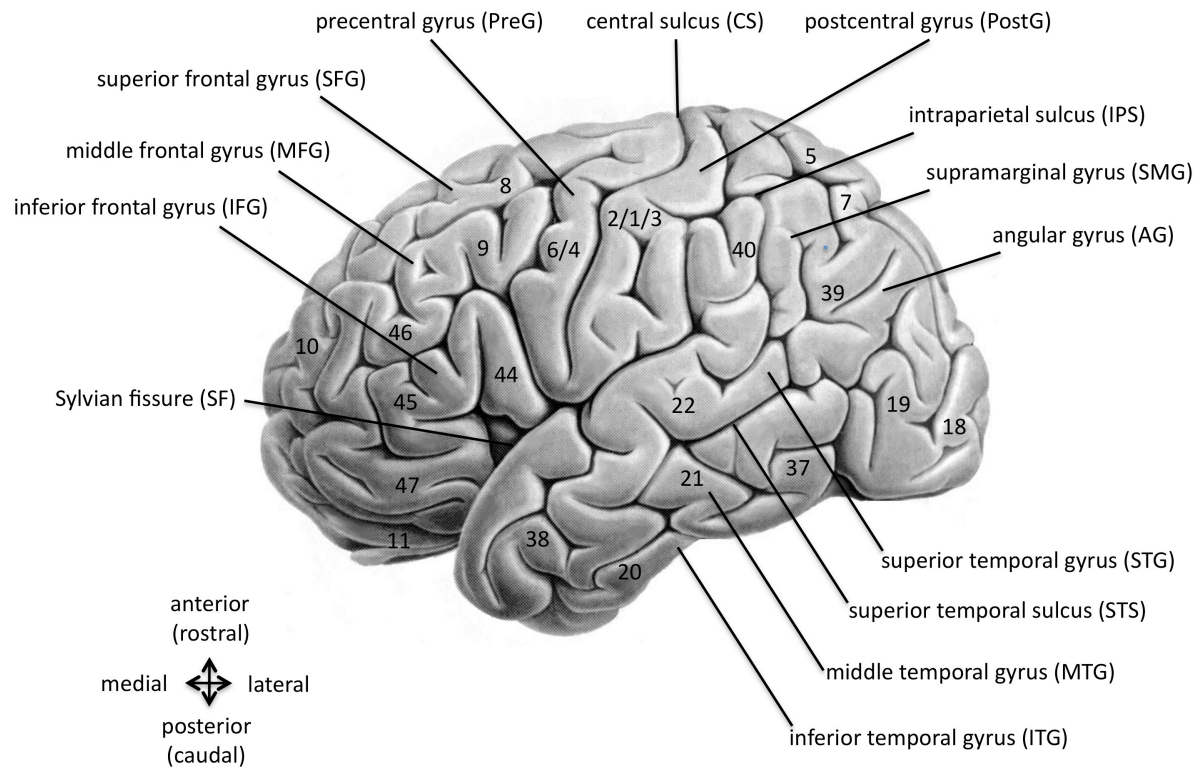
Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 393-402.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, Volume 2, Article 255.

- Jefferies, E. (2013). The neural basis of semantic cognition: Converging evidence from neuropsychology, neuroimaging, and TMS. *Cortex*, 49, 611-625.
- Kemmerer, D. (2014). *The cognitive neuroscience of language: An introduction*. New York: Psychology Press.
- Price, C.J. (2010). The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, 1191, 62-88.
- Purcell, J.J., Turkeltaub, P.E., Eden, G.F., & Rapp, B. (2011). Examining the central and peripheral processes of written word production through meta-analysis. *Frontiers in Psychology*, 2, Article 239.
- Segaert, K., Menenti, L., Weber, K., Petersson, K.M., & Hagoort, P. (2012). Shared syntax in language production and language comprehension—an fMRI study. *Cerebral Cortex*, 22, 1662-1670.
- Turkeltaub, P.E., Coslett, H.B. (2010). Localization of sublexical speech perception components. *Brain and Language*, 114, 1-15.

Figure 1. Gyral-sulcal and cytoarchitectonic organization of the left hemisphere of the human brain. Labeled lines point to major gyri and sulci, and numbered regions indicate cytoarchitectonically defined Brodmann areas. (A) Left lateral view. (B) Left ventral view.

