

Cognitive control during sentence generation

Malathi Thothathiri, Michelle Rattinger & Bhairvi Trivedi

To cite this article: Malathi Thothathiri, Michelle Rattinger & Bhairvi Trivedi (2015): Cognitive control during sentence generation, Cognitive Neuroscience, DOI: [10.1080/17588928.2015.1090421](https://doi.org/10.1080/17588928.2015.1090421)

To link to this article: <http://dx.doi.org/10.1080/17588928.2015.1090421>



Accepted author version posted online: 01 Oct 2015.



Submit your article to this journal [↗](#)



Article views: 9



View related articles [↗](#)



View Crossmark data [↗](#)

Publisher: Taylor & Francis

Journal: *Cognitive Neuroscience*

DOI: 10.1080/17588928.2015.1090421

Cognitive control during sentence generation

Malathi Thothathiri

Michelle Rattinger

Bhairvi Trivedi

Address correspondence to:

Malathi Thothathiri

Department of Speech and Hearing Science

Hall of Government, Room 203

The George Washington University

Washington, D.C.

malathi@gwu.edu

Abstract

Sentence generation involves selecting the right words out of multiple co-activated representations for the right utterance positions. This selection process is hypothesized to involve cognitive control. Previous research has linked cognitive control during language processing to regions within the frontal lobe, including the anterior cingulate cortex (ACC) and left ventrolateral prefrontal cortex (VLPFC). We examined the neural correlates of word and sentence generation to determine whether these cognitive control regions were activated more for sentences than words. We also manipulated the phonological overlap between nouns in the sentence to determine whether cognitive control recruitment increased under phonological competition. Whole-brain analysis revealed increased activation in the ACC and bilateral frontal cortex for sentences over words. Region of interest analysis revealed an effect of phonological competition within the ACC. Psychophysiological interaction (PPI) analysis further showed that functional connectivity within a task maintenance network increased under phonological competition. These results suggest a role for conflict detection and monitoring during sentence generation, consistent with a theoretical framework incorporating domain-general cognitive control processes into language processing.

Keywords: sentence generation; cognitive control; phonology; anterior cingulate cortex; prefrontal cortex; fMRI

Acknowledgments

We thank Katherine O'Connor for assistance with stimuli, and John VanMeter and Manish Khatiwada for scanning help. This research was partly funded via a university facilitating fund from the George Washington University to the first author.

Introduction

Sentence generation involves additional processes beyond word generation. While syntax is an obvious differentiating factor, recent theoretical frameworks have also explored the role of more general cognitive control functions, which may be important for selecting between and integrating multiple co-activated linguistic representations (Hagoort & Indefrey, 2014; Novick, Trueswell, & Thompson-Schill, 2010). Such functions may be even more relevant when the co-activated representations compete or conflict with one another (see e.g., Thothathiri, Kim, Trueswell, & Thompson-Schill, 2012; Thothathiri, Schwartz, & Thompson-Schill, 2010). Additionally, during language production, cognitive control might be important for self-monitoring in order to ensure that speech matches what the speaker intends to say (Nozari, Dell, & Schwartz, 2011; Piai, Roelofs, Acheson, & Takashima, 2013). In this study, we examined whether sentence generation recruits cognitive control regions more than word generation, and whether this recruitment is heightened under phonological competition.

Behavioral research suggests that speakers plan ahead when producing a phrase or sentence, i.e. they access multiple phonological representations and plan their articulation rather than proceed solely incrementally or one word at a time (Costa & Caramazza, 2002;

Schnur, 2011). A consequence of planning ahead during sentence production is that interference between different phonological representations must be resolved in order to select the right representation at the right time. The presence of phonological interference during multi-word production is readily evident in paradigms that manipulate phonological overlap between words. Several previous studies have found that *onset* overlap, in particular, leads to interference (Jaeger, Furth, & Hilliard, 2012a; 2012b; O'Seaghdha & Marin, 2000). For instance, speakers are slower to produce a target word (e.g., "story") after an onset-related prime (e.g., "storage") than after an unrelated prime (e.g., "collar") (O'Seaghdha & Marin, 2000). Similar inhibitory effects are also observed in sentence production tasks. When describing visual scenes, speakers tend to avoid onset overlap between adjacent words and opt for other options instead (e.g., avoid "Hannah handed..." and use "Hannah gave...") (Jaeger et al., 2012a). When an utterance does contain onset-overlapping words, it is more likely to include disfluencies, such as pauses and restarts, than if there is no overlap (Jaeger et al., 2012b). These results suggest that sentence-level planning might require interference resolution and more so in cases where the activated representations overlap phonologically.

