

Attention to Spoken Word Planning: Chronometric and Neuroimaging Evidence

Ardi Roelofs*

Nijmegen Institute for Cognition and Information

Abstract

This article reviews chronometric and neuroimaging evidence on attention to spoken word planning, using the WEAVER++ model as theoretical framework. First, chronometric studies on the time to initiate vocal responding and gaze shifting suggest that spoken word planning may require some attention, even though it is one of our most highly practiced psychomotor skills. Second, neuroimaging studies that localized brain activity during word planning suggest that attention enhances activity in perisylvian cortical areas. What word to enhance and for how long is determined by an executive control system located in anterior cingulate and lateral prefrontal cortex. Commonalities of anterior cingulate function across verbal vocalizations (speech) and nonverbal vocalizations (e.g. human crying, laughter, and monkey calls) are outlined.

Introduction

Attention and language play a key role in everyday life. Attention allows our perceptions, thoughts, and actions to be accurate and speeded in the face of distraction (e.g. Posner 1978; Posner & Raichle 1994; LaBerge 1995), whereas language is our most important means of symbolic communication (e.g. Deacon 1997). Attention and spoken language production are intimately related. On the one hand, producing spoken language usually aims at directing the attention of a listener to the communicative intention of the speaker (Clark 1996). On the other hand, as with other voluntary activities, the intentional production of language cannot happen without paying some attention (Levelt et al. 1999). The amount of attention required is presumably not much, given that producing words is one of our most highly exercised psychomotor skills. With 1-h talking per day, we produce some four million words per year (Roelofs 2002), which seems to happen automatically and effortlessly. Still, speaking may require some attention, as is perhaps most evident from the effort associated with talking while driving a car in heavy traffic or talking in a foreign language. Recent behavioral studies have experimentally confirmed that there may be detrimental consequences of speaking for other

highly practiced activities, such as driving a car (Kubose et al. 2006; Levy et al. 2006; Strayer et al. 2006). The amount of attention we have to pay to language production thus sheds light on the limits of automaticity in general.

The word 'attention' is an umbrella term covering many aspects of self-regulation (e.g. Pashler 1998; Posner 2004; Styles 2006). In an influential approach, Posner and colleagues (Posner and Raichle 1994; Posner and Rothbart 2007) distinguish between three central aspects of attention, each associated with separate networks of brain areas: alerting, orienting, and executive control. The alerting network is involved in achieving and maintaining an alert state, and is associated with the locus coeruleus in the midbrain and regions of the frontal and parietal cortices. The orienting network is involved in directing processing towards a location in space by overtly shifting gaze or covertly shifting the focus of processing while keeping the eyes fixed. Orienting is associated with the superior colliculus, pulvinar, right parietal cortex, and frontal eye fields. The executive control network is involved in selection among competitors and controlled memory retrieval, among other activities, and is associated with the anterior cingulate cortex (ACC), lateral prefrontal cortex (LPFC), and basal ganglia. It is assumed that the executive control network also regulates the orienting network. The present article concentrates on the executive control aspect of attention, and briefly addresses overt orienting of attention (gaze shifting).

Being able to pay attention to language production, such as in naming objects, has two major benefits for a speaker: accuracy and speed (Roelofs 1992, 2007). Accuracy in object naming is ordinarily not a problem when the object is the only item in the visual field and has a single good name. Orienting to the object may then achieve much of the required selectivity. Difficulties arise when other objects are in the neighborhood or several good names exist for the object. Also, the accurate planning and execution of naming may run into difficulties when two or more naming responses must be coordinated. Executive control is then required to prevent cross-talk between alternative responses. It may also increase the speed with which word planning takes place. When a word is required, the executive control system seems to elevate activity in brain areas that code the particular word. These orienting and executive control aspects of attention to word planning are discussed in the present article. Attention to word planning also includes self-monitoring, which involves assessing whether planning and articulation are consistent with intent. Self-monitoring is not addressed here (for recent discussions, see Özdemir et al. 2007 and Roelofs 2004).

The article reviews chronometric and neuroimaging evidence on attention to spoken word planning. First, chronometric studies on the time to initiate vocal responding and gaze shifting suggest that spoken word planning may require some attention, even though it is one of our most highly practiced psychomotor skills. Attention is not only required by

