9 Modeling the Attentional Control of Vocal Utterances: From Wernicke to WEAVER++

Ardi Roelofs

INTRODUCTION

In *Die Sprache*, Wundt (1900) criticized the now classic model of normal and aphasic utterance production of Wernicke (1874, 1885, 1886) by arguing that producing verbal utterances is an active goal-driven process rather than a passive associative process proceeding from stimulus to vocal response, as held by the model. According to Wundt (1900, 1904), an attentional process located in the frontal lobes of the human brain actively controls an utterance perception and production network located in perisylvian brain areas, described by the Wernicke model. Modern models of vocal utterance production such as WEAVER++ (Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992, 1997, 2003) build in many respects on the Wernicke model, but also address Wundt's critique by implementing assumptions on how the production–perception network is controlled. Characteristics of vocal utterance production, such as production onset latencies, errors, and corresponding brain activity, arise from the interplay of the production–perception network and the attentional control system. For example, patterns of speech errors by normal and aphasic speakers seem to be determined, at least in part, by self-monitoring, which is an important attentional control function (Roelofs, 2004). Models can benefit aphasia therapy. As Basso and Marangolo (2000) stated, "Clearly articulated and detailed hypotheses about representations and processing of cognitive functions allow rejection of all those strategies for treatment that are not theoretically justified. The more detailed the cognitive model, the narrower the spectrum of rationally motivated treatments" (p. 228).

The remainder of this chapter is organized as follows. I start by describing some of the key characteristics of the classic Wernicke model and outline Wundt's critique that the model lacks attentional control mechanisms. According to Wundt, understanding attentional control is important for aphasia therapy, because control processes may partly compensate the negative effects of lesions on language performance. Next, I describe vocal utterance production and perception in the WEAVER++ model (Levelt et al., 1999; Roelofs, 1992, 1997, 2003) as well as the model's assumptions on attentional control. I then review brain imaging evidence on the attentional control of word production, which has confirmed Wundt's suggestion that control processes are localized in the frontal lobes. Controversy exists about the role of one of the frontal areas, the anterior cingulate cortex (ACC). Researchers generally agree that the ACC plays a role in the contextual regulation of nonverbal vocal utterances, including monkey calls and human crying, laughing, and pain shrieking. However, no agreement exists on the role of the ACC in spoken word production. Some researchers deny any role for the ACC in word production (e.g., Jürgens, 2002), while others assume involvement of the human ACC but disagree on whether the ACC plays a regulatory role (Posner & Raichle, 1994; Posner & Rothbart, 2007; Roelofs & Hagoort, 2002) or a role in detecting conflict or predicting error likelihood (Brown & Braver, 2005; Miller & Cohen, 2001; Sohn, Albert, Jung, Carter, & Anderson, 2007). I review brain imaging evidence from my own laboratory for a regulatory role of the human ACC in attentional control. Finally, avenues for future research are indicated.

WERNICKE'S MODEL AND WUNDT'S CRITIQUE

In a small monograph published in 1874, called *The aphasia symptom-complex: A psychological study on an anatomical basis.* Wernicke presented a model for the functional neuroanatomy of vocal utterance production and comprehension. During the past century, the model has been extremely influential in directing and organizing research results on aphasic and normal language performance. According to Wernicke (1874), verbal vocal utterances require both cortical and brainstem mechanisms, whereas nonverbal vocal utterances, such as crying, only need brainstem circuits. Figure 9.1 illustrates the structure of the model.

At the heart of the model, auditory images for words are linked to motor images for words. The auditory images were presumed to be stored in what is today called Wernicke's area, which includes the left posterior superior temporal lobe. The motor images were assumed to be stored in Broca's area, which consists of the left posterior inferior frontal lobe. The model assumes that when a word is heard (Wernicke used the example of hearing the word BELL), auditory signals from sensory brainstem nuclei travel to the primary auditory cortex and then activate the auditory images for words. The auditory images activate associated concept images, which include sensory images of the visual and tactile impressions of the corresponding object. This leads to comprehension of the word. In repeating a heard word, the auditory images activate the corresponding motor images in Broca's area, which then activate the motor nuclei in the brainstem via primary motor cortex. In naming a pictured bell, concept images corresponding to the bell are activated, which then activate the motor images. The motor image activates the auditory image for the word, which in turn activates the motor image. This reverberation of activation stabilizes the activation of the motor images, which serves a monitoring function.