The involvement of frontally mediated cognitive control during single word generation is well documented. For example, in the verb generation task, subjects are presented with a noun stimulus and asked to generate a semantically related verb. Left ventrolateral prefrontal cortex (VLPFC) is activated more when the nouns have many associated verbs (e.g., ball: kick, throw, hit) compared to a single associated verb (scissors: cut) (Nelson, Reuter-Lorenz, Persson, Sylvester & Jonides, 2009; Snyder, Banich, & Munakata, 2011). This is consistent with the proposal that this frontal region is involved in selecting between competing alternatives (Novick

et al., 2010). Neuropsychological evidence further suggests that damage to this region impairs selection in patients (Thompson-Schill, Swick, Farah, D'Esposito, Kan, & Knight, 1998). Other tasks suggest the involvement of the anterior cingulate cortex (ACC). For example, in the picture-word interference task, participants are asked to name a picture and ignore a superimposed word that is either semantically related or unrelated. The ACC shows greater activation for the related versus the unrelated condition (Piai et al., 2013. See also de Zubicaray, McMahon, Eastburn, & Pringle, 2006; Hirschfeld, Jansma, Bolte, & Zwitterlood, 2008). Similarly, in the vocal Stroop task, the ACC is recruited when participants have to name aloud the ink color while ignoring the competing word representation (Piai et al., 2013). These results suggest that regions implicated in conflict detection, monitoring, and resolution might be involved in selecting words during speech (see also Botvinick, Cohen & Carter, 2004; Schnur, Schwartz, Kimberg, Hirshorn, Coslett, & Thompson-Schill, 2009).

Neuroimaging research at the sentence level has focused more on comprehension than production. For comprehension, some studies have associated the left VLPFC with syntactic operations (Friederici, 2012; Segaert, Menenti, Weber, Petersson, & Hagoort, 2012). However, others have linked this region to more general cognitive control by examining conflict resolution during sentence processing (January, Trueswell, & Thompson-Schill, 2009; Thothathiri, Kim, et al., 2012). For example, interpretation of sentences containing conflict between syntactic and semantic cues (e.g., *The journalist was interviewed by the undergraduate*) recruits left VLPFC and this recruitment varies parametrically with conflict rather than syntactic demands (Thothathiri, Kim, et al., 2012). For production, Geranmeyeh and colleagues compared free-form descriptive speech to counting, nonverbal decision-making, and

rest (Geranmayeh, Wise, Mehta, & Leech, 2014). Multivariate analyses identified distinct networks for different tasks. A left-lateralized fronto-temporo-parietal (FTP) network, including VLPFC, paracingulate cortex and superior parietal lobule was more active during descriptive speech than during counting and decision-making. In contrast, a different cingulo-opercular network comprised of ACC, bilateral frontal operculum, bilateral frontal poles and visual and motor areas was more active during decision-making than during speech. However, this contradicts the results from single word production described above as well as evidence from monolinguals and bilinguals, which suggests that the ACC is a common locus for cognitive control during verbal and non-verbal tasks (Abutalebi et al., 2011).

Neuropsychological studies of language production also implicate cognitive control in spontaneous speech and multiword production. For example, Robinson and colleagues have argued that production impairments in patients with dynamic aphasia arise from their inability to select between competing verbal response options (Robinson, Shallice, & Cipolotti, 2005). Patients with damage to VLPFC and severely reduced spontaneous speech showed impairments when asked to generate a verbal response under conditions of high but not low competition (Robinson et al., 2005). Similarly, Schwartz and Hodgson (2002) reported that patient MP, who was diagnosed with non-fluent aphasia, struggled to produce words in the context of other words, despite the fact that the syntactic demands of the tasks were minimal. The authors suggested that this deficit arose from an inability to resolve interference when multiple words are co-activated (Schwartz & Hodgson, 2002). Thothathiri et al. (2010) extended this framework to a multiword priming task. Left VLPFC patients were asked to name two pictures using simple “<name> and <name>” phrases. The authors manipulated “selection for position” demands by