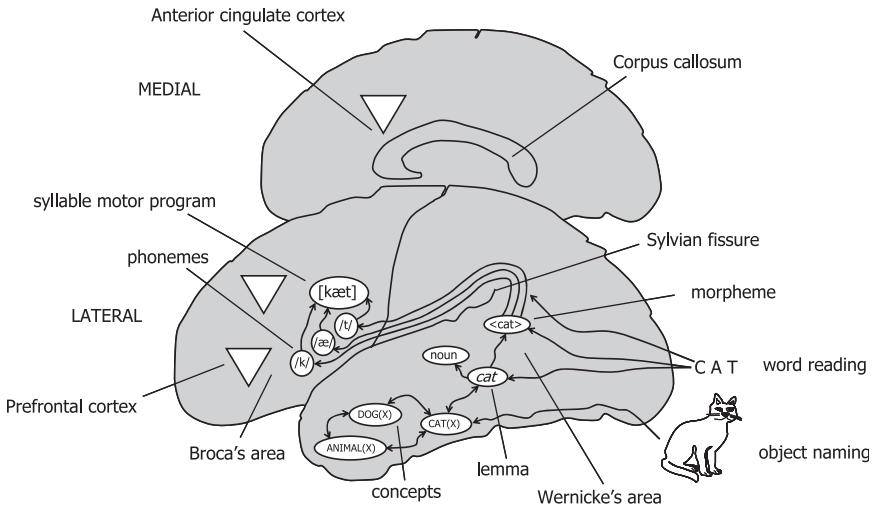


Fig. 1. Cortical anatomy of attention to word planning illustrated by medial and lateral views of the human brain. Representations of concepts [e.g. CAT(X)], lemmas (e.g. *cat* specifying that the word is a noun), morphemes (e.g. <cat>), phonemes (e.g. /k/, /æ/, and /t/), and syllable motor programs (e.g. [kæt]) are located in a left-lateralized perisylvian network of brain areas. An executive control system located in anterior cingulate cortex and in ventro- and dorsolateral prefrontal cortex (indicated by triangles) exerts attentional influences on the lexical network.

conceptualizing what to say, presumably the least automatized aspect of speaking, but also by lexical selection and form-encoding processes. Second, neuroimaging studies that localized brain activity provide evidence on the cortical brain regions involved when speakers pay attention to word planning. These regions are perisylvian areas of spoken word planning and frontal areas of control (Roelofs and Hagoort 2002; Roelofs 2003; Roelofs et al. 2006), illustrated in Figure 1. Attention to spoken word planning seems to enhance the activation flow in cortical pathways corresponding to the target word (e.g. Raichle et al. 1994; Snyder et al. 1995; de Zubicaray et al. 2001). Attentional enhancements occur even when an object is presented in the absence of neighboring distractors, that is, the same system properties that underlie selectivity when visual distractors are present (e.g. naming a pictured cat with the word DOG superimposed, for example, de Zubicaray et al. 2001) also underlie the attentional enhancement of the target when no visual distractors are present (e.g. Raichle et al. 1994; Snyder et al. 1995). What word to enhance and for how long seems to be determined by an executive control system located in the ACC and LPFC. The LPFC is situated off the sides of the frontal lobes of the brain, while the ACC lies in the center, as illustrated in Figure 1. The cingulate cortex surrounds the corpus callosum on the upper side. The implication of the ACC means that executive control over

spoken word production is achieved by a brain area that also occupies the upper rung in the hierarchy of control over the evolutionary forerunner of speech, namely, nonverbal vocalization. The latter includes human crying, laughing, and pain shrieking, as well as monkey calls. The system for controlling verbal vocalizations seems superimposed onto the system for nonverbal vocalizations. The neural pathways along which the ACC exerts its control differ between verbal and nonverbal vocalizations (Deacon 1997; Jürgens 2002).

In their classic article 'Attention to action: willed and automatic control of behavior', Norman and Shallice (1986) made a distinction between 'horizontal threads' and 'vertical threads' in the control of behavior. 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. In the present article, attention to word planning is addressed using the theoretical framework of the *WEAVER++* model (Roelofs 1992, 1997, 2003, 2004, 2007), which computationally implements specific claims about how the horizontal and vertical threads are woven together in planning spoken words. A central claim embodied by *WEAVER++* is that attention to word planning is regulated by condition-action rules. Information about words is stored in a large associative network. *WEAVER++*'s lexical network is accessed by spreading activation while the condition-action rules determine what is done with the activated lexical information depending on the goal. When a goal is placed in working memory, processing in the system is focused on those rules that include the goal among their conditions. The rules mediate attentional influences by selectively enhancing the activation of target nodes in the network in order to achieve mappings of concepts onto articulatory programs. In naming an object in the presence of a visual distractor (e.g. naming a pictured cat with the word *DOG* superimposed), the condition-action rules achieve both output and input control (Roelofs 2003). The rules ensure that the object rather than the distractor is named by selectively enhancing the activation of lexical representations appropriate for naming the object (output control) and by suppressing the distractor input to the lexical network (input control).

I begin by reviewing the functional architecture of the spoken word planning system. Next, I discuss evidence that an executive control system determines the duration of attention to spoken word planning with the aim of speeded and accurate responding. The evidence comes from studies that measured the latencies of vocal responding and gaze shifting in simple dual-task situations involving word production. I then discuss evidence on the brain regions underlying word planning and evidence for a role of the ACC and LPFC in attention to spoken word planning. Commonalities of ACC function across verbal and nonverbal vocalization systems are outlined.

Structure of the Word Planning System

The naming of objects seems to involve the activation of concepts, lemmas, morphemes, phonemes, and syllable motor programs in associative memory (e.g. Levelt et al. 1999), as illustrated in Figure 1. For example, in the WEAVER++ model (Roelofs 1992, 1997, 2003; Levelt et al. 1999), naming a pictured cat involves the activation in a lexical network of the representation of the concept CAT(X), the lemma of *cat* specifying that the word is a noun (for languages such as Dutch, lemmas also specify grammatical gender), the morpheme <cat>, the phonemes /k/, /æ/, and /t/, and the syllable motor program [kæt]. 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 flow from concepts to phonological forms is limited unless attentional enhancements are involved to boost the activation (Roelofs 1992, 2003). This may also hold for similar models of word planning (e.g. Dell 1986; Dell et al. 1997). Some attention theorists (e.g. Posner 1978; Posner and Dehaene 1994; Posner and Raichle 1994) maintain that helping activation to cross a large network distance is an important function of attention in general.