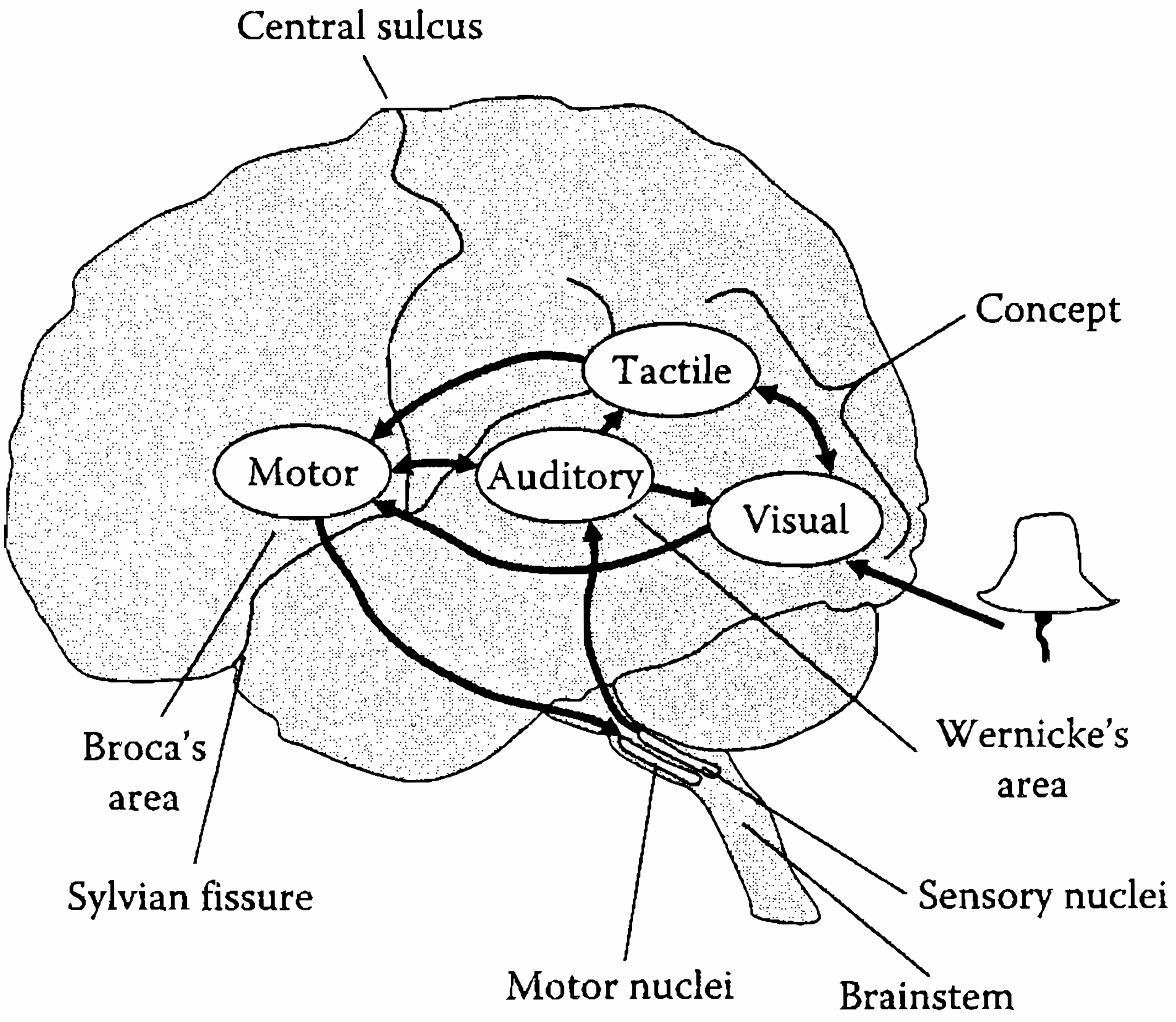

FIGURE 9.1 Illustration of the functional neuroanatomy for vocal utterance production and perception assumed by Wernicke's model. Concept, auditory, and motor images are located in a left-lateralized perisylvian network of brain areas.

Wernicke (1874, 1885, 1886) developed the model to explain various types of aphasia, which were presumed to be the result of different loci of brain damage. For example, according to the model, damage to the auditory images gives rise to speech comprehension deficits (today called Wernicke's aphasia), whereas damage to the motor images gives rise to deficits in speech production (Broca's aphasia). The reverberation of activation between motor and auditory images in speech production explains why brain-damaged patients with speech recognition deficits (people with Wernicke's aphasia) often have fluent but phonemically disordered speech production. When the auditory word images are lesioned, activity of the motor images no longer sufficiently stabilizes, explaining the phonemic paraphasias.

It is outside the scope of the present chapter to evaluate the scientific merits of Wernicke's model. For a description of the impact of the model during the past century, I refer to Shallice (1988). Overviews of modern theorizing on aphasia may be found in, for example, Caplan (1992), Martin (2003), Nickels (1997), and Rapp (2001). Relevant for the present chapter is the critique on the model advanced by Wundt (1900). According to Wundt, the retrieval of words from memory is an active goal-driven process rather than a passive associative process, as held by Wernicke's model. In particular, an attentional process located in the frontal lobes of the human brain controls the word perception and production network located in perisylvian brain areas, described by the Wernicke model. Consequently, characteristics of normal and impaired vocal performance arise from interactions between the production–perception network and the attentional control system.

Wundt (1900) maintained that such interactions may partly compensate the negative effects of lesions on language performance. In support of this claim, he describes an anomic patient who used the strategy of naming objects by first writing the object name down and then reading the written name aloud (today, recognized as a common compensatory strategy, cf. Nickels, 2002). Furthermore, the patient failed to name attributes of objects both when asked for the attribute ("what is the color of blood?") and when the attribute was shown (a red patch) but not when he actively got hold of the object together with the attribute (i.e., a drop of blood). This suggests that actively going after an enriched input may help remedy word retrieval problems. Wundt speculated that the strategic use of alternative routes through the perception-production network could lead to new network associations substituting the damaged ones. He recommended extended practice on using the alternative route as a form of aphasia therapy. Today, this is one of the approaches to therapy for naming disorders (e.g., Nickels, 2002). Studies of phenomena such as central nervous system repair, cortical reorganization after brain damage, and the improvement of language function by behavioral therapy, support the view that patients may regain lost capabilities by extensive training (e.g., Taub, Uswatte, & Elbert, 2002). Surprisingly, although Wundt's critique on Wernicke's model seems fundamental, assumptions about attentional control have typically not been part of computational models of word production and perception that have been developed during the past century (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Rapp & Goldrick, 2000).

In a review of the literature on treatment for word-retrieval disorders, Nickels (2002) concludes that

> There can be no doubt that therapy for word-retrieval impairments can be highly successful, resulting in long-term improvements which can be of great communicative significance for the individual with aphasia. However, predicting the precise result of a specific treatment task with a specific individual with certainty is still not possible. (p. 935)

According to her, "If one is ever to achieve (or even attempt) prediction in treatment, between task and impairment, a clearly articulated theory of the levels of processing that can be impaired is essential" (p. 955). One such theoretical effort is reviewed next.

THE WEAVER++ MODEL

In a seminal article, Norman and Shallice (1986) made a distinction between "horizontal threads" and "vertical threads" in the control of perception and action. Horizontal threads are strands of processing that map perceptions onto actions and vertical threads are attentional influences on these mappings. Behavior arises from interactions between horizontal and vertical threads. WEAVER++ (Roelofs, 1992, 1997, 2003, 2004, 2007) is a model that computationally implements specific claims about how the horizontal and vertical threads are woven together in the planning and comprehending of spoken words. Different from Wernicke's model, WEAVER++ was designed to explain evidence from word production latencies (see Levelt et al., 1999, for a review). I first describe the functional claims of the model, and then the presumed neuroanatomical correlates, as assessed by functional brain imaging studies rather than lesion-deficit analyses.

FUNCTIONAL ASPECTS

Whereas the Wernicke model assumes an associative network of concept, auditory, and motor images for words, WEAVER++ distinguishes concepts, lemmas, morphemes, phonemes, and syllable motor programs, as illustrated in Figure 9.2. For example, naming a pictured bell involves the activation of the representation of the concept BELL(X), the lemma of *bell* specifying that the word is a noun (for languages such as Dutch, lemmas also specify grammatical gender), the morpheme <bell>, the phonemes /b/, /e/, and /l/, and the syllable motor program [bel]. In the model, activation spreads from level to level, whereby each node sends a proportion of its activation to connected nodes. Consequently, network activation induced by perceived objects decreases with network distance. The activation