priming one of the two nouns in one phrasal position and subsequently requiring patients to either produce the noun in the same position or switch it to a different position. Patients with damage to posterior VLPFC (BA 44/6) in particular showed impairment in flexibly switching nouns between different positions. In a subsequent study, the same patients were found to be impaired in a sequence reproduction task when the stimuli were letters but not when the stimuli were colors (Thothathiri, Gagliardi, & Schwartz, 2012). The authors speculated that this pattern might arise if left BA 44/6, and more broadly posterior left VLPFC, were specialized for phonological but not semantic selection. Such specialization could explain why patients with BA 44/6 damage might struggle with flexibly sequencing phonological, but not semantic, representations. This subdivision is consistent with other findings suggesting a posterior frontal locus for processing phonology (Hartwigsen, Price, Baumgaertner, Geiss, Koehnke, Ulmer, & Siebner, 2010; McDermott, Petersen, Watson, & Ojemann, 2003).

To summarize, previous neuroimaging research has suggested that single word generation recruits cognitive control processes subserved by the left VLPFC and the ACC, and that connected speech recruits cognitive control processes subserved by a left-lateralized FTP network (which includes the left VLPFC) (de Zubicaray et al., 2006; Geranmayeh et al., 2014; Hirschfeld et al., 2008; Piai et al., 2013). Neuropsychological results from patients with aphasia converge in implicating the left VLPFC in the generation of multiword utterances (Robinson et al., 2005; Thothathiri et al., 2010).¹ The current study sought to answer the following questions using functional neuroimaging in healthy adults: (1) whether sentence generation recruits

¹ The role of the ACC has not been investigated much in this population because lesions to the ACC are not common in left MCA strokes.

cognitive control regions more than word generation; (2) whether both lateral and medial cognitive control regions are implicated; and (3) which cognitive control regions and networks show increased recruitment under conditions of phonological competition. Participants viewed visual stimuli and were asked to covertly generate single words or sentences describing the stimuli. Within the sentence generation stimuli, we manipulated phonological overlap between the nouns corresponding to the characters. Based on previous research, we predicted that sentence generation would require additional cognitive control than single word generation, and that phonological overlap between co-activated representations would induce additional demands.

Methods

Participants

Fourteen right-handed native English speakers from the Washington, D.C. area completed the experiment (18-36 years. Mean = 22.13. 11 female) and were paid \$20 for their participation. Participants self-reported handedness and language history, and were screened for MRI safety. None reported previous head injury or psychiatric or neurological disorders. All gave consent under a protocol approved by the George Washington University.

Procedure

Participants were instructed to view scenes and describe them silently. The scenes used in the experimental conditions depicted transitive actions involving two people (Figure 1A, Figure 1B) while those used in the baseline comparison condition depicted a single person (Figure 1C). Thus, the experimental conditions differed from the baseline in requiring the covert production

of a complete sentence versus a single noun. Within the experimental conditions, we manipulated the degree of phonological overlap between the two nouns in the sentence. In the Overlap condition, the two nouns had similar phonological onsets (e.g., “The **surfer** chased the **surgeon**”. Figure 1A); in the Non-Overlap condition, they did not (e.g., “The **surfer** chased the **gymnast**”. Figure 1B). Due to phonological similarity between the nouns, the Overlap condition was expected to induce phonological interference and higher selection demands than the Non-Overlap condition.

During each Overlap or Non-Overlap trial, participants first saw the scene and described it covertly, and then completed a written sentence using a ` . During the manual response phase, participants saw a partially complete sentence describing the scene (e.g., “The _____ chased the surgeon”) at the bottom and two response options at the top (e.g., “baby”, “surfer”). They filled in the blank using the first or the second option by pressing the corresponding button. Participants were instructed to use the left thumb for button 1 and the right thumb for the button 2. In order to encourage covert description of the entire scene and not just sub-parts, the to-be-filled item was the first noun, the verb or the second noun an equal number of times. During the baseline trials, participants first saw the scene containing a single person, described it covertly, and then provided a manual response. During the manual response phase, there were two response options at the top (e.g., “robber”, “clown”). As in the experimental conditions, participants selected the first or the second option by pressing the corresponding button. In all cases, the two response options included the correct answer and a distractor. Across trials, the number of times that each button was expected to be used was counterbalanced (32 button 1, 31 button 2 in each run). Additionally, neither response type

repeated more than 3 times in a row. Participants were instructed to respond as quickly and accurately as possible.

