

2

Excursions

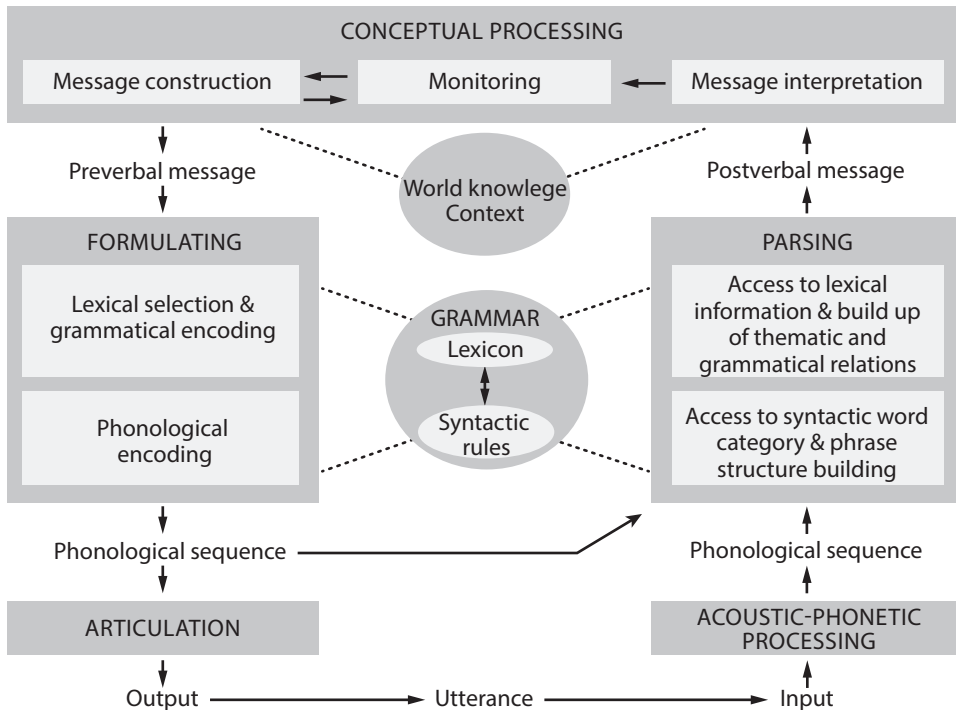
In the first chapter I discussed the brain basis of auditory language comprehension. According to a theoretical view (Chomsky, 1995b; Berwick, Friederici, Chomsky, and Bolhuis, 2013) language is assumed to consist of three larger components, a core language system that contains the syntactic rules and the lexicon, and two interface systems. These two interface systems (as I mentioned in the introduction) are the external sensory-motor interface and the internal conceptual-intentional interface. The former interface guarantees perception and production, whereas the latter interface guarantees the relation of the core system to mental concepts and intentions.

In this chapter I will touch on two issues. First, in section 2.1, I will evaluate the relation of language comprehension and production with respect to the theoretically assumed common knowledge base of the language system, consisting of syntactic rules and words. Second, in section 2.2, I will go beyond the core language system and discuss communication including intentional and situational aspects. These aspects are not part of the core computational language system but are important in everyday communication, and therefore deserve mention here.

2.1 Language Comprehension and Production: A Common Knowledge Base of Language

So far we have only discussed language comprehension and its neurobiological bases. However, there is a clear assumption that language comprehension and language production access a common knowledge base consisting of a lexicon and syntactic rules (Chomsky, 1965; Friederici and Levelt, 1988). The model by Friederici and Levelt (1988) was formulated on the basis of this theoretical consideration as well as empirical data available at the time. Here I will first describe the model as depicted in figure 2.1 and then turn to recent empirical data to back up the model.

The model assumes language production (left side of figure 2.1) and language comprehension (right side of figure 2.1) to consist of different subprocesses that are partly sequentially organized. These processes access a central knowledge base, known as the grammar, which consists of the lexicon and syntactic rules (depicted as a circle in the middle of the

**Figure 2.1**

Model of language production and comprehension. Central knowledge bases are represented as circles, processes are represented as boxes. For a detailed description see text. Adapted from Friederici and Levelt (1988). *Sprache*. In *Psychobiologie. Grundlagen des Verhaltens*, ed. Klaus Immelmann, Klaus Scherer, Christian Vogel and Peter Schmoock, 648–671. Stuttgart: Gustav Fischer.

figure). Also common to production and comprehension is the level of conceptual processing at which messages are constructed and at which perceived messages are interpreted. This is done with respect to context and the available world knowledge. Message interpretation and construction are permanently monitored. The intermediate representations of a preverbal message and postverbal message mediate between the conceptual system and the language system (production/comprehension). The subprocesses during comprehension—including the stages of syntactic phrase structure building on the basis of word category information and the subsequent processes of assigning syntactic and thematic relations—have been described in detail in chapter 1. This entire process is often called parsing.