It is often assumed that perceived objects have direct access to concepts, for example, CAT(X), and only indirect access to word forms (e.g. <cat> and /k/, /æ/, and /t/), whereas perceived words have direct access to word forms and only indirect access to concepts (cf. Roelofs 1992, 2003, 2006). Consequently, naming objects requires concept selection, whereas words can be read aloud without concept selection. The latter is achieved by mapping input word forms (e.g. the visual word CAT) directly onto output word forms (e.g. <cat> and /k/, /æ/, and /t/), without engaging concepts and lemmas, as illustrated in Figure 1. With such direct form-to-form mapping, activation has to travel a much shorter network distance from input to output than with a mapping via concepts and lemmas. In word reading through the form-to-form route, target morphemes are enhanced (Roelofs 2003). Given the shorter network distance with reading, the attentional enhancements may be less for word reading than object naming (see Roelofs 2003, p. 125), and successful word reading relies much less on the enhancement than does object naming (Roelofs 2003).

Enhancements of Word Planning in WEAVER++

The activation enhancements in WEAVER++ are 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 (cf. Roelofs 2003, p. 100):

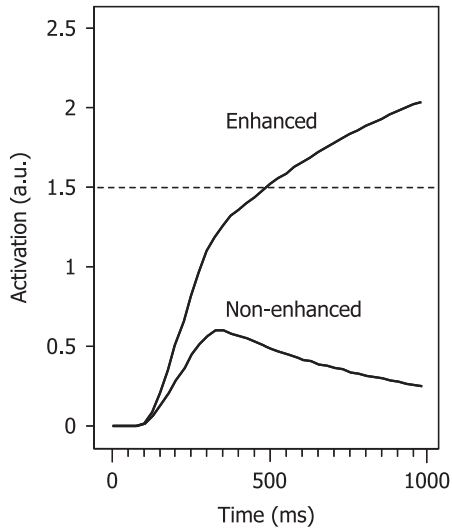


Fig. 2. Illustration of the effect of enhancement in the WEAVER++ model. Shown are the activations over time of the syllable motor program [kæʔ] in response to a pictured cat with and without attentional enhancements. The dashed line indicates the availability threshold. The activation of the motor program exceeds threshold with but not without enhancement. a.u., arbitrary units.

- (1) 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 motor programs have been activated sufficiently, that is, above an availability threshold. The executive 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).

Figure 2 illustrates the need for attentional enhancement in the WEAVER++ model in naming isolated objects. Shown are activations over time of the syllable motor program [kæʔ] in response to a pictured cat. The activation curves were obtained by running the simulations of object naming reported in Roelofs (2004), which are typical for the model. The cascade version of the model was used (Roelofs 2003: 103). To assess the effect of enhancement on word planning in the model, the simulations were run with and without attentional enhancements. The perceived cat provided activation to the network for 200 ms. The horizontal dashed line in Figure 2 indicates the availability threshold for motor programs in the model. In all previous published reports on the model, the threshold was set at 1.5. Figure 2 shows that the activation of the

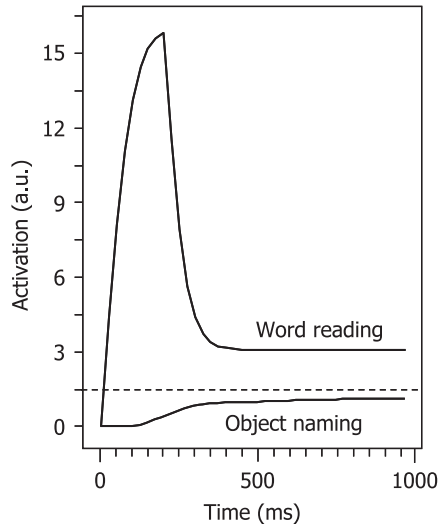


Fig. 3. Illustration of the difference in dependence on the enhancements between object naming and word reading in the WEAVER++ model. Shown are the activations over time of the syllable motor program [kæt] in naming a pictured cat or reading its name. The magnitude of the enhancement was reduced by half relative to the other simulations. The reduced enhancement was still sufficient to make the motor program available for word reading, whereas it was insufficient for object naming. a.u., arbitrary units.

motor program exceeds threshold only with the attentional enhancement. Thus, the enhancement is required for successful word planning in the model. This conclusion may also hold for similar models (e.g. Dell 1986; Dell et al. 1997).

Figure 3 illustrates the difference between object naming and word reading in the dependence on the enhancements in WEAVER++. Shown are the activations over time of the syllable motor program [kæt] in response to a pictured cat (object naming) and its written name (word reading). The magnitude of the enhancement for both tasks was reduced by half relative to the other simulations. Figure 3 shows that a reduced enhancement may still be sufficient to make the motor program available for word reading, whereas the enhancement does not suffice for object naming.