Experimental trials lasted 9 seconds. Participants were asked to view the scene and describe it silently in the first 4.5 seconds and then to fill in the blank for the sentence on the subsequent screen within the next 4.5 seconds. If a response was detected before 4.5 seconds, a fixation cross appeared for the remainder of the trial (see Figures 1A and 1B). Baseline trials lasted 3 seconds. Participants were asked to view the scene and describe it silently in the first 1.5 seconds and then to complete their manual response on the subsequent screen within the next 1.5 seconds. If a response was detected before 1.5 seconds, a fixation screen appeared for the remainder of the trial (see Figure 1C).

Each participant completed 2 runs. Run order was counterbalanced across subjects (7 Run1 first, 7 Run2 first). Within each run, there were 21 trials in each condition (overlap, non-overlap, and baseline) presented in an event-related design. The order of trials and null events was randomized using optseq2 (<http://surfer.nmr.mgh.harvard.edu/optseq/>). Null events showed a fixation cross on a white screen for a variable duration (3, 6 or 9 seconds). The total run duration was 567 seconds.

Prior to entering the scanner, participants completed a short practice session. First, they were familiarized with the characters used in the scenes. They viewed each character (e.g., surfer) and chose the correct noun out of two options (e.g., “surfer”, “ghost”). Next, they were familiarized with the actions depicted in the scenes. They viewed each scene (e.g., robber chasing witch) and chose the correct verb out of two options (e.g., “touched”, “chased”).

Finally, after being familiarized separately with the nouns and the verbs, participants practiced describing the entire scene. They viewed the same scenes as in the verb familiarization above, described it covertly, and then completed a written sentence by filling in the blank. All scenes used during the practice were different from those used in the scanner tasks.

Stimuli

The scene stimuli were drawn in black on a white background using a Wacom Bamboo Fun Tablet. Scenes for the experimental conditions depicted 1 of 6 possible transitive actions (bite, chase, hug, kick, punch, touch) and involved 2 of 14 possible characters (archer, artist, baby, baker, gymnast, gypsy, nun, nurse, priest, prince, shepherd, sheriff, surfer, surgeon). All characters were drawn with neutral expressions. Actions were depicted from left to right. Scenes for the baseline condition involved 1 of 14 possible characters (clown, dancer, diver, fairy, florist, ghost, hippie, hobo, lifeguard, maid, mailman, robber, teacher, witch) drawn standing alone, depicting no transitive action. These characters were chosen such that the corresponding nouns did not share phonological onsets with the critical nouns.

Each of the 14 experimental condition characters (e.g., surfer) appeared 12 times in the visual scenes (6 times as agent of action, 6 times as patient of action. 6 times per run. 6 times each in overlap and non-overlap conditions). Each of the 6 transitive actions appeared 14 times (6 or 8 times per run). However, each combination of action and 2 characters appeared only once during the entire experiment. Each of the 14 baseline condition characters (e.g., clown) appeared 3 times in the visual scenes (1 or 2 times per run).

Scenes for the overlap condition involved characters whose names shared phonological onset (e.g., prince-priest, surfer-surgeon). The two nouns in each overlap pair had the same two or more initial phonemes. The two nouns did not rhyme. We created the corresponding non-overlap stimuli by yoking each noun in the overlap pair with a noun from a different pair (e.g., prince-nurse, surfer-gymnast). In all cases, verbs were paired with nouns such that there was no phonological onset overlap between the verb and the nouns (e.g., “punch” was never paired with “prince” or “priest”). The two nouns within an overlap or non-overlap pair had the same number of syllables (1 or 2). Because we wanted the overlap and non-overlap pairs to differ primarily in the extent of phonological overlap, we evaluated the semantic relatedness of nouns within the two types of pairs. Overlap and non-overlap nouns did not differ in their semantic similarity as computed via latent semantic analysis (LSA; <http://lsa.colorado.edu>). Mean pairwise comparison score = .06 (overlap) and .05 (non-overlap). $F < 1$, $p > .7$).

fMRI Data Acquisition and Analysis

Structural and functional images were acquired using a 3T Siemens Trio scanner at the Center for Functional and Molecular Imaging at Georgetown University. Structural images were acquired using a sagittal T1-weighted MPRAGE sequence (TR=1900 ms, TE=2.52 ms, flip angle=9°, TI=900ms, voxel=1x1x1mm). Functional images were acquired using an echo-planar imaging sequence (TR=3000 ms, TE=30 ms, flip angle=90°, voxel=3x3x3mm).