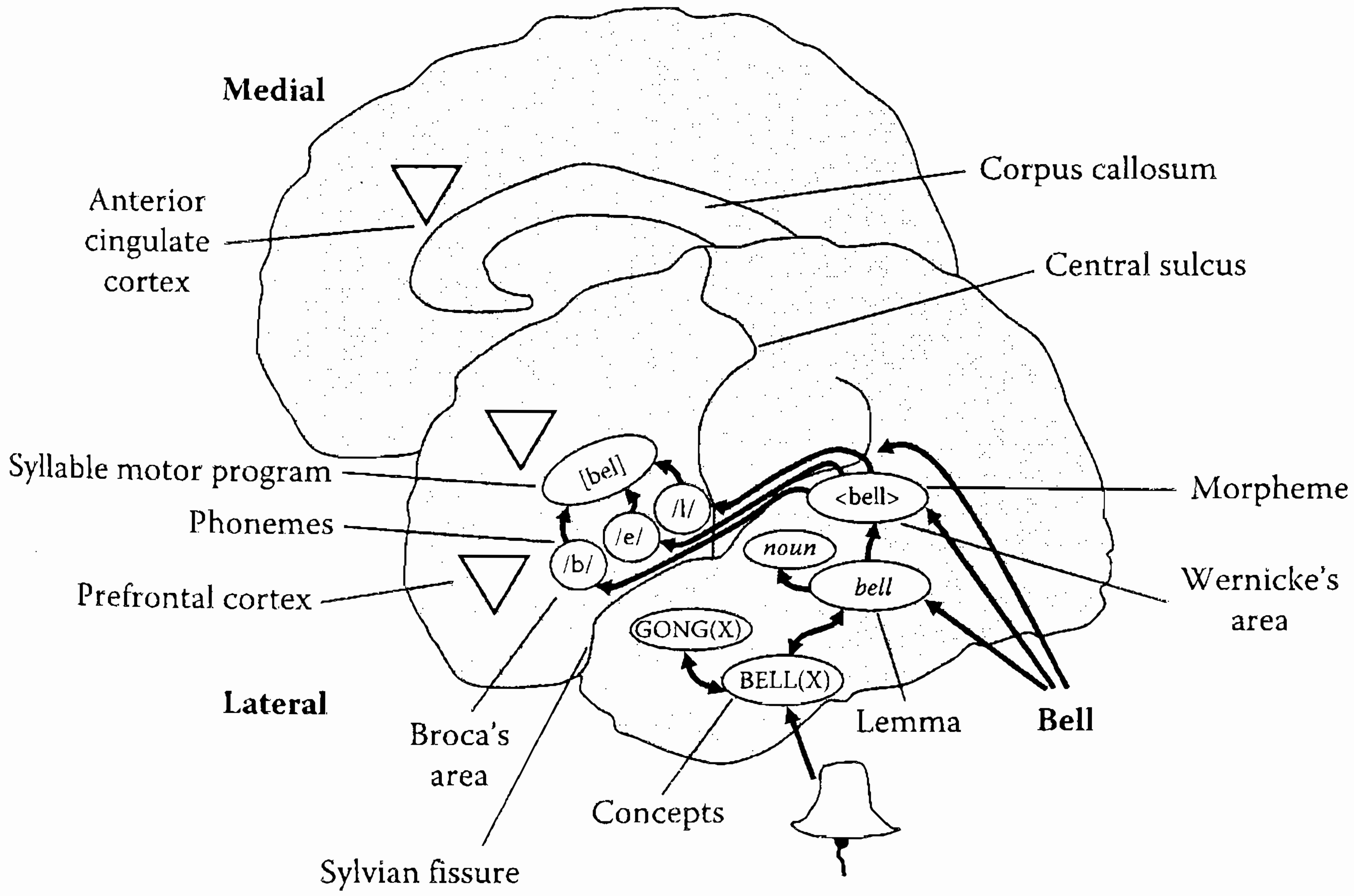

FIGURE 9.2 Illustration of the functional neuroanatomy for vocal utterance production and perception assumed by the WEAVER++ model. Only cortical areas are shown. Representations of concepts (e.g., BELL(X)), lemmas (e.g., *bell* specifying that the word is a noun), morphemes (e.g., <bell>), phonemes (e.g., /b/, /e/, and /l/), and syllable motor programs (e.g., [bel]) are located in a left-lateralized perisylvian network of brain areas. An attentional control system located in anterior cingulate cortex and in ventro- and dorsolateral prefrontal cortex (indicated by triangles) exerts regulatory influences over the lexical network.

flow from concepts to phonological forms is limited unless attentional enhancements are involved to boost the activation (Roelofs, 1992, 2003).

Following Wernicke's model, WEAVER++ assumes that perceived objects have direct access to concepts (e.g., BELL(X)) and only indirect access to word forms (e.g., <bell> and /b/, /e/, /l/), whereas perceived spoken and written words have direct access to word forms and only indirect access to concepts (cf. Roelofs, 1992, 2003, 2005, 2006, 2007). Consequently, naming objects requires concept selection, whereas spoken words can be repeated without concept selection. The latter is achieved by mapping input word-forms (e.g., the input phonological or orthographic form of BELL) directly onto output word-forms (e.g., <bell> and /b/, /e/, /l/), without engaging concepts and lemmas, as illustrated in Figure 9.2.

Whereas Wernicke's model lacks attentional control mechanisms, these are present in WEAVER++. In particular, a goal-driven selection of information from the lexical network in WEAVER++ is regulated by a system of condition-action rules. When a goal is placed in working memory, word planning is controlled by those rules that include the goal among their conditions. For object naming, a rule would specify that:

IF the goal is to say the name of the object,
and the concept corresponds to the object,
THEN select the concept,
and enhance its activation.

The activation enhancements are required until appropriate syllable motor programs have been activated above an availability threshold. The attentional control system determines how strongly and for how long the enhancement occurs (Roelofs, 2003, 2007). A speaker may assess the required duration of the enhancement by monitoring the progress on word planning (Roelofs, 2004, 2007). Condition-action rules allow for the specification of alternative naming routes. For example, the compensatory writing strategy discussed earlier may be specified as:

IF the goal is to say the name of the object,
and naming fails,
THEN set the goal to write the name of the object,
and read aloud the written name.

A compensatory strategy that has been studied in much detail is letter-by-letter reading in pure alexia (e.g., Cohen et al., 2004; Henry et al., 2005; Shallice, 1988), which is a frequent consequence of left occipitotemporal cortex damage. This area seems to implement the abstract identity of strings of letters. Pure alexic patients often retain single letter recognition abilities, and develop an effortful letter-by-letter strategy, which is the basis of most rehabilitation techniques. The strategy consists of silently sounding out the letters of the word from left to right. Consequently, word reading latency increases linearly with the number of letters in the word. Using brain imaging, Cohen et al. (2004) observed increased activation of the right occipitotemporal cortex in reading by a pure alexic patient compared to healthy controls, suggesting that letters were identified in the right rather than the left hemispheric area, as is normally the case. Moreover, the patient showed stronger than normal activation in left frontal and parietal areas that are implicated in phonological recoding and working memory, suggesting that the letter-by-letter strategy more strongly engages these functions. Examination of the patient 8 months later revealed decreased word reading latencies and decreased activation in the right occipitotemporal cortex (Henry et al., 2005), suggesting that the area became better at identifying letters with practice. The work of Cohen et al. (2004) and Henry et al. (2005) demonstrates the utility of functional brain imaging in assessing the effect of strategy use. Moreover, it provides some evidence for Wundt's conjecture that compensatory strategies may establish new routes through the perception-production network.