A. Left lateral view



B. Left ventral view

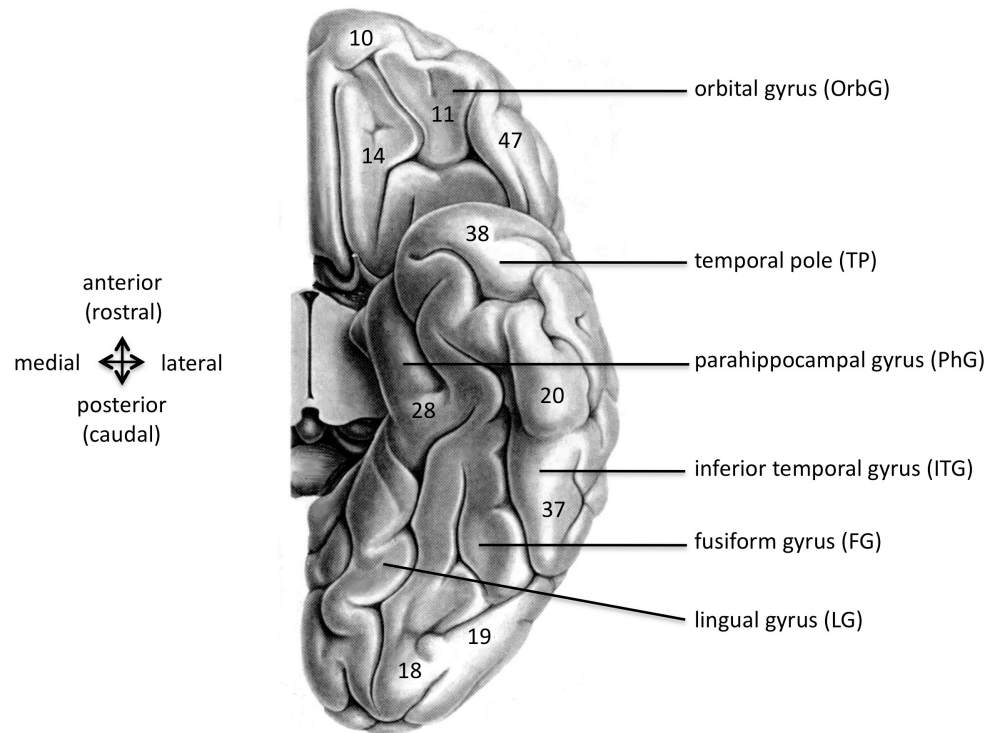
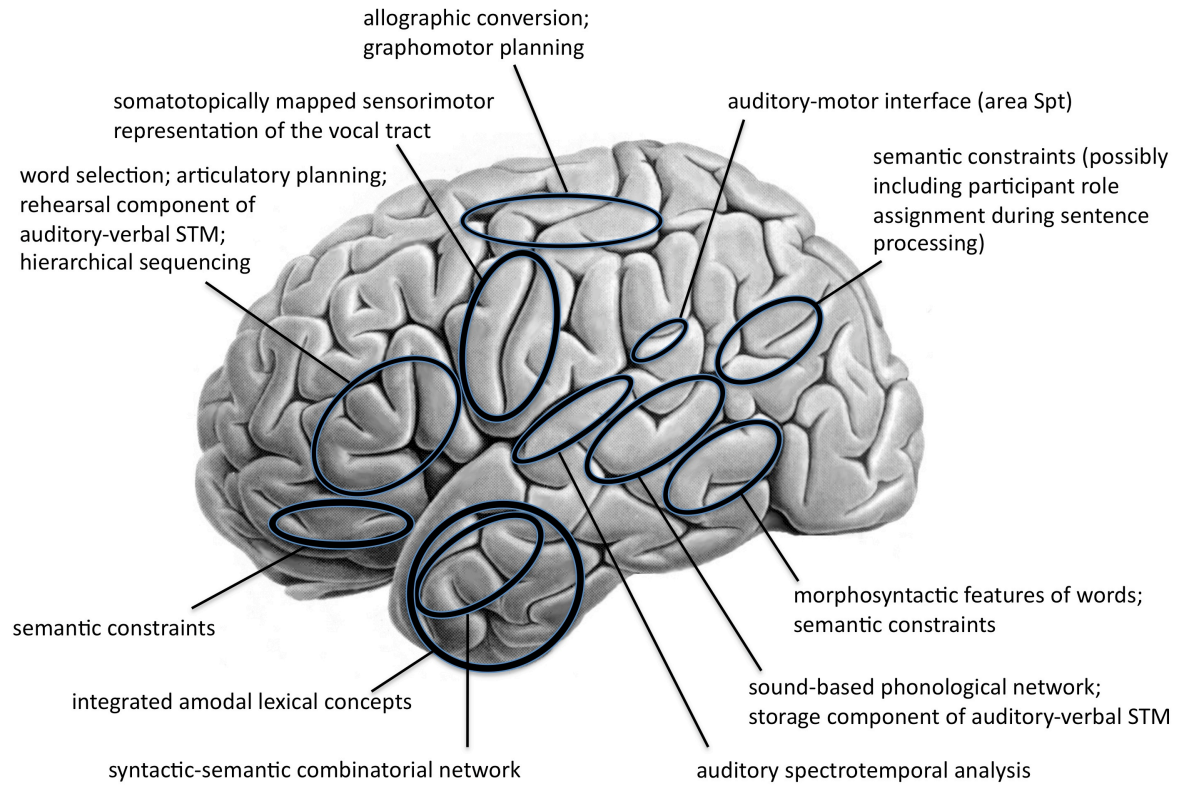


Figure 2. Illustration of the major functional-anatomical correspondences discussed in the text. (A) Left lateral view. (B) Left ventral view.

A. Left lateral view



B. Left ventral view