Chronometric Evidence on Attention to Word Planning

The assumption that word planning requires attentional activation enhancements is supported by evidence from response times (see Roelofs, 2003, 2006, for reviews). In a recent study by Roelofs et al. (2007), participants were shown pictured objects (e.g. a cat) while hearing a tone or a spoken word presented 600 ms after picture onset. When a spoken word was presented (e.g. *cup* or *house*), participants indicated whether it contained a prespecified phoneme (e.g. /k/) by pressing a button. When the tone

was presented, they indicated whether the object name contained the phoneme (the first experiment) or they named the object (the second and third experiment). Phoneme monitoring latencies for the spoken words were shorter when the object name contained the prespecified phoneme (e.g. *cat* – *cup*) compared to when it did not (e.g. *cat* – *house*). However, no priming of phoneme monitoring was obtained (the fourth experiment) when the objects required no response, but were only passively viewed. This held not only for the fastest monitoring responses to the spoken words, but also for the slowest ones. For the slowest responses to the words, there should have been sufficient time for activation to spread from concepts to phonemes. Thus, passive object viewing does not lead to significant phonological activation. These results suggest that attentional enhancements are a precondition for obtaining phonological activation from perceived objects.

In the passive viewing condition of Roelofs et al. (2007), speakers may have paid some attention to the object, but apparently not long enough to induce phonological activation. To assess how long attention needs to be sustained to a pictured object, eye movements and response times to the object may be measured. While individuals can shift the focus of attention without an eye movement (covert orienting), they cannot move their eyes to one location while paying full attention to another (i.e. shifts of eye position require shifts in attention). Thus, a gaze shift (overt orienting) indicates a shift of attention. Measurements of gaze shifts have been important in studying attention in infants and children (Posner and Rothbart 2007). Recently, latencies of gaze shifts have also been measured while adult speakers performed language production tasks.

Research on spoken word planning has shown that there is a close link between the duration of word planning and gaze shifts in object naming (e.g. Meyer et al. 1998; Griffin 2001; Korvorst et al. 2006). For example, when speakers are asked to name two spatially separated objects (e.g. one to the left and the other to the right), they look longer at first to-be-named objects with disyllabic names (e.g. *lion*) than with monosyllabic names (e.g. *cat*) even when the object recognition times are the same (Meyer et al. 2003). The effect of the phonological length suggests that the shift of gaze from one object to the other is initiated only after the phonological form of the name for the object has been planned sufficiently and the corresponding articulatory program is available. The executive control system appears to instruct the orienting system to shift gaze depending on phonological encoding. The phonology-dependent gaze shifts may promote naming speed and accuracy by preventing interference from the other object name. Moreover, the phonology-dependent gaze shift may diminish resource consumption. Articulating a word such as ‘cat’ or ‘lion’ can easily take half a second or more. If gaze shifts are initiated after identifying the first object, the planning of the name for the second object may be completed well before articulation of the name for the

first object has been finished. This means that the second vocal response needs to be buffered for a relatively long time. By starting perception of the second object when planning the first object name is completed sufficiently, the use of buffering resources can be limited.

However, the phonology-dependent gaze shifts are also obtained when the second naming response is replaced by a manual response to a left- or right-pointing arrow (Roelofs 2007). That is, gaze shifts still depend on phonological encoding when speakers name an object and manually discriminate an arrow. This finding suggests that the avoidance of response buffering and the prevention of interference from the second response are not the only reasons for a phonology-dependent gaze shift. Rather, some aspect of spoken word planning itself would appear to be the critical factor. If attentional enhancements are required until the word has been planned far enough, as illustrated in Figure 2, this would explain why attention, indexed by eye gazes, is sustained to word planning until the phonological form is planned (Roelofs 2007). This should hold regardless of the need for response buffering and the prevention of interference, as empirically observed (Roelofs 2007).

Evidence suggests that shifts of gaze occur closer to articulation onset in naming objects than in reading their names (Roelofs 2007). Figure 4

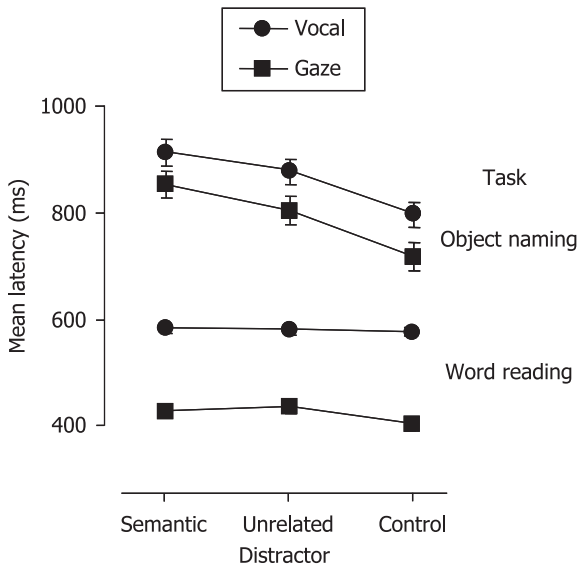


Fig. 4. Chronometric signatures of attention to word planning. Shown are the mean latencies of vocal responding and gaze shifting in object naming and word reading in semantic, unrelated, and control conditions (Roelofs, 2007). The error bars indicate 95% confidence intervals (for reading, the intervals are too small to be visible in the figure). Gaze shifts occurred much earlier in word reading than in object naming relative to the onset of the vocal responses, regardless of whether or not distractors were present.