NEUROANATOMICAL ASPECTS

Following Wernicke's model, WEAVER++ assumes that the activation of representations underlying object naming proceeds from percepts in posterior cortical areas to articulatory programs in anterior areas, as illustrated in Figure 9.2. Using WEAVER++ as framework, Indefrey and Levelt (2004) performed a meta-analysis of 82 neuroimaging studies on word production, which suggested that the following cortical areas are involved. Information on the time course of word production in relation to these areas came from magnetoencephalographic studies. The meta-analysis included object naming (e.g., say "bell" to a pictured bell), word generation (producing a use for a noun, e.g., say "ring" to the spoken or written word BELL), word repetition or reading (e.g., say "bell" to BELL), and pseudoword repetition or reading (e.g., say "bez" to BEZ). Activation of percepts and concepts in object naming happens in occipital and inferiotemporal regions of the brain. The middle part of the left middle temporal gyrus seems to be involved with lemma retrieval. When the total object naming time is about 600 ms, activity in these areas occurs within the first 275 ms after an object is presented. Next, activation spreads to Wernicke's area, where the morphological code (i.e., lexical phonological code) of the word seems to be retrieved. Activation is then transmitted to Broca's area for phoneme processing and syllabification, taking some 125 ms. During the next 200 ms, syllable motor programs are accessed. The sensorimotor areas control articulation. Word repetition and reading may be accomplished by activating the areas of Wernicke and Broca for aspects of form encoding, and motor areas for articulation.

Neuroimaging studies on word planning have confirmed Wundt's (1900, 1904) suggestion that the perisylvian production-perception network is controlled by attentional control mechanisms located in the frontal lobes. In particular, attentional control processes engage the lateral prefrontal cortex (LPFC) and the ACC, as illustrated in Figure 9.2. The ACC and LPFC are more active in word generation (say "ring" to BELL) when the attentional control demands are high than in word repetition (say "bell" to BELL) when the demands are much lower (Petersen, Fox, Posner, Mintun, & Raichle, 1988; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). The increased activity in the frontal areas disappears when word selection becomes easy after repeated generation of the same use to a word (Petersen, van Mier, Fiez, & Raichle, 1998). Moreover, activity in the frontal areas is higher in object naming when there are several good names for an object so that selection difficulties arise than when there is only a single appropriate name (Kan & Thompson-Schill, 2004). Also, the frontal areas are more active when retrieval fails and words are on the tip of the tongue than when words are readily available (Maril, Wagner, & Schacter, 2001). Frontal areas are also more active in naming objects with semantically related words superimposed (e.g., naming a pictured bell combined with the word GONG) than without word distractors, as demonstrated by de Zubicaray, Wilson, McMahon, and Muthiah (2001). Thus, the neuroimaging evidence suggests that medial and lateral prefrontal areas exert attentional control over word planning. Along with the increased frontal activity, there is an elevation of activity in perisylvian areas (e.g., de Zubicaray et al., 2001; Raichle et al., 1994; Snyder, Abdullaev, Posner, & Raichle, 1995).

Evidence for the involvement of frontal areas in the attentional control of word production also comes from impaired performance. Semantic retrieval problems due to lesions of temporal areas of the human brain typically preserve the ability to generate category terms. For example, a patient may be able to say "instrument" to a bell, without being able to say "bell." Humphreys and Forde (2005) reported evidence on a patient with combined frontal-temporal damage, who had, instead, a specific impairment of generating category terms. According to Humphreys and Forde, the unusual impairment resulted because categorizing requires the attentional control provided by the frontal lobes.

Although both the ACC and LPFC are involved in the attentional control of word planning, the areas seem to play different roles. Evidence suggests that the dorsolateral prefrontal cortex (DLPFC) is involved in maintaining goals in working memory (for a review, see Kane & Engle, 2002). WEAVER++'s assumption that abstract condition-action rules mediate goal-oriented retrieval and selection processes in prefrontal cortex is supported by evidence from single cell recordings and hemodynamic neuroimaging studies (e.g., Bunge, 2004; Bunge, Kahn, Wallis, Miller, & Wagner, 2003; Wallis, Anderson, & Miller, 2001). Moreover, evidence suggests that the ventrolateral prefrontal cortex play a role in selection among competing response alternatives (Thompson-Schill et al., 1997), the control of memory retrieval, or both (Badre, Poldrack, Paré-Blagoev, Insler, & Wagner, 2005). The ACC seems to exert regulatory influences over these processes.

In the light of Darwin's continuity hypothesis (i.e., new capabilities arise in evolution by modification and extension of existing ones), the involvement of the ACC in the attentional control of spoken word production seems plausible, because the area also controls nonverbal vocal utterances, considered by many to be the evolutionary forerunner of speech (e.g., Deacon, 1989, 1997; Jürgens, 1998; Ploog, 1992). Vocal utterances of nonhuman primates (monkeys and apes) consist of innate emotional vocalizations, such as fear, aggression, alarm, and contact calls. The two most stereotypical innate vocalizations in humans are crying and laughing (e.g., Newman, 2007). Evidence suggests that the ACC plays a critical role in the voluntary initiation and suppression of these nonverbal vocal utterances (e.g., Aitken, 1981). The area does so by sending regulatory signals to the periaqueductal gray in the caudal midbrain. The periaqueductal area links emotional signals from the amygdala and other areas to the corresponding vocal responses. Also, the area links sensory stimuli, such as a heard vocal utterance, to corresponding vocal motor programs, thereby providing a low-level audio-vocal interface. Neighboring areas in the midbrain contain sensorimotor-orienting circuits underlying the automatic shift of gaze and attention and the turning of the head toward the sensory stimuli. The ACC signals the periaqueductal gray to initiate or withhold the motor program, depending on the context. The motor programs are embodied by premotor and motor nuclei in the lower brainstem and spinal cord (as assumed by Wernicke). The premotor nuclei coordinate the activity of the motor nuclei controlling the larynx, respiratory apparatus, and supralaryngeal tract. The three levels of vocal control (ACC, periaqueductal gray, lower brainstem nuclei) seem to be present in mammalian species as different as the cat and the bat (see Jürgens, 2002, for a review). For example, the ACC exerts control over the echolocation of bats (Duncan & Henson, 1994; Gooler & O'Neill, 1987), showing that the area also regulates noncommunicative use of the voice.

In the human speech system, the motor region of the posterior ventrolateral cortex directly projects onto the brainstem premotor and motor nuclei for the control of the oral, vocal, and respiratory muscles, bypassing the periaqueductal gray. Still, the ACC may exert regulatory influences over the speech system through its connections with ventrolateral prefrontal, premotor, and motor cortex (Deacon, 1997; Jürgens, 2002; Paus, 2001). The ACC seems implicated in enhancing the activation of target representations in the ventrolateral frontal areas until retrieval and selection processes have been accomplished in accordance with the goals maintained in DLPFC (Roelofs, van Turennout, & Coles, 2006). According to this view, the ACC plays a role in the regulation of both verbal and nonverbal vocal utterances (cf. Deacon, 1989, 1997; Posner & Raichle, 1994), although through different neural pathways, as illustrated in Figure 9.3.