Images were processed and analyzed using FSL (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012). Non-brain voxels were removed using BET. Images were motion-corrected using MCFLIRT, spatially smoothed using a Gaussian kernel (FWHM=5 mm) and high-pass filtered

(100 Hz). Statistical maps were normalized to MNI-152 space. Functional activation data were corrected for differences in slice timing and analyzed using a general linear model that contained regressors for each event type (overlap, non-overlap and baseline) convolved with a double gamma hemodynamic response function. The covert response generation phase was modeled separately from the manual response phase for each event type. Motion parameters were not included as regressors because motion is likely to be correlated with task-related activity during covert production.

All results are reported for the covert production phase only. Whole-brain Sentence > Word analysis was subjected to cluster-level familywise error correction within FSL ($Z > 2.3$, corrected cluster $p < .05$). Cognitive control related regions of interest (ROIs) were extracted as follows: the ACC ROI (393 voxels. Peak=8,14,36) was obtained by masking suprathreshold clusters from the whole-brain analysis with the anterior cingulate gyrus (Harvard-Oxford atlas probability > 40). The left BA 44/6 ROI (432 voxels. Peak=-44,10,26) was obtained by masking suprathreshold clusters with the intersection of left precentral gyrus and left pars opercularis. Additionally, to examine whether areas outside of the ROIs were recruited by phonological competition, we conducted a supplementary whole-brain Overlap > Non-overlap analysis.

PPI analyses were conducted using the generalized PPI (gPPI) approach (McLaren, Ries, Xu, & Johnson, 2012). Mean time series for the ROI was extracted using `fslmeants`. The general linear model contained psychological regressors for each condition and phase, the physiological regressor, and interaction terms between each psychological regressor and the physiological regressor. A planned contrast of the overlap-physiological interaction and the non-overlap-

physiological interaction was used to assess functional connectivity ($Z > 2.3$, corrected cluster $p < .05$).

Results

Behavioral

Mean accuracies (standard deviation (SD) in parentheses) for the three conditions were: overlap = 95.75% (3.64), non-overlap = 97.45% (4.01), baseline = 98.13% (2.32). A repeated measures ANOVA revealed no effect of condition [$F(2,26)=2.23$, $p > .1$]. Mean reaction time (RT) and SD for the three conditions were: overlap = 1284.00 ms (257.26), non-overlap = 1311.85 ms (284.73), baseline = 716.14 ms (116.41). Repeated-measures ANOVA revealed a significant effect of condition [$F(2,26)=76.09$, $p < .001$]. Planned Bonferroni-corrected comparisons between all pairs revealed significantly faster RTs in the baseline than in the overlap and non-overlap conditions [p 's $< .001$] but no reliable difference between the overlap and non-overlap conditions [$p > .9$].

Sentence > Word Generation

In the whole brain analysis, we examined which regions showed greater activation for sentence over word generation, and whether this list included cognitive control regions. As shown in Table 1, increased activation for the sentence generation conditions compared to the baseline word generation condition was found in bilateral visual cortices, paracingulate/cingulate cortex, posterior left frontal cortex and posterior right frontal cortex.

Overlap > Non-Overlap

From the whole brain analysis, we extracted ROIs implicated in cognitive control and examined whether these ROIs showed additional activation for Overlap over Non-overlap sentence generation. The ACC ROI was extracted from the paracingulate/cingulate cortex cluster and the left BA 44/6 ROI was extracted from the left precentral/prefrontal cortex cluster.² As seen in Figure 2, the ACC ROI showed Overlap > Non-overlap activation [$F(1,13)=6.84$, $p<.05$]. The left BA 44/6 ROI did not [$F(1,13)=.12$, $p>.7$].

The whole brain Overlap > Non-overlap contrast revealed no clusters at the pre-determined $Z>2.3$ threshold. At a more liberal $Z>1.65$ threshold, we detected a single cluster located within the ACC (948 voxels; peak=2 -2 46).