shows the mean latencies for the vocal responses and gaze shifts in object naming and word reading in a semantic condition (e.g. a pictured cat combined with the word DOG), an unrelated condition (e.g. a pictured cat combined with the word HOUSE), and a control condition (e.g. a pictured cat combined with XXX for object naming or the word CAT in an empty picture frame for word reading). A distractor effect is obtained in object naming but not in word reading, suggesting differences in attentional demands between the two tasks (Roelofs 2003). In all three distractor conditions, the gaze shifts occurred about 66 ms before articulation onset in object naming, whereas they happened already about 156 ms before articulation onset in word reading. Given the shorter network distance for word reading than object naming (see Figure 1), attentional enhancements may be less for reading than naming (see Figure 3). If enhancements are required until the word has been planned sufficiently, this explains why attention, as indexed by eye gazes, is sustained longer to word planning in object naming than in word reading, regardless of whether or not distractors are present.

Evidence suggests that the allocation of attention in dual-task performance is not fixed but strategically determined (cf. Meyer and Kieras 1997). Attention is sustained to a task just as long as is needed to achieve acceptable levels of speed and accuracy (e.g. Roelofs 2007). Recent experiments in my laboratory suggest that whether speakers sustain attention to word planning until the completion of phonological encoding may depend on the nature of the secondary task (for a different view, see Ferreira and Pashler 2002; Dell'Acqua et al. 2007). When speakers name pictured objects (primary task) and manually discriminate arrows or tones (secondary task), phonological encoding for word production delays the manual responses to the arrows but not to the tones. This suggests that speakers shift attention earlier to the tones than to the arrows, presumably because vocal response planning hampers auditory perception (Houde et al. 2002). Consequently, the tone task needs to be protected against interference from speech planning. This might be achieved through attentional enhancement of the processing of the tones. However, there was a cost to the earlier shifts of attention to the tones, namely, an increase in object naming errors.

To conclude, chronometric evidence on object naming and word reading in simple dual-task situations suggests that spoken word planning may require some attention. Moreover, it seems that attention may to some degree flexibly be allocated to word planning to promote word planning speed and accuracy.

Neuroimaging Evidence on Attention to Word Planning

In the remainder of this article, I review neuroimaging evidence on the brain areas realizing the attention to word planning. Hereby, it is important

that the attentional influences in cortical pathways of word planning be distinguished from the brain structures that exert control over those influences. I first discuss the areas underlying word planning and then those implicated in executive control.

In the human brain, the activation of representations underlying object naming proceeds from percepts in posterior cortical areas to articulatory programs in anterior areas, as illustrated in Figure 1. A meta-analysis of 82 neuroimaging studies on word production by Indefrey and Levelt (2004) suggested that the following cortical areas are involved. Information on the time course of word production in relation to these areas came from magnetoencephalographic neuroimaging studies. The meta-analysis included object naming (e.g. say 'cat' to a pictured cat), word generation (producing a use for a noun, for example, say 'stroke' to the word CAT), word reading (e.g. say 'cat' to the word CAT), and pseudoword reading (e.g. say 'caz' to CAZ). 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, motor programs are accessed. The sensorimotor areas control articulation. Word reading may be accomplished by activating occipital and inferiotemporal regions for orthographic processing, the areas of Wernicke and Broca for aspects of form encoding, and motor areas for articulation.

Neuroimaging studies on word planning have shown that the ACC and LPFC are activated in a variety of circumstances requiring attention. The ACC and LPFC are more active in word generation (say 'stroke' to the word CAT) when the executive control demands are high than in word reading (say 'cat' to CAT) when the demands are much lower (Petersen et al. 1988; Thompson-Schill et al. 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 et al. 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 and 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 et al. 2001). Frontal areas are also more active in naming objects with semantically related words superimposed (e.g. naming a pictured cat combined with the word DOG) than without word distractors (e.g. a pictured cat combined with XXX), as demonstrated by de Zubizaray et al. (2001). Thus, the neuroimaging evidence suggests that medial and

lateral prefrontal areas exert control over word planning. Along with the increased frontal activity, there is an elevation of activity in perisylvian areas (e.g. Raichle et al. 1994; Snyder et al. 1995; de Zubicaray et al. 2001).

Although both the ACC and LPFC are involved in executive aspects of attention to word planning, the areas seem to play different roles. Much evidence suggests that the dorsolateral prefrontal cortex is involved in maintaining goals in working memory (for a review, see Kane and 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. Wallis et al. 2001; Bunge et al. 2003; Bunge 2004). Moreover, evidence suggests that the ventrolateral prefrontal cortex plays a role in selection among competing response alternatives (Thompson-Schill et al. 1997), the control of memory retrieval, or both (Badre et al. 2005). The ACC seems to exert regulatory influences over these processes, as I discuss next.