The subprocesses during language production, known as formulating, are described in more detail and substantiated by empirical evidence in this section. The phonological sequences at the bottom of the figure are again an intermediate representation between the language system and the input/output processes. The arrow going from the phonological

sequence representation during production to the comprehension system is evidenced by the observation that we as humans are able to stop the production process before articulation in case we realize that it would be better not to utter the constructed message. Once it is articulated, not only our own comprehension system (as indicated by the arrow at the very bottom), but also the comprehension system of others, can perceive the utterance.

Neuroscientific work on language production—at least at the sentence level—is still conducted less frequently than that on comprehension. This stands in clear contrast to research on the mechanisms of speech production with its motoric and articulatory aspects. I refer readers to a recent book, *Neural Control of Speech* (Guenther, 2016), which describes these processes in much detail. My focus in this section is rather on those processes of language production that precede articulatory speech output, and the relation of these to language comprehension.

I will first discuss a number of studies that led to the model by Friederici and Levelt (1988), and then subsequent studies in support of this model. As studies of language comprehension were presented in detail in chapter 1, I will focus here primarily on studies of language production.

Over the last few decades, researchers have taken different approaches to investigate language production. One method was to study language breakdown after brain lesions, a second way was to systematically investigate speech errors in healthy speakers, and a third approach was to trigger the to-be-formulated message by means of visually or auditorily presented material in experimental settings. I will address these different approaches in turn and show that although the neuroscientific database is limited, they converge to a model of language production that appears to display the reverse order of the three phases/stages that we discussed for language comprehension.

Language Breakdown in Patients

Historically, it was Paul Broca (1865) who, after investigating a patient with severe production deficits and dissecting his brain, claimed that the inferior frontal gyrus was responsible for language production (see details in the introduction of this book). Arnold Pick (1913) described an interesting production deficit after brain lesion, called agrammatic aphasia, as a deficit characterized by an inability to produce grammatical function words (conjunctions, auxiliaries, etc.) and grammatical morphemes (verb inflections and plural morphemes) in a sentence leading to a telegraphic style. The output of these patients was non-fluent with a distortion in the prosodic range of speech. Pick already assumed different stages in the process of language production because Broca's aphasia, a deficit in language production, could be separated from dysarthria, a deficit in articulation. Moreover, he postulated that syntactic and prosodic planning must go hand-in-hand during language production. Work with progressive non-fluent aphasia, that is, patients who suffer from cortical atrophy in prefrontal regions bilaterally and perisylvian regions in the left hemisphere, has demonstrated that speech errors at the phonemic level of language production are not

caused by a motor planning impairment (Ash et al., 2010). These data provide evidence for a separation of levels of language planning and motoric articulation. Other types of aphasia (Wernicke's aphasia and so-called conduction aphasia) were characterized by a seemingly fluent language production, but with sound and word substitutions as well as a misordering of these elements in the speech output (Goodglass and Kaplan, 1972). These different types of aphasia already suggested that syntax, lexical processing, and articulation are to be thought of as different subcomponents or processing stages in language production.

Different levels of processing during language production were also evidenced by language impairments caused by stroke. At the phonological level Broca's aphasics with lesions in the frontal cortex, often including motor areas, show production errors that suggest deficits at the level of articulatory planning. They show impairments in the timing relations required for the production of consonants (Blumstein, Cooper, Goodglass, Statlender, and Gottlieb, 1980; Kurowski, Hazen, and Blumstein, 2003). The characteristic deficit of Broca's aphasia, however, lies at the level of grammatical encoding. In their production they often leave out function words and inflectional morphemes, resulting in utterances that appear like a telegram.

Aphasia research was the first to provide neurobiological evidence for a common knowledge base, called grammar in the model presented in figure 2.1. In two seminal papers (Zurif, Caramazza, and Myerson, 1972; Caramazza and Zurif, 1976) researchers demonstrated that Broca's aphasics, characterized by their deficit in syntactic processing during language production, were also deficient in sentence comprehension when sentences could only be understood on the basis of syntactic cues, e.g., *The lion was chased by the tiger*. For such a sentence the correct interpretation relies on the processing of the function words *was* and *by*. The conclusion from these observations was in line with the linguistic theory postulating a common knowledge base of syntax for language production and language comprehension.