The activation enhancements provided by the ACC constitute a kind of driving force behind vocal utterance production. This fits with the idea that for action control, it is not enough to have goals in working memory, but one should be motivated to attain them. Anatomically, the ACC is in a good position to provide such a driving force (cf. Paus, 2001). The necessary arousal may be provided through the extensive projections from the thalamus and reticular brainstem nuclei to the ACC. The information on what goals to achieve may be provided through the extensive connections between the ACC and dorsolateral prefrontal cortex. Access to the motor system by the ACC is provided by the dense projections of the motor areas of the cingulate sulcus onto the brainstem and motor cortex. The idea that the ACC provides a kind of driving force behind vocal utterance production agrees with the effect of massive damage to the ACC.

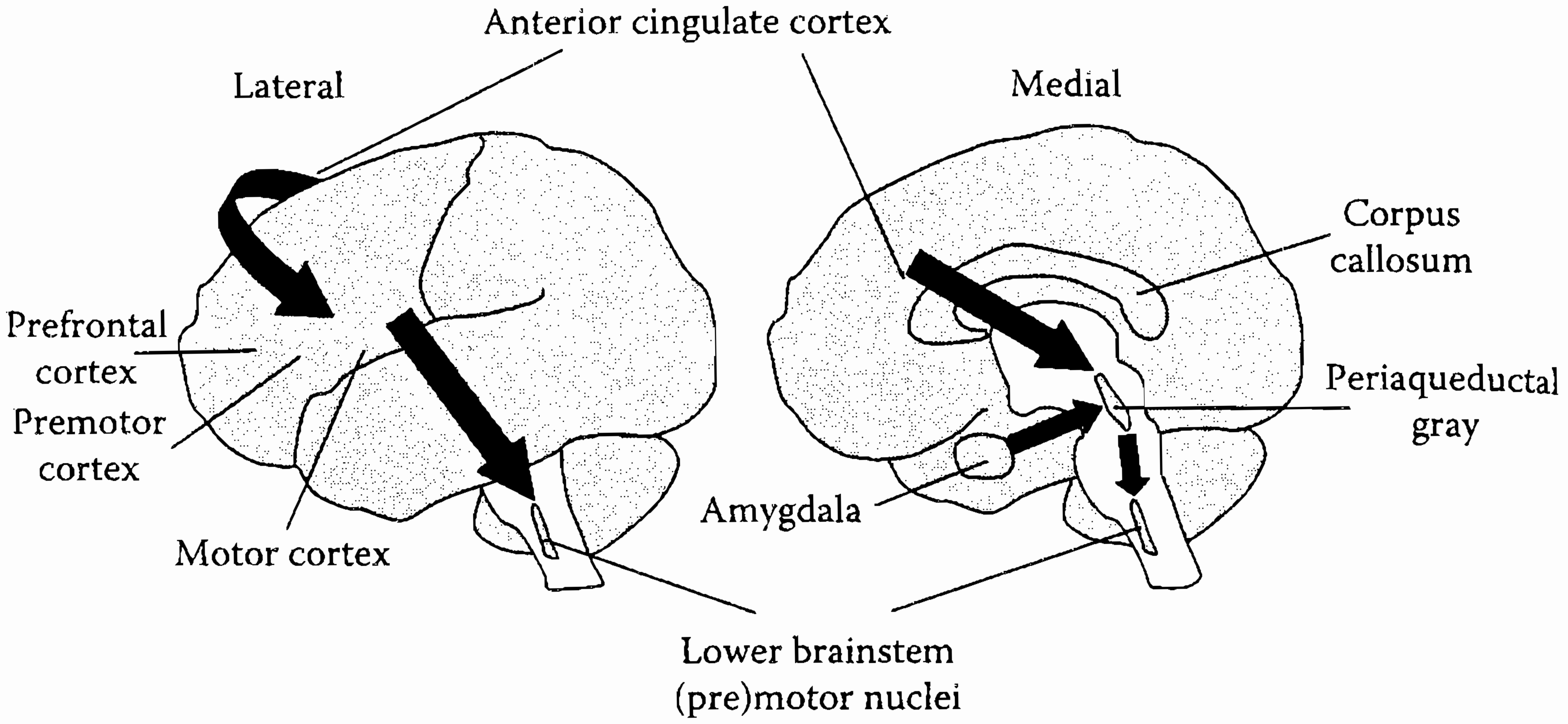

FIGURE 9.3 Illustration of the regulatory pathways of the anterior cingulate cortex in verbal (lateral view) and nonverbal vocal utterances (medial view). In verbal utterances, the ACC exerts control over lateral prefrontal, premotor, and motor cortex, which directly controls the pontine and medullary (pre)motor nuclei. In nonverbal utterances, the ACC exerts control over the periaqueductal gray.

Damage of the medial frontal cortex including the ACC typically results in transient akinetic mutism, which is characterized by reduced frequency of spontaneous speech with a preserved ability to repeat what is said (i.e., when externally triggered). The mutism arises with left or bilateral medial lesions, but damage to the right ACC may also result in transient speech aspontaneity (Chang, Lee, Lui, & Lai, 2007). Jürgens and von Cramon (1982) reported a case study of a patient with a lesion to the medial frontal cortex on the left side, which included damage to the ACC. The patient first exhibited a state of akinetic mutism during which no spontaneous vocal utterances were produced. After a few months, the patient could whisper but not produce voiced verbal utterances, which was restored only later. During the following months, the frequency of spontaneous utterances increased, but emotional intonation (e.g., an angry voice) remained impaired. A patient reported by Rubens (1975) presented with voiceless laughing and crying in the initial phase of mutism. Later, the ability to produce voiced verbal utterances was regained, but the intonation remained monotonous. According to Jürgens and von Cramon (1982), these findings indicate a role for the ACC in the volitional control of the emotional aspect of spoken utterances, but not their verbal aspect.

ROLE OF THE ACC IN ATTENTIONAL CONTROL

Although the role that I proposed for the ACC in the attentional control of both verbal and nonverbal vocal utterances seems plausible, this claim is controversial. Whereas researchers generally agree that the ACC plays a role in the attentional control of nonverbal vocal utterances, they have found no agreement on the role of the human ACC in the production of verbal vocal utterances. Some researchers deny any role for the ACC in spoken word production except for emotional aspects (e.g., Jürgens, 2002). Other researchers assume involvement of the human ACC in attentional control but disagree on whether the ACC plays a regulatory role (Posner & Raichle, 1994; Posner & Rothbart, 2007; Roelofs & Hagoort, 2002), as in call production, or a role in detecting conflict or predicting error-likelihood (Brown & Braver, 2005; Miller & Cohen, 2001; Sohn et al., 2007). I discuss the different views on the role of the human ACC in attentional control, and review brain imaging evidence from my own laboratory supporting a regulatory role of the human ACC.