The involvement of the ACC in the control of spoken word production may not come as a surprise, because the area also controls nonverbal vocalization, considered by many to be the evolutionary precursor of speech (e.g. Deacon 1997; Jürgens 1998). Nonhuman primates (monkeys and apes) do not have speech but only innate emotional vocalizations, such as fear, aggression, alarm, and contact calls. Evidence suggests that the ACC plays a critical role in the voluntary initiation and suppression of these calls (e.g. Aitken 1981). The area does so by sending regulatory signals to the periaqueductal grey matter and bordering tegmentum in the caudal midbrain. Here, the regulatory signals from the ACC are combined with emotional signals from the amygdala and other areas, linking specific emotional states to the corresponding vocal responses. The motor programs for these responses are embodied by premotor nuclei in the pontine and medullary reticular formation and motor nuclei of the brainstem and spinal cord. The premotor nuclei of the reticular formation coordinate the activity of the motor nuclei controlling the larynx, respiratory apparatus, and supralaryngeal tract. In the human speech system, the motor region of the posterior ventrolateral cortex directly projects onto the brainstem reticular formation and motor nuclei, bypassing the periaqueductal area. Still, the ACC may exert direct regulatory influences over the speech system through its connections with ventrolateral prefrontal, premotor, and motor cortex (Deacon 1997; Paus 2001; Jürgens 2002). 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 dorsolateral prefrontal cortex (DLPFC; Roelofs et al. 2006). According to this view, the ACC plays a regulatory role in both verbal and nonverbal vocalization (cf. Posner & Raichle 1994; Deacon 1997), although through different neural pathways.

In addition to regulating the output of acquired and innate vocalization systems, the ACC seems involved in regulating external influences on these vocalization systems (Posner and Raichle 1994; Roelofs and Hagoort 2002; Roelofs 2003; Posner and Rothbart 2007), that is, input control. Single-cell recordings in monkeys have revealed that the ACC is involved in the suppression of activity in auditory cortex during vocalization (Müller-Preuss et al. 1980; Müller-Preuss and Ploog 1981). The suppression already begins several hundred milliseconds prior to the onset of vocalization (Eliades and Wang 2003). Similarly, functional neuroimaging studies on humans have demonstrated that during self-produced speech, the auditory cortex suppresses its response to acoustic signals (Houde et al. 2002). These results suggest a role for the ACC in directing attention away from sounds in the environment to protect emotional vocalization and the planning and articulation of speech. Other neuroimaging research has suggested that the ACC may also play an active role the other way around, by specifically enhancing the activation of these external sounds or visual information (Benedict et al. 2002; Crottaz-Herbette and Menon 2006).

The proposed role of the ACC in output and input regulation (Posner and Raichle 1994; Roelofs and Hagoort 2002; Posner and Rothbart 2007) is not the only view about ACC function in the literature. According to a prominent alternative view, the ACC is involved in performance monitoring (Miller and Cohen 2001). According to this view, ACC activity reflects the detection of response conflict and acts as a signal that engages executive control processes subserved by LPFC. 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 a pictured cat combined with the word DOG). ACC activity should not differ between congruent trials (e.g. a pictured cat combined with the word CAT) and neutral trials (e.g. a pictured cat combined with XXX), because competing response alternatives are absent on both trial types. In contrast, the regulatory hypothesis (Posner and Raichle 1994; Roelofs and Hagoort 2002; Posner and Rothbart 2007) 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 (*cat*) is already activated by the distractor (CAT) 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 study. Participants were scanned while they were presented with arrow-word combinations. The arrows were pointing to the left or right and the words were LEFT or RIGHT. The participants indicated by a left or right button press the direction denoted by the word. The use of manual instead of vocal responses is justified by a meta-analysis of the existing neuroimaging literature and the results from a new

neuroimaging experiment by Barch et al. (2001), which indicated that Stroop-like tasks activate the same region of the ACC, regardless of whether the response modality is spoken or manual. 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 responding. This conclusion was corroborated by successful WEAVER++ simulations of the chronometric and neuroimaging findings (Roelofs et al. 2006).

Conclusions

This article reviewed chronometric and neuroimaging evidence on attention to spoken word planning. First, chronometric studies on the time to initiate vocal responding and gaze shifting suggest that spoken word planning may require some attention, even though it is one of our most highly practiced psychomotor skills. Second, neuroimaging studies that localized brain activity provide evidence on the cortical regions involved. Attention to word planning enhances activity in perisylvian areas. What word to enhance and for how long is determined by an executive control system located in ACC and LPFC. Commonalities of ACC function across verbal and nonverbal vocalization systems were indicated. Future chronometric and neuroimaging research on attention and word planning will continue to illuminate the functions of the cortical areas in which these planning events occur, the modulatory effects of attention on these planning processes, and the control exerted over these modulations.

Acknowledgements

The preparation of the article was supported by a VICI grant from the Netherlands Organization for Scientific Research.