Linguistic theory even made more fine-grained assumptions with respect to the two functional subcomponents of grammar: lexicon and syntactic rules (Chomsky, 1965). A crucial test of these would be to show that the same phonological word would be processed differently depending on its grammatical function. For example, the ambiguous word *can* should be processed differently in the phrasal context of *the can* compared to *I can*. In the latter context the word *can* is syntactic function word, whereas in the former context it is lexical item. It was demonstrated that an agrammatic Broca's patient was able to read the word *can* when presented in the *the can* context but not in the *I can* context (Andreewsky and Seron, 1975). A similar function-based processing differentiation was demonstrated in German for the prepositional form *auf*, which can either function as syntactic obligatory preposition carrying no lexical meaning (*Peter hofft auf den Sommer* / Peter hopes for the summer) or as a lexical preposition carrying referential meaning (*Peter steht auf dem Stuhl* / Peter stands on the chair) (Friederici, 1982, 1985). While Broca's aphasics were able to

produce and process prepositions in their lexical role, they failed to produce and process syntactically based prepositions.

Production errors made by patients with Wernicke's aphasia, which is caused by lesions in the temporal cortex, are partly located at the lexical level. These patients produce semantic paraphasias and have problems with the correct naming of pictures representing objects. They show deficits in accessing the lexical item in the lexicon not only during production but also during perception when judging and selecting lexical items (Milberg, Blumstein, and Dworetzky, 1987). These findings again are in line with the view of a central knowledge base of words—the lexicon—for language production and comprehension and thereby with the model proposed by Friederici and Levelt (1988).

These findings impressively demonstrate that the brain appreciates and reflects crucial differentiations made in linguistic theory. Moreover, they provide neurological evidence for a common knowledge base of language for both comprehension and production.

Speech Errors and Models of Language Production

A second line of research in language production provided crucial information for model building. This was the research on speech errors, often referred to as *slip of the tongue*. Accepting speech errors as linguistic evidence was a successful way to advance our understanding of the processing stages underlying language production. The speech error corpora of Meringer and Mayer (1895) mark the beginning of this approach. Fromkin (1973) continued to use speech errors as a systematic reflection of how sounds and words of a language are organized in the mind prior to speech output. The analysis of speech errors is still viewed as a means to uncover the workings of a highly complex system that sometimes only allows insights once the system breaks down. Therefore, speech errors are of high explanatory value.

Garrett (1980) developed a model of language production based on a systematic classification of different types of speech errors (anticipations, shifts, exchanges, substitutions, and blends). In particular, exchange errors—*the room to my door* instead of *the door to my room* (Garrett, 1980)—suggest that during production lexical items are put into “syntactic frames.” The syntactic frames come with the fixed positions of function words and syntactic word category slots and in such a frame (only) words that belong to the same syntactic category can be exchanged. The cited speech error is only one example of how the analysis of a particular error type can be used to infer the underlying processing steps. Other errors lead to the assumption of other processing stages. The resulting language production model differentiates three processing stages intervening between the conceptual message level and the level of articulation. First, semantic factors select lexical items and their grammatical relations (leading to the functional level of representation). Second, syntactic factors select positional sentence frames with their function words and grammatical morphemes (leading to the positional level of representation). Third, phonetic details of both lexical and grammatical items are specified prior to articulation (at the sound level). In his 1980

paper, Garrett discussed the relation between language production and language comprehension, arguing that both processes must access a common knowledge base of syntactic rules and words. This relation between language production and comprehension was specified in a model by Friederici and Levelt (1988) (see figure 2.1).

Experimental Work

A third approach used in psycholinguistics and in experimental neuroscience is to ask speakers to produce a word or a phrase that describes a given picture. In the context of this book I will only consider those studies on word production that investigate aspects beyond the motor act of articulation. For models of articulation, please refer to Guenther (1995, 2016), Hickok (2012), and Guenther and Hickok (2015). Their work discusses the programming of phoneme and syllable units and their motor output. The neural network supporting the speech output and its timing includes the cerebellum, the thalamus and thalamocortical circuits, the basal ganglia, the supplementary motor area, and the motor cortex and its interplay in time (Kotz and Schwartz, 2015).

Psycholinguistic research including speech error analyses and experimental work has led to solid models of language production (Garrett, 1980; Dell, 1986; Levelt, 1989). These models all agree that there are at least three separate processing levels: meaning, form, and articulation. Some psycholinguistic studies on healthy individuals have tried to investigate the seriality and timing of different processing stages during language production and also found evidence for three different processing phases. These studies used a picture-word-interference paradigm. This paradigm asks a speaker to name an object depicted in a picture, and during the production process the speaker receives an “interfering” word auditorily. The interfering word could be semantically similar (*goat/sheep*), a noun with a similar syntactic gender (article), or a phonologically similar noun. These words are presented at different time points before the target word (i.e., the name of the object) is articulated. Interestingly, interference of these three different conditions is separated in time: first (and furthest away from articulation) a semantic interference is observed; second, the syntactic gender is effective; and finally, immediately prior to articulation, phonological interference can be observed (Levelt, Schriefers, Vorberg, Meyer, Pechmann, and Havinga, 1991; Jescheniak and Levelt, 1994). These results clearly support models assuming a serial and incremental activation of different processing levels during language production, such as those language production models formulated by Garrett (1980) and Levelt (1989).