EMOTIONAL VOCALIZATION

Based on three decades of electrophysiological studies of monkey vocalization, Jürgens and colleagues developed a model of the role of the ACC in vocal utterances (see Jürgens, 1998, 2002, 2009, for reviews), which has become a leading model in the literature on mammalian vocalization. This work has shown that the anterior cingulate is the only cortical area that is directly involved in call production by monkeys. Although extensive connectivity between the ACC and lateral prefrontal, premotor, and motor cortex is acknowledged (e.g., Jürgens, 2002), a role of the ACC in the control of speech is denied (Jürgens, 2002, 2009). The extensive connectivity between the ACC and lateral frontal cortex may serve functions unrelated to speech. For example, the ACC is implicated in the voluntary initiation of swallowing (Maeda et al., 2006; Watanabe, Abe, Ishikawa, Yamada, & Yamane, 2004).

However, given the neuroimaging evidence on word production reviewed above, it seems difficult to maintain that the ACC plays no role in spoken word production. Vocalization, like swallowing, requires the coordinated activity of the oral, vocal, and respiratory muscles. Since swallowing and breathing are exclusive activities, the larynx plays a critical role in gating access to the respiratory tract. MacNeilage (1998) suggested that ingestion-related capabilities in lateral frontal cortex associated with chewing, sucking, and licking were modified in human evolution to serve a role in speaking. The extensive connectivity between de ACC and lateral frontal cortex, including Broca's area and the larynx motor cortex, may have become exploited in the control of speech production (cf. Ploog, 1992). Clearly, the ACC is involved in speech production only under certain circumstances, namely when utterance production requires attentional control. However, a circumscribed role of the ACC is also observed in call production, where the area plays a role only in the contextual regulation of vocalization, but not in their production per se.

REGULATION VERSUS CONFLICT DETECTION

Whereas Jürgens (2002, 2009) denies any role for the ACC in spoken word production, other researchers assume involvement of the human ACC but disagree on whether the area plays a regulatory role (Posner & Raichle, 1994; Posner & Rothbart, 2007; Roelofs & Hagoort, 2002) or a role in detecting conflict or predicting error-likelihood. According to the latter view, the ACC is involved in performance monitoring (Miller & Cohen, 2001). In particular, ACC activity reflects the detection of response conflict and acts as a signal that engages attentional control processes subserved by LPFC.

In a review of the monkey literature on the control of gaze by the frontal lobes, Schall and Boucher (2007) conclude that neurons in the supplementary eye field (part of the medial frontal cortex) exhibit activity specifically related to response conflict, but neurons in the ACC do not. The ACC neurons only reflect the amount of control exerted. This conclusion was surprising given that the conflict detection hypothesis is the dominant view on ACC function in the literature on humans. According to Schall and Boucher, the difference between monkeys (control-related activity but no conflict-related activity in the ACC) and humans (ACC conflict detection) may reflect differences in species, tasks, or effectors. For example, monkey studies have typically employed saccadic stop-signal tasks, in which conflict is evoked by infrequently presenting a stop signal while preparation of a saccade is in progress. In contrast, human studies have typically employed Stroop-like tasks with vocal or manual responding. For example, in the color-word Stroop task, participants name the ink color of congruent or incongruent color words (e.g., the word RED printed in green), whereby vocal responding is slower in the incongruent than the congruent condition. In the arrow-word version of this task, the stimuli consist of incongruent and congruent combinations of left- or right-pointing arrows and the words LEFT or RIGHT, and the participants respond by pressing a left or right response button. According to Schall and Boucher, it is possible that the saccadic responses in monkey studies yielded less conflict than the vocal and manual responses in human studies, explaining the difference in results between monkeys and humans.

However, it is also possible that the difference in results between monkeys and humans has another ground, namely a confound between conflict and control in the human studies. Incongruent and congruent stimuli not only differ in the amount of conflict they evoke, but also in the amount of attentional control required by the corresponding responses. Thus, the conflict-related activity in Stroop-like tasks is compatible with both the regulative and conflict detection hypotheses on ACC function. It is possible to discriminate between the two hypotheses empirically by using neutral stimuli. A critical prediction made by the conflict hypothesis is that ACC activity should be increased only when conflicting response alternatives are present (e.g., in responding to the word LEFT combined with a right-pointing arrow). ACC activity should not differ between congruent trials (e.g., the word LEFT combined with a left-pointing arrow) and neutral trials (e.g., the word LEFT only), because competing response alternatives are absent on both trial types. In contrast, the regulatory hypothesis (Posner & Raichle, 1994; Posner & Rothbart, 2007; Roelofs & Hagoort, 2002) not only predicts more ACC activity on incongruent than on neutral trials, but also less ACC activity on congruent than on neutral trials. Less ACC activity is predicted because the correct response (left) is already activated by the distractor (a left-pointing arrow) on congruent trials and therefore less enhancement of the target is required.

To test between the conflict detection and regulation hypotheses about ACC function, Roelofs et al. (2006) conducted a functional magnetic resonance imaging (fMRI) study. Participants were scanned while they were presented with arrow-word combinations. The participants were asked to communicate the direction denoted by the word by pressing a left or right button using the index and middle fingers of their left hand. A meta-analysis of the existing neuroimaging literature and the results from a new neuroimaging experiment by Barch et al. (2001) has shown that Stroop-like tasks activate the dorsal ACC regardless of whether the direction of the word is communicated through a spoken or manual response. On incongruent trials in the experiment of Roelofs et al. (2006), the word and the arrow designated opposite responses. On congruent trials, the word and arrow designated the same response. On neutral trials, a word was presented in combination with a straight line, so only one response was designated by the stimulus. Congruent, incongruent, and neutral trials were presented rapidly in a randomly intermixed order. The response time data showed that, consistent with earlier findings (e.g., Roelofs, 2003 for a review), responses to the words were much slower on incongruent than on neutral trials and fastest on congruent trials. The neuroimaging data demonstrated that activity in the ACC was larger on incongruent than on congruent trials. The same held for activity in the LPFC. Importantly, ACC activity was larger for neutral than for congruent stimuli, in the absence of response conflict. This result demonstrates the engagement of the ACC in the regulation of communicative responses. This conclusion was corroborated by successful WEAVER++ simulations of the chronometric and neuroimaging findings (Roelofs et al., 2006).

ANTICIPATORY ADJUSTMENTS

People are often faced with circumstances in which certain vocal behaviors are inappropriate, such as laughing at a funeral or talking aloud in a library. This raises the question whether the ACC is also involved in adjusting the control settings for responding (e.g., raising the response thresholds) depending on the communicative situation. Sohn et al. (2007) proposed that the ACC plays a role in signaling upcoming response conflict. Brown and Braver (2005) argued that the ACC signals upcoming error likelihood, independent of response conflict. More generally, environmental cues may provide information about which type of stimulus is coming and, as a consequence, about which control setting is most appropriate for responding to the stimulus. However, these contextual cues do not necessarily have to predict response conflict or error likelihood. This raises the question whether anticipatory activity in the ACC may be obtained independent of upcoming conflict or error likelihood. Aarts, Roelofs, and Van Turennout (2008) conducted an fMRI experiment that examined this issue.