Short Biography

Ardi Roelofs is a Senior Researcher at the Nijmegen Institute for Cognition and Information and a Research Fellow of the F.C. Donders Centre for

Cognitive Neuroimaging in Nijmegen, the Netherlands. He earned BA, MA, and PhD degrees (all three with highest honor) in Experimental Psychology from Radboud University Nijmegen. After a postdoctoral year in the Department of Brain and Cognitive Sciences at MIT, he has been a Researcher at the Max Planck Institute for Psycholinguistics in Nijmegen and Reader in Cognitive Science at the University of Exeter in England. Currently, he supervises a project called 'Goal-referenced control of verbal and nonverbal actions' at Radboud University Nijmegen. The project uses response time, eye tracking, electrophysiological, hemodynamic neuroimaging, molecular genetic, and computational modeling techniques to study attention and language use.

Notes

* Correspondence address: Ardi Roelofs, Nijmegen Institute for Cognition and Information, Montessorilaan 3, Nijmegen, Gelderland 6525 HR, the Netherlands. E-mail: a.roelofs@nici.ru.nl.

Works Cited

- Aitken, P. G. 1981. Cortical control of conditioned and spontaneous vocal behavior in rhesus monkeys. *Brain and Language* 13.171–84.
- Badre, D., R. A. Poldrack, E. Paré-Blagoev, R. Z. Insler, and A. D. Wagner. 2005. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* 47.907–18.
- Barch, D. M., T. S. Braver, E. Akbudak, T. Conturo, J. Ollinger, and A. Snyder. 2001. Anterior cingulate cortex and response conflict: Effects of response modality and processing domain. *Cerebral Cortex* 11.837–48.
- Benedict, R. H. B., D. W. Shucard, M. P. Santa Maria, J. L. Shucard, J. P. Abara, M. L. Coad, D. Wack, J. Sawusch, and A. Lockwood. 2002. Covert auditory attention generates activation in the rostral/dorsal anterior cingulate cortex. *Journal of Cognitive Neuroscience* 14.637–45.
- Bunge, S. A. 2004. How we use rules to select actions: a review of evidence from cognitive neuroscience. *Cognitive, Affective, and Behavioral Neuroscience* 4.564–79.
- Bunge, S. A., I. Kahn, J. D. Wallis, E. K. Miller, and A. D. Wagner. 2003. Neural circuits subserving the retrieval and maintenance of abstract rules. *Journal of Neurophysiology* 90.3419–28.
- Clark, H. H. 1996. *Using language*. Cambridge, UK: Cambridge University Press.
- Crottaz-Herbette, S., and V. Menon. 2006. Where and when the anterior cingulate cortex modulates attentional response: combined fMRI and ERP evidence. *Journal of Cognitive Neuroscience* 18.766–80.
- Deacon, T. W. 1997. *The symbolic species: The co-evolution of language and the brain*. New York, NY: Norton.
- Dell, G. S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93.283–321.
- Dell, G. S., M. F. Schwartz, N. Martin, E. M. Saffran, and D. A. Gagnon. 1997. Lexical access in aphasic and nonaphasic speakers. *Psychological Review* 104.801–38.
- Dell'Acqua, R., R. Job, F. Peressotti, and A. Pascali. 2007. The picture-word interference effect is not a Stroop effect. *Psychonomic Bulletin and Review* 14.717–22.
- Eliades, S. J., and X. Wang. 2003. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *Journal of Neurophysiology* 89.2194–207.
- Ferreira, V., and H. Pashler. 2002. Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28.1187–99.

- Griffin, Z. M. 2001. Gaze durations during speech reflect word selection and phonological encoding. *Cognition* 82.B1–B14.
- Houde, J. F., S. S. Nagarajan, K. Sekihara, and M. M. Merzenich. 2002. Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience* 14.1125–38.
- Indefrey, P., and W. J. M. Levelt. 2004. The spatial and temporal signatures of word production components. *Cognition* 92.101–44.
- Jürgens, U. 1998. Speech evolved from vocalization, not mastication. *Behavioral and Brain Sciences* 21.519–20.
- . 2002. Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews* 26.235–58.
- Kan, I. P., and S. L. Thompson-Schill. 2004. Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cognitive, Affective, and Behavioral Neuroscience* 4.43–57.
- Kane, M. J., and R. W. Engle. 2002. The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. *Psychonomic Bulletin and Review* 9.637–71.
- Korvorst, M., A. Roelofs, and W. J. M. Levelt. 2006. Incrementality in naming and reading complex numerals: evidence from eyetracking. *Quarterly Journal of Experimental Psychology* 59.296–311.
- Kubose, T. T., K. Bock, G. S. Dell, S. M. Garnsey, A. F. Kramer, and J. Mayhugh. 2006. The effects of speech production and speech comprehension on simulated driving performance. *Applied Cognitive Psychology* 20.43–63.
- LaBerge, D. 1995. *Attentional processing: the brain's art of mindfulness*. Cambridge, MA: Harvard University Press.
- Levelt, W. J. M., A. Roelofs, and A. S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22.1–38.
- Levy, J., H. Pashler, and E. Boer. 2006. Central interference in driving: Is there any stopping the psychological refractory period? *Psychological Science* 17.228–35.
- Maril, A., A. D. Wagner, and D. L. Schacter. 2001. On the tip of the tongue: An event-related fMRI study of semantic retrieval failure and cognitive conflict. *Neuron* 31.653–60.
- Meyer, A. S., A. Roelofs, and W. J. M. Levelt. 2003. Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language* 48.131–47.
- Meyer, A. S., A. M. Sleiderink, and W. J. M. Levelt. 1998. Viewing and naming objects. *Cognition* 66.B25–33.
- Meyer, D. E., and D. E. Kieras. 1997. A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review* 104.3–65.
- Miller, E. K., and J. D. Cohen. 2001. An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience* 24.167–202.
- Müller-Preuss, P., J. D. Newman, and U. Jürgens. 1980. Anatomical and physiological evidence for a relationship between the 'cingular' vocalization area and the auditory cortex in the squirrel monkey. *Brain Research* 202.307–15.
- Müller-Preuss, P., and D. Ploog. 1981. Inhibition of auditory cortical neurons during phonation. *Brain Research* 215.61–76.
- Norman, D. A., and T. Shallice. 1986. Attention to action: Willed and automatic control of behavior. *Consciousness and self-regulation: advances in research and theory*, ed. by R. J. Davidson, G. E. Schwartz and D. Shapiro, 1–18. New York, NY: Plenum Press.
- Özdemir, R., A. Roelofs, and W. J. M. Levelt. 2007. Perceptual uniqueness point effects in monitoring internal speech. *Cognition* 105.457–65.
- Pashler, H. E. 1998. *The psychology of attention*. Cambridge, MA: MIT Press.
- Paus, T. 2001. Primate anterior cingulate cortex: Where motor control, drive and cognition interface. *Nature Reviews Neuroscience* 2.417–24.
- Petersen, S. E., P. T. Fox, M. I. Posner, M. Mintun, and M. E. Raichle. 1988. Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature* 331.585–9.