A brain-based study used motor readiness brain potential during a push-button reaction time paradigm to investigate the temporal sequence of syntactic and phonological processes during language production (van Turennout, Hagoort, and Brown, 1998). The readiness potential indicates the preparation of a motor action prior to its execution and can, therefore, be used as a response-related measure. Participants were requested to name colored pictures of objects in a noun phrase consisting of a morpho-syntactically inflected

adjective in Dutch, e.g. *rode tafel* (*red table*). But before producing the noun phrase they were instructed to perform a syntactic-phonological classification task requiring them to push the “syntax” or “phonological” classification button (i.e., to perform a motor action). Syntactic classification involved determining the syntactic gender of the noun, in turn determining the adjective’s inflection in Dutch. Phonological classification required the determination of the initial phoneme. The electrophysiological recording was time-locked to the picture onset. The lateralized readiness potential related to the preparation of the motor push-button action revealed a primacy of the “syntactic” response, which was followed in about 40 ms by the “phonological” response. These data provided electrophysiological evidence for the temporal sequence of syntactic preceding phonological processes during language production.

These fine-grained temporal results are evidenced even at the neurotopological level by a study using intracranial electrophysiology (Sahin, Pinker, Cash, Schomer, and Halgren, 2009). With this method, local field potentials are recorded from ensembles of neurons using electrodes implanted over language-related areas including Broca’s area. Participants were requested to read words aloud or to grammatically inflect them. Electrodes over Broca’s area revealed a sequential activation for lexical processes, grammatical (syntactic) processes, and phonological processes after visual word presentation and before the respective utterance. The observed sequence of processes recorded with the time-sensitive method of electrophysiology is in line with the model of language production as formulated by Levelt, Roelofs, and Meyer (1999), and that postulated in figure 2.1 of Friederici and Levelt (1988).

In addition to these electrophysiological studies revealing the temporal structure for language production, functional brain activation studies tried to localize production processes. Brain activation studies by means of functional magnetic resonance imaging have applied those experimental paradigms used in behavioral studies in language production, such as picture naming (as a single word or in a determiner phrase), reading aloud, verb generation (choose a verb that fits a given noun, e.g., apple: *eat*) and word fluency (generate nouns of a given semantic category (animal: *dog, cat*, etc.). These paradigms focus on single-word or phrase production, although they differ in their input to the production system. Note that in these paradigms, the message to be produced is not triggered internally but by external input. Nonetheless, these paradigms may allow us to investigate the process of language production and the neural network supporting this process to a certain extent. In a meta-analysis of 82 activation studies focusing on word production, Indefrey and Levelt (2004) report a set of regions that were reliably found for picture naming and word generation. The network includes the posterior portion of the left inferior frontal gyrus, the left precentral gyrus, the supplementary motor area, the mid and posterior parts of the left superior temporal gyrus and middle temporal gyrus, the left fusiform gyrus, the left anterior insula, the left thalamus, the right mid middle temporal gyrus, and the cerebellum. This is a large network covering all aspects of word production including articulation.

The ultimate brain-based model of language production, however, would want to localize the different aspects of language production in this neural network. These aspects are lexical selection, grammatical (syntactic) encoding, phonological encoding, and articulation. Localizations are partly proposed on the basis of comparing different production studies or on comparing different production conditions in one study.

Lexical selection. Concerning lexical selection, the following conclusions were drawn. As the left middle temporal gyrus was found in picture naming and word generation (Indefrey and Levelt, 2004), but not in studies on word reading (Turkeltaub, Eden, Jones, and Zeffiro, 2002), it was hypothesized that the left middle temporal gyrus could be the neural substrate of lexical selection in word production (Indefrey, 2007). The process of lexical selection was investigated by means of magnetoencephalography, which allows us to measure brain activation with a high temporal resolution (Maess, Friederici, Damian, Meyer, and Levelt, 2002). In this study a semantic category interference paradigm was used to target the lexical selection process. Participants were requested to name pictures of objects in two different blocks: blocks that contained objects of the same semantic category or blocks containing objects of different categories. Behaviorally, naming in same-category blocks is usually slower than naming in different-category blocks due to semantic competition between items in the same-category block. At the neural level, a significant activation difference between same-category and different-category conditions, which was observed in the time window of 150–225 ms after stimulus presentation, was localized in the mid part of the left middle temporal gyrus. Since this time window falls within the time window of lexical selection as indicated by behavioral studies, the left mid part of the middle temporal gyrus is understood to be the neural substrate for lexical selection.