As in the study of Roelofs et al. (2006), participants were scanned while they were presented with arrow-word combinations. Again, the index and middle fingers of the left hand were used for responding. On each trial, the participants were now informed about the arrow-word stimulus conditions by means of symbolic cues, which were presented well before the arrow-word stimulus. The cue was a colored square that indicated whether the upcoming arrow-word stimulus was congruent, incongruent, or neutral, or the cue provided no information about the upcoming condition. Green squares were always preceding congruent stimuli, red squares preceded incongruent stimuli, and yellow squares preceded neutral stimuli. The uninformative cues were grey squares, which could be followed by any of the stimulus types.

If the ACC plays a role in anticipatory adjustments in control, ACC activity should be higher in response to informative cues than to uninformative cues. If the adjustments are independent of response conflict or error likelihood, enhanced ACC activity should be obtained for cues preceding congruent stimuli. Adjustments are expected in premotor cortex, where response rules are implemented (Wallis & Miller, 2003). An informative cue preceding an incongruent stimulus might encourage participants to weaken the connections between the arrows and their responses, because the arrows elicit the wrong response. However, an informative cue preceding a congruent target might encourage participants to strengthen the connections between the arrows and the corresponding responses, because the arrows now elicit the correct response. In WEAVER++, the adjustments may be achieved by condition-action rules specifying that,

> IF the goal is to indicate the direction denoted by the word,
> and the cue is green,
> THEN strengthen the connection between arrows and responses.

If such advance adjustments are successful, ACC activity should exhibit smaller differences among target conditions in response to the arrow-word stimuli after informative cues (when control was adjusted in advance) than following uninformative cues (when control was not adjusted in advance).

Aarts et al. (2008) observed that participants responded faster to the arrow-word stimuli after informative than uninformative cues, indicating cue-based adjustments in control. Moreover, ACC activity was larger following informative than uninformative cues, as would be expected if the ACC is involved in anticipatory control. Importantly, this activation in the ACC was observed for informative cues even when the information conveyed by the cue was that the upcoming arrow-word stimulus evokes no response conflict and has low error likelihood. This finding demonstrates that the ACC is involved in anticipatory control processes independent of upcoming response conflict or error likelihood. Moreover, the response of the ACC to the target stimuli was critically dependent upon whether the cue was informative or not. ACC activity differed among target conditions after uninformative cues only, indicating ACC involvement in actual control adjustments. Taken together, these findings argue strongly for a role of the ACC in anticipatory control independent of anticipated conflict and error likelihood, and also show that such control can eliminate conflict-related ACC activity during target processing.

Premotor cortex activity should reflect the operation of control in response to informative cues. Therefore, we expected a positive correlation between cue-related ACC and premotor activity. The correlation should be confined to the right premotor cortex, contralateral to the response hand. Correlation analyses confirmed that cue-based activity in the ACC was positively correlated with activity in the dorsal premotor cortex and the supplementary motor area contralateral to the response hand. Although several other frontal areas were active in response to the cues, no correlations between cue-based ACC activity and the other regions were found. These results provide evidence for a direct influence of the ACC over premotor cortex (cf. Figure 9.3).

SUMMARY AND CONCLUSIONS

This chapter outlined the classic model of Wernicke (1874) for the functional neuroanatomy of vocal utterance production and comprehension, and Wundt's (1900, 1904) critique that the model lacks attentional control mechanisms, which he localized in the frontal lobes. Next, the WEAVER++ model (Roelofs, 1992, 2003, 2007) was described, which builds in many respects on Wernicke's ideas but also addresses Wundt's critique by implementing assumptions on attentional control. Characteristics of utterance production by healthy and brain-damaged individuals arise from the interplay of a perisylvian production-perception network and the frontal attentional control system. I indicated that controversy exists about the role of one of the frontal areas, the ACC. Whereas some researchers deny any role for the ACC in spoken word production, other researchers assume involvement of the area but disagree on whether it plays a regulatory role, as in call production, or a role in detecting conflict or predicting error-likelihood. I reviewed evidence for a regulatory role of the ACC.

Aphasiologists agree that a good theoretical model is important for therapy (e.g., Basso & Marangolo, 2000; Nickels, 1997, 2002). However, according to Nickels (2002), "One of the limitations remains that while theories of language processing are becoming increasingly specified (e.g., Levelt et al., 1999), how these models will function once damaged is not at all clear (but see Dell et al., 1997, for a computationally implemented theory that has investigated the effects of "lesioning"" (p. 955). Given the importance of modeling for therapy, future research should further theoretically analyze and model vocal utterance production and its attentional control, impairments, and their interactions. I hope this chapter has provided some helpful hints for this research and for clinical practice.

ACKNOWLEDGMENT

The preparation of the chapter was supported by a VICI grant from the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

Aarts, E., Roelofs, A., & Van Turennout, M. (2008). Anticipatory activity in anterior cingulate cortex can be independent of conflict and error likelihood. *Journal of Neuroscience*, *28*, 4671–4678.

Aitken, P. G. (1981). Cortical control of conditioned and spontaneous vocal behavior in rhesus monkeys. *Brain and Language*, *13*, 171–184.

Badre, D., Poldrack, R. A., Paré-Blagoev, E., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, *47*, 907–918.

Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of response modality and processing domain. *Cerebral Cortex*, *11*, 837–848.

Basso, A., & Marangolo, P. (2000). Cognitive neuropsychological rehabilitation: The emperor's new clothes? *Neuropsychological Rehabilitation*, *10*, 219–230.

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*, 1118–1121.

Bunge, S. A. (2004). How we use rules to select actions: A review of evidence from cognitive neuroscience. *Cognitive, Affective, & Behavioral Neuroscience*, *4*, 564–579.

Bunge, S. A., Kahn, I., Wallis, J. D., Miller, E. K., & Wagner, A. D. (2003). Neural circuits subserving the retrieval and maintenance of abstract rules. *Journal of Neurophysiology*, *90*, 3419–3428.

Caplan, D. (1992). *Language: Structure, processing, and disorders*. Cambridge, MA: MIT Press.

Chang, C.-C., Lee, Y. C., Lui, C.-C., & Lai, S.-L. (2007). Right anterior cingulate cortex infarction and transient speech aspontaneity. *Archives of Neurology*, *64*, 442–446.

Cohen, L., Henry, C., Dehaene, S., Martinaud, O., Lehéricy, S., Lemer, C., & Ferrieux, S. (2004). The pathophysiology of letter-by-letter reading. *Neuropsychologia*, *42*, 1768–1780.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.

Deacon, T. W. (1989). The neural circuitry underlying primate calls and human language. *Human Evolution*, *4*, 367–401.

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. New York, NY: Norton.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*, 801–838.

de Zubicaray, G. I., Wilson, S. J., McMahon, K. K., & Muthiah, S. (2001). The semantic interference effect in the picture-word paradigm: An event-related fMRI study employing overt responses. *Human Brain Mapping*, *14*, 218–227.