- Petersen, S. E., H. van Mier, J. A. Fiez, and M. E. Raichle. 1998. The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences, USA* 95.853–60.
- Posner, M. I. 1978. *Chronometric explorations of mind*. Hillsdale, NJ: Erlbaum.
- . (ed.) 2004. *Cognitive neuroscience of attention*. New York, NY: Guilford.
- Posner, M. I., and S. Dehaene. 1994. Attentional networks. *Trends in Neurosciences* 17.75–9.
- Posner, M. I., and M. E. Raichle. 1994. *Images of mind*. New York, NY: W. H. Freeman.
- Posner, M. I., and M. K. Rothbart. 2007. *Educating the human brain*. Washington, DC: APA Books.
- Raichle, M. E., J. A. Fiez, T. O. Videen, A.-M. K. MacLeod, J. V. Pardo, P. T. Fox, and S. E. Petersen. 1994. Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex* 4.8–26.
- Roelofs, A. 1992. A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42.107–42.
- . 1997. The WEAVER model of word-form encoding in speech production. *Cognition* 64.249–84.
- . 2002. Storage and computation in spoken word production. Storage and computation in the language faculty, ed. by S. Nootboom, F. Weerman and F. Wijnen, 183–216. Dordrecht, The Netherlands: Kluwer.
- . 2003. Goal-referenced selection of verbal action: modeling attentional control in the Stroop task. *Psychological Review* 110.88–125.
- . 2004. Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: Comment on Rapp and Goldrick (2000). *Psychological Review* 111.561–72.
- . 2006. Context effects of pictures and words in naming objects, reading words, and generating simple phrases. *Quarterly Journal of Experimental Psychology* 59.1764–84.
- . 2007. Attention and gaze control in picture naming, word reading, and word categorizing. *Journal of Memory and Language* 57.232–51.
- Roelofs, A., and P. Hagoort. 2002. Control of language use: Cognitive modeling of the hemodynamics of Stroop task performance. *Cognitive Brain Research* 15.85–97.
- Roelofs, A., R. Özdemir, and W. J. M. Levelt. 2007. Influences of spoken word planning on speech recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33.900–13.
- Roelofs, A., M. van Turenout, and M. G. H. Coles. 2006. Anterior cingulate cortex activity can be independent of response conflict in Stroop-like tasks. *Proceedings of the National Academy of Sciences, USA* 103.13884–9.
- Snyder, A. Z., Y. G. Abdullaev, M. I. Posner, and M. E. Raichle. 1995. Scalp electrical potentials reflect regional cerebral blood flow responses during processing of written words. *Proceedings of the National Academy of Sciences, USA* 92.1689–93.
- Strayer, D. L., F. A. Drews, and D. J. Crouch. 2006. A comparison of the cell phone driver and the drunk driver. *Human Factors* 48.381–91.
- Styles, E. A. 2006. *The psychology of attention*. New York, NY: Psychology Press.
- Thompson-Schill, S. L., M. D'Esposito, G. K. Aguirre, and M. J. Farah. 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences, USA* 94.14792–7.
- Wallis, J. D., K. C. Anderson, and E. Miller. 2001. Single neurons in prefrontal cortex encode abstract rules. *Nature* 411.953–6.
- de Zubicaray, G. I., S. J. Wilson, K. K. McMahon, and S. Muthiah. 2001. The semantic interference effect in the picture-word paradigm: an event-related fMRI study employing overt responses. *Human Brain Mapping* 14.218–27.