Grammatical encoding. The process of grammatical encoding can be investigated experimentally in healthy people, both at the phrase level and at the sentential level. At the phrase level functional magnetic resonance imaging work points to an involvement of Broca's area (BA 44, BA 45) in syntactic processes during language production as demonstrated in a German functional imaging study on grammatical gender selection when naming a picture in a determiner phrase (Heim, Opitz, and Friederici, 2002). This condition was compared to two other conditions: a control condition (saying *jaja*) and a naming condition (i.e., naming a picture without a determiner). The control condition of articulation alone activated the premotor cortex, the precentral gyrus (BA 6), the anterior insula, and the parietal operculum. Lexical access reflected in naming a picture (without grammatical gender) compared to articulation revealed a network consisting of BA 45, fusiform gyrus, and inferior temporal gyrus. Phrase production (with gender-marked determiner) in contrast to word production (without gender-marked determiner) revealed activation not only in Broca's area but also in the superior temporal gyrus and middle temporal gyrus. These data suggest that a network of areas—which, in addition to Broca's area, involves the middle

temporal gyrus where the lexicon is thought to be localized—support determiner phrase production. This is what one might expect, because syntactic gender information of a noun is arbitrary and stored together with the respective item in the lexical entry.

The process of grammatical encoding at the sentential level was targeted in a study using a scene description paradigm (Indefrey et al., 2001). Participants had to describe actions of colored geometrical figures in full sentences (*The red square is pushing the green triangle away*), in single phrases (*red square, green triangle, push, away*) or in single words (*square, red, triangle, green, push, away*). The posterior part of Broca's area (BA 44) and the adjacent BA 6 were activated more strongly in the sentence condition compared to the word list condition. The study was conducted in German, in which the adjectives in phrases and sentences are morphologically marked for syntactic gender agreement with the noun, thus requiring morphosyntactic processes. Given this, the authors take their finding to show that BA 44 and BA 6 are involved in grammatical encoding (Indefrey et al., 2001; see also Indefrey, Hellwig, Herzog, Seitz, and Hagoort, 2004 for a replication of this finding). Other functional magnetic resonance imaging studies on grammatical encoding during production also found activation in BA 44 together with the anterior part of Broca's area (BA 45) when comparing sentence generation to word generation in German (Haller, Radue, Erb, Grodd, and Kircher, 2005) and in English for free sentence generation (Kemeny, Ye, Birn, and Braun, 2005). The additional activation of BA 45 in these studies may be due to increased demands on lexical processes, be it due to lexical access as such or be it due to lexical search and selection in free sentence generation.

Another approach to grammatical encoding during sentence production used a functional magnetic resonance imaging adaptation paradigm (Menenti, Gierhan, Segaert, and Hagoort, 2011; Menenti, Segaert, and Hagoort, 2012). This paradigm is tied to the phenomenon that activity of neural populations—and thereby the functional magnetic resonance imaging signal—decreases once a stimulus is repeated (Krekelberg, Boynton, and van Wezel, 2006). In the language production study Menenti et al., (2012) presented photographs of actions performed by an actor (thereby taking the thematic role of an agent) on an undergoer (thereby taking the thematic role of a patient) and the participants were required to produce a sentence describing the picture. The action was restricted to three verbs (*kiss, help, strangle*) and the agent-noun and patient-noun were either a man or a woman. The study varied the verb (thematic-semantic role: agent, patient), the lexical item (verbs, nouns), and the syntactic structure (active/passive construction). The functional magnetic resonance imaging data reported suggest that thematic-semantic aspects involve the middle temporal gyrus bilaterally but that lexical processes only recruit the left posterior middle temporal gyrus. The activation pattern for the syntactic manipulation involved prefrontal, inferior frontal, temporal, and parietal brain regions.¹ At this point it is interesting to note that a functional magnetic resonance imaging adaptation study on production and comprehension revealed a large overlap of those brain regions demonstrating the adaptation effects in production and in comprehension (Menenti et al., 2011). This can be taken as

empirical support for the model presented in figure 2.1 that assumes a common knowledge base in the service of production and comprehension processes.

Phonological encoding and articulation. Since all studies reported Broca's area (BA 45 and/or BA 44) activation for phrase and sentence level production compared to the respective control conditions, we can conclude that this area is involved in grammatical encoding. However, some but not all studies reported activation in the premotor cortex, BA 6. Area BA 6 may not necessarily be part of the core language system, but possibly part of the output system. The role of Broca's area and the adjacent premotor cortex and motor cortex during language production has indeed been the focus of an ongoing discussion. The role of Broca's area during language production has actually been debated since Paul Broca's first publication in 1861. While some researchers proposed that Broca's area accesses *phonological representations* of words whereas the motor regions are responsible for articulations, other researchers see Broca's area to be involved in processing *articulatory representations* (Hickok, 2012). However, in a patient study it was shown that articulatory planning deficits were correlated with brain regions in a discrete region of the left precentral gyrus of the insula, rather than Broca's area (Dronkers, 1996).