Duncan, G. E., & Henson, O. W. (1994). Brain activity patterns in flying, echolocating bats (*Pteronotus parnellii*): Assessed by high resolution autoradiographic imaging with [^{3}H]2-deoxyglucose. *Neuroscience*, *59*, 1051–1070.

Gooler, D. M., & O'Neill, W. E. (1987). Topographic representation of vocal frequency demonstrated by microstimulation of anterior cingulate cortex in the echolocating bat, *Pteronotus parnelli parnelli*. *Journal of Comparative Physiology A*, *161*, 283–294.

Henry, C., Gaillard, R., Volle, E., Chiras, J., Ferrieux, S., Dehaene, S., & Cohen, L. (2005). Brain activations during letter-by-letter reading: A follow-up study. *Neuropsychologia, 43*, 1983–1989.

Humphreys, G. W., & Forde, E. M. E. (2005). Naming a giraffe but not an animal: Basic-level but not superordinate naming in a patient with impaired semantics. *Cognitive Neuropsychology, 22*, 539–558.

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*, 101–144.

Jürgens, U. (1998). Speech evolved from vocalization, not mastication. *Behavioral and Brain Sciences, 21*, 519–520.

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews, 26*, 235–258.

Jürgens, U. (2009). The neural control of vocalization in mammals: A review. *Journal of Voice, 23*, 1–10.

Jürgens, U., & von Cramon, D. (1982). On the role of the anterior cingulate cortex in phonation: A case report. *Brain and Language, 15*, 234–248.

Kan, I. P., & Thompson-Schill, S. L. (2004). Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cognitive, Affective, & Behavioral Neuroscience, 4,* 43–57.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review, 9*, 637–671.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–38.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences, 21*, 499–511.

Maeda, K., Takashi, O., Shinagawa, H., Honda, E., Kurabayashi, T., & Ohyama, K. (2006). Role of the anterior cingulate cortex in volitional swallowing: An electromyographic and functional magnetic resonance imaging study. *Journal of Medical and Dental Sciences, 53*, 149–157.

Maril, A., Wagner, A. D., & Schacter, D. L. (2001). On the tip of the tongue: An event-related fMRI study of semantic retrieval failure and cognitive conflict. *Neuron, 31*, 653–660.

Martin, R. C. (2003). Language processing: Functional organization and neuroanatomical basis. *Annual Review of Psychology, 54*, 55–89.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Newman, J. D. (2007). Neural circuits underlying crying and cry responding in mammals. *Behavioural Brain Research, 182*, 155–165.

Nickels, L. (1997). *Spoken word production and its breakdown in aphasia*. Hove, UK: Psychology Press.

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology, 16*, 935–979.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwarts, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (pp. 1–18). New York, NY: Plenum Press.

Paus, T. (2001). Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nature Reviews Neuroscience, 2*, 417–424.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature, 331*, 585–589.

Petersen, S. E., van Mier, H., Fiez, J. A., & Raichle, M. E. (1998). The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences USA*, *95*, 853–860.

Ploog, D. W. (1992). The evolution of vocal communication. In H. Papoušek, U. Jürgens, & M. Papoušek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 6–30). Cambridge, UK: Cambridge University Press.

Posner, M. I., & Raichle, M. E. (1994). *Images of mind.* New York, NY: W. H. Freeman.

Posner, M. I., & Rothbart, M. K. (2007). *Educating the human brain.* Washington, DC: APA Books.

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A.-M. K., Pardo, J. V., Fox, P. T., & Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, *4*, 8–26.

Rapp, B. (Ed.). (2001). *What deficits reveal about the human mind/brain: A handbook of cognitive neuropsychology*. Philadelphia, PA: Psychology Press.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*, 460–499.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42*, 107–142.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*, 249–284.

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, *110*, 88–125.

Roelofs, A. (2004). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: Comment on Rapp and Goldrick (2000). *Psychological Review*, *111*, 561–572.

Roelofs, A. (2005). The visual-auditory color-word Stroop asymmetry and its time course. *Memory & Cognition, 33*, 1325–1336.

Roelofs, A. (2006). Context effects of pictures and words in naming objects, reading words, and generating simple phrases. *Quarterly Journal of Experimental Psychology*, *59*, 1764–1784.

Roelofs, A. (2007). Attention and gaze control in picture naming, word reading, and word categorizing. *Journal of Memory and Language, 57*, 232–251.

Roelofs, A., & Hagoort, P. (2002). Control of language use: Cognitive modeling of the hemodynamics of Stroop task performance. *Cognitive Brain Research*, *15*, 85–97.

Roelofs, A., van Turennout, M., & Coles, M. G. H. (2006). Anterior cingulate cortex activity can be independent of response conflict in Stroop-like tasks. *Proceedings of the National Academy of Sciences USA*, *103*, 13884–13889.

Rubens, A. B. (1975). Aphasia with infarction in the territory of the anterior cerebral artery. *Cortex*, *11*, 239–250.

Schall, J. D., & Boucher, L. (2007). Executive control of gaze by the frontal lobes. *Cognitive, Affective, & Behavioral Neuroscience*, *7*, 396–412.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge, UK: Cambridge University Press.

Snyder, A. Z., Abdullaev, Y. G., Posner, M. I., & Raichle, M. E. (1995). Scalp electrical potentials reflect regional cerebral blood flow responses during processing of written words. *Proceedings of the National Academy of Sciences USA*, *92*, 1689–1693.

Sohn, M.-H., Albert, M. V., Jung, K., Carter, C. S., & Anderson, J. R. (2007). Anticipation of conflict monitoring in the anterior cingulate cortex and the prefrontal cortex. *Proceedings of the National Academy of Sciences USA*, *104*, 10330–10334.

Taub, E., Uswatte, G., & Elbert, T. (2002). New treatments in neurorehabilitation founded on basic research. *Nature Review Neuroscience*, *3*, 228–236.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences USA*, *94*, 14792–14797.

Wallis, J. D., & Miller, E. K. (2003). From rule to response: Neuronal processes in the premotor and prefrontal cortex. *Journal of Neurophysiology*, *90*, 1790–1806.

Wallis, J. D., Anderson, K. C., & Miller, E. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, *411*, 953–956.

Watanabe, Y., Abe, S., Ishikawa, T., Yamada, Y., & Yamane, G. (2004). Cortical regulation during the early stage of initiation of voluntary swallowing in humans. *Dysphagia*, *19*, 100–108.

Wernicke, C. (1874). *The aphasia symptom-complex: A psychological study on an anatomical basis.* Breslau, Germany: Cohn & Weigert.

Wernicke, C. (1885). Some new studies on aphasia. *Fortschritte der Medizin*, *3*, 824–830.

Wernicke, C. (1886). Some new studies on aphasia. *Fortschritte der Medizin*, *4*, 371–377, 463–469.

Wundt, W. (1900). *Language*. Leipzig, Germany: Verlag von Wilhelm Engelmann.

Wundt, W. (1904). *Principles of physiological psychology*. London, UK: Swan Sonnenschein.