A recent study using electrocorticography recording directly from the surface of the cortex which allows to monitor brain activation in the millisecond domain has provided a clear answer to this debate (Flinker et al., 2015). In this study participants were asked to repeat a word they heard. The observed activations sequentially involved the superior temporal gyrus (auditory perception of word), Broca's area (phonological word representation), and the motor cortex (articulation of word). The brain activation in Broca's area (peaking at about 340 ms after the word was heard) clearly preceded activation in the motor cortex during articulation (peaking at about 580 ms) by about 241 ms. Data from this study suggest that during language production Broca's area may generate a phonological/phonetic representation, which is then implemented for output by the motor cortex.

Summary

Historically, first production models were built on the basis of language deficits in patients with brain lesions, and later on the basis of speech errors in healthy people. More recently, attempts have been made to apply neuroscientific methods, such as functional magnetic resonance imaging and electrocorticography, using pictures or perceived words at controlled input to the production system. The available data suggest that language production, apart from brain structures supporting the motor act of speaking, involves Broca's area in addition to temporal regions. In this section we started from the assumption that both language comprehension and language production access a common knowledge base, called grammar, consisting of the lexicon and syntactic rules. The empirical data discussed suggest that the different processing steps assumed for language comprehension reversely

mirror those of language production. They, moreover, are in line with the assumption of a common knowledge base.

2.2 Language Comprehension and Communication: Beyond the Core Language System

We have seen already that there are a number of psychological aspects of language use that are not part of the language system, but which are relevant for, and play a role in, daily communication. These include emotional prosody, discussed in section 1.8, as well as aspects of the communicative context and communicative hand gestures accompanying an utterance. In a model of language as proposed by Berwick et al. (2013), these aspects are not part of the core language system. I will touch upon the communicative aspects here because readers may expect to find such a discussion in these pages. But my excursion will be a brief because these aspects are not captured by the neuroanatomical modal of language developed in this book.

Pragmatics during Communication

Contextual knowledge during communication is covered by a research domain called pragmatics. It concerns the way in which context or contextual knowledge contributes to meaning. The understanding in a communicative situation may depend on the communication partner's ability to infer the speaker's intended meaning. To learn more about the field of pragmatics in general, I refer readers to books that cover this research field well (Levinson, 1983; Clark, 1996).

In the context of this book, the question arises how our brains deal with pragmatic information. There are at least two pragmatic domains that have been investigated with respect to their brain representations: one is the recognition of the communicative intention of the speaker and the other is the inference process taking place during discourse understanding.

Interpreting a speaker's message often requires processing the specific communicative context, which means that the listener has to process more than the words and phrases of a given utterance. The interpretation might require inferring the speaker's intention from the contextual situation or other non-linguistic information available. This would also suggest that the neurobiology of this process must extend beyond the core language system. Functional magnetic resonance imaging research indeed revealed an extended network recruiting regions typically involved in reasoning about the mental states of others, such as the medial prefrontal cortex and the temporo-parietal junction area, as well as the anterior temporal lobes bilaterally (Frith and Frith, 2003; Saxe and Kanwisher, 2003).

Some studies have investigated the brain basis of pragmatics in patients with cortical lesions. Many of these, such as the study by Ferstl, Guthke, and von Cramon (2002), investigated patients with brain lesions in the left hemisphere and right hemisphere. These authors found that patients with lesions in the left frontal brain regions, but not patients with lesions in the left temporal or right frontal region, had a deficit in drawing inferences

from a text. It is, in particular, the dorsomedial prefrontal cortex that appears to support inferencing (Siebörger, Ferstl, and von Cramon, 2007) often together with the posterior cingulate cortex (for a review, see Ferstl, 2015).

One of the few functional magnetic resonance imaging studies in this domain focused on processing indirect replies rather than direct replies in a dialog situation. A large network was found to be active: it involved the dorsomedial prefrontal cortex, right temporo-parietal junction area, and the insula as well as the right medial temporal and the bilateral inferior frontal gyrus (Basnakova, Weber, Petersson, van Berkum, and Hagoort, 2014). These different regions in the network are said to support different aspects of the inference process. The dorsomedial prefrontal cortex, the right temporo-parietal junction, and the insula may be recruited during mentalizing processes and theory-of-mind mentalizing (Saxe and Kanwisher, 2003). The left anterior insula, in particular, which is known to be involved in social affective processes and empathy, may have a role in this context (Singer and Lamm, 2009). All these processes might come into play when processing in a pragmatically complex situation, but they are certainly not part of the language system.

Thus it appears that processes drawing on contextual knowledge or information that can only be inferred indirectly from the text involve a large neural network that goes well beyond the core language network.

Gestures during Communication

In a direct interpersonal communication, the role of expressive hand gestures can be very important. And indeed speakers usually move their hands when they talk. These hand movements or gestures may have different functions: they may be automatic hand movements with no direct communicative function (for example, when speaking on the telephone), but they can also convey additional meaning during face-to-face communication.

Different types of co-speech gestures can be deployed during the act of communication: (1) so-called hand signs, like the “OK” sign, are meaningful hand postures with a clear message in a given culture (Nakamura et al., 2004); (2) iconic gestures are hand movements that represent object attributes, spatial relations, or actions; (3) deictic or pointing gestures serve as indicators to objects, events, or spaces and are used in mother-child interaction as a joined attention cue impacting language learning; (4) beats are rhythmic gestures that indicate the significance of an accompanying word or phrase. These different gesture types are used in different situations during online processing to enhance comprehension.

Iconic gestures are shown to play a supportive role in first-language acquisition (Morford and Goldin-Meadow, 1992; Capone and McGregor, 2004; Iverson and Goldin-Meadow, 2005) and in second-language learning in adults (Kelly, McDevitt, and Esch, 2009). A functional magnetic resonance imaging study on word learning in adults revealed that words learned together with iconic gestures could be learned and memorized better than those accompanied by meaningless gestures (Macedonia, Müller, and Friederici, 2011). A comparison of the brain activation for the differentially learned words found activity in the

premotor cortices for words encoded with iconic gestures. Words encoded with meaningless gestures activated a network associated with cognitive control. Moreover, other studies showed that iconic gestures enhance native speech comprehension directly. They can be used to disambiguate speech (Holle and Gunter, 2007) and are beneficial in countering the effects of difficult communication conditions independent of whether the difficulties are due to external (babble noise) or internal (hearing impairment) factors (Obermeier, Dolk, and Gunter, 2012).

The neural basis of co-speech gestures has been described by Dick and Broce (2015) based on the dual-stream model of language processing, as proposed by Hickok and Poeppel (2007). Note that this model deviates from the model presented in this book, in that these authors assume only one dorsal stream, namely, the connection from the posterior temporal cortex to the premotor cortex, which subserves a sensory-motor function. This model is considered as a framework for the neurobiology of co-speech gestures (Dick and Broce, 2015).

A number of studies used ERPs to investigate co-speech gestures and their interplay with language during speech comprehension. In some of these studies in which iconic gestures and speech were presented together, the gesture was either semantically congruent or incongruent with the speech. Incongruency led to an N400 component known from semantic incongruency in language-only studies (Kelly, Kravitz, and Hopkins, 2004; Holle and Gunter, 2007; Özyürek, Willems, Kita, and Hagoort, 2007; Wu and Coulson, 2007). The N400 findings in these studies suggest an interaction between gesture and speech at the semantic level during online comprehension.

There are a number of recent imaging studies on the processing of gesture and language, mostly touching the semantic level of sentence processing. Imaging studies on the processing of semantic information in gesture and speech, similar to the electrophysiological studies, manipulated the semantic relation between gesture and speech. In general, these studies report a neural network similar to that known from language-only studies. This includes the inferior frontal gyrus, mostly the pars triangularis, the posterior superior temporal sulcus, and the middle temporal gyrus (Willems, Özyürek, and Hagoort, 2007, 2009; Holle, Gunter, Rüschemeyer, Hennenlotter, and Iacoboni, 2008; Holle, Obleser, Rüschemeyer, and Gunter, 2010; Dick, Mok, Beharelle, Goldin-Meadow, and Small, 2014).

Functionally, this raises a most interesting question concerning gesture and speech: Where in the brain does the integration between gesture and speech occur? A couple of functional magnetic resonance imaging studies address this issue directly. Some researchers suggest the posterior temporal cortex as the region where the integration of iconic gesture and speech takes place (Holle et al., 2008, 2010; Green et al., 2009). Other researchers view the left inferior frontal gyrus as the region to support the integration of gesture and speech (Willems et al., 2007, 2009). The latter two studies compared brain activation of speech with incongruent iconic gestures to speech with iconic gestures. In contrast, the former studies, which found their integration effects in the posterior temporal cortex,

compared speech with gestures of different degrees of meaningfulness (dominant and subdominant meaning) to speech with meaningless, that is, grooming hand movements (Holle et al., 2008). In a similar vein, the findings from Gunter, Kroczeck, Holle, and Friederici (2013) suggest a major role of the posterior temporal cortex and not of the inferior frontal gyrus as a region where integration of gestures and speech takes place.

Meaningful gestures led to increased activation in the premotor cortex and in the parietal cortex when meaningful gestures were compared to grooming hand movements (Holle et al., 2008). An involvement of the premotor cortex was also found in two other studies on the processing of co-speech gestures, thereby providing supportive evidence (Willems et al., 2007; Josse, Joseph, Bertasi, and Giraud, 2012). However, other studies did not find these activations (Green et al., 2009; Willems et al., 2009), thus leaving a number of open questions. Further research must be carried out to show the exact role of the premotor cortex and the parietal cortex during the processing of meaningful gesture and speech during comprehension.

Whereas some gestures indicate a meaning by itself, in beat gestures the speaker tries to put a certain element of the sentence in focus. Such gestures were highlighted in an electrophysiological study in which the impact of a beat gesture on the processing of syntactically ambiguous German sentences was investigated (Holle et al., 2012). The ambiguous sentences in the study allowed for either a simple subject-first or a more complex, less-preferred, object-first structure. The ERP component P600, usually observed for the more complex object-first structure, was reduced by the beat gesture—indicating an enhanced processing of the complex sentence in the presence of a beat gesture. Note, however, that the P600 reflects a late processing stage during which different types of information are integrated. This result was extended by a functional magnetic resonance imaging version of the experiment (Gunter et al., 2013). Here, clear main effects of syntactic complexity (complex compared to simple syntax) were found in the left inferior frontal gyrus in the pars opercularis (BA 44), the left posterior middle temporal gyrus, the left pre-supplementary motor area, the left precentral gyrus, and bilateral insulae. All these areas, except the left inferior frontal gyrus, showed a significant interaction with beat gesture in such a way that a beat on the first noun phrase facilitates the processing of the easy subject-first sentence structure and inhibits the complex object-first structure, and vice versa for a beat on the second noun phrase. These findings indicate that gesture and speech can interact when trying to overcome preferences in syntax structure. Interestingly, however, no such interaction takes place in BA 44, the only area for which a main effect of syntax is found.

The available data indicate that meaningful gestures can enhance speech comprehension. At the neurophysiological level the integration of meaningful gesture and speech appears to interplay between the posterior temporal cortex and the inferior frontal gyrus. Beat gestures—when relevant for the assignment in syntactic structure—also interact with speech in prefrontal and temporal regions, but not in BA 44. This region known to support syntactic structure building reveals a syntax-only effect. This latter result is of particular

interest, since it demonstrates that syntax, as a major part of the core language system, is independent of communicative gestures.

Summary

There are important aspects to be considered for communication that are beyond the core language system we reviewed in this section. These are contextual knowledge, known as pragmatics, as well as communicative hand gestures, which may interact with language during communication. At the neuroscientific level a number of brain regions beyond those involved in language, such as the dorsomedial prefrontal cortex and the temporo-parietal junction, have been identified to support aspects of social communication. The interplay between meaningful gestures with language is thought to take place in the posterior temporal cortex at the junction of the parietal cortex. Interestingly, BA 44 as the main syntactic processing region remains unaffected by communicative gestures. At the moment the available data on the communication aspects of language in general are sparse, but a recent neuroanatomical model of social communication (Catani and Bambini, 2014) may serve as a road map for future research.

This is a section of [doi:10.7551/mitpress/11173.001.0001](https://doi.org/10.7551/mitpress/11173.001.0001)

Language in Our Brain

The Origins of a Uniquely Human Capacity

By: Angela D. Friederici

Citation:

Language in Our Brain: The Origins of a Uniquely Human Capacity

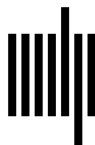
By: Angela D. Friederici

DOI: 10.7551/mitpress/11173.001.0001

ISBN (electronic): 9780262342964

Publisher: The MIT Press

Published: 2017



The MIT Press

© 2017 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

This book was set in Syntax LT Std and Times New Roman by Toppan Best-set Premedia Limited. Printed and bound in the United States of America.

The cover displays the white matter fiber tracts of the human brain for the left and the right hemispheres provided by Alfred Anwander, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.

Library of Congress Cataloging-in-Publication Data

Names: Friederici, Angela D., author.

Title: Language in our brain : the origins of a uniquely human capacity / Angela D. Friederici ; foreword by Noam Chomsky.

Description: Cambridge, MA : The MIT Press, [2017] | Includes bibliographical references and index.

Identifiers: LCCN 2017014254 | ISBN 9780262036924 (hardcover : alk. paper)

Subjects: LCSH: Cognitive grammar. | Cognitive learning. | Cognitive science. | Brain--Localization of functions. | Brain--Language. | Brain--Physiology. | Brain--Locations of functionality. | Psycholinguistics.

Classification: LCC P165 .F64 2017 | DDC 401/.9--dc23 LC record available at <https://lccn.loc.gov/2017014254>

10 9 8 7 6 5 4 3 2 1