Functional Connectivity

For the ACC ROI, which showed differential activation for the Overlap over the Non-overlap sentence condition, we conducted a PPI analysis to determine which regions showed greater functional connectivity with the ACC during overlap compared to non-overlap. This revealed increased connectivity between the ACC and the left cerebellum, right planum polare, and left frontal pole (Figure 3, Table 2).

Post-hoc Analysis

While the study's design explicitly manipulated phonological overlap, semantic similarity between nouns might also be expected to recruit cognitive control. In a post-hoc analysis we added the mean-centered LSA score as a covariate to the model and extracted mean contrast

² Masking the right precentral/prefrontal cortex cluster from the whole brain analysis with the intersection of right BA 44 and right BA 6 resulted in no voxels.

estimates of this covariate for individual subjects within the above-mentioned BA 44/6 ROI. A one-sample t-test ($\mu_0=0$) of these estimates revealed a significant effect [Mean=113.27; standard error=34.72; $t(13)=3.26$; $p<.01$].³

General Discussion

Generating a multi-word utterance involves resolving interference and selecting between multiple co-activated words. Therefore, we hypothesized that sentence generation would recruit cognitive control. Below we discuss results from the whole-brain, ROI and PPI analyses, which are consistent with this hypothesis.

The whole-brain analysis revealed four clusters—bilateral visual cortices, cingulate cortex, left frontal cortex and right frontal cortex—which were significantly more active for sentence than word generation. Activation in the bilateral visual cortices is likely related to the greater visual complexity of the scenes used in the sentence conditions. Of more direct interest was activation found in the ACC and bilateral frontal cortices. The ACC is widely regarded as a neural substrate of domain-general cognitive control, although there are nuanced debates regarding its specific control function (e.g., See Carter & van Veen, 2007 for a review). For the present purposes, the activation of this region for sentence versus word generation is consistent with the idea that formulating a sentence involves controlled processing of co-activated representations (see more below).

Increased activation in the left frontal cortex for sentences versus words could be accounted for by a number of previously postulated functions, including cognitive control, but

³ Consistent with a cognitive control account, the ACC ROI also showed a significant effect [Estimate=174.82, $p<.01$] while a control low-level visual ROI (suprathreshold occipital pole, probability > 40) did not [Estimate=42.21, $p>.1$].

also alternatives such as working memory, phonological processing and syntactic processing (Baddeley, 2003; Friederici, 2012; McDermott et al., 2003). In fact, it has been suggested that adjacent sub-regions within the left frontal cortex—particularly Broca’s area—might differ in their degree of domain-generalty (Fedorenko, Duncan, & Kanwisher, 2012; 2013). In this context, it is informative to consider the precise anatomical location indicated by our results—near the junction between the inferior frontal and the inferior precentral sulcus. Quantitative meta-analyses link this region—also known as the inferior frontal junction (IFJ)—to cognitive control (Derrfuss, Brass, Neumann, & von Cramon, 2005). The peak location identified for the Stroop task (-40, 4, 32) lies within our left frontal cluster. More broadly, the area surrounding the inferior frontal sulcus is activated for a broad set of non-linguistic tasks, in contrast to ventral regions argued to be more language-specific (Fedorenko et al., 2012; 2013). The BA 44/6 locus also aligns with findings linking lesions to this region with impairments in “selection for position” (Thothathiri et al., 2010). Further, post-hoc analysis showed greater activation within BA 44/6 when nouns in a trial were more semantically similar. These results in conjunction with the anatomical locus support a cognitive control explanation for left frontal activation during the sentence conditions.

We had no a priori expectation regarding right frontal involvement in the experimental task. It is possible that with a cognitive control framework, the left hemisphere is involved in selective linguistic processing and the right hemisphere in selective non-verbal (e.g., visuospatial) processing (Nozari & Thompson-Schill, 2013).

The ROI and the supplementary whole brain analyses indicated increased ACC activation for phonological overlap versus non-overlap sentence generation. This is consistent with a conflict monitoring account (Botvinick et al., 2004). Linguistic competitors during verbal response generation are known to recruit the ACC (de Zubicaray et al., 2006; Piai et al., 2013). Moreover, ERP results show modulation of early signals, hypothesized to originate from the ACC, based on whether the distractors are linguistic or non-linguistic. This suggests that this region may be sensitive to the extent to which the distractors are likely to interfere with the correct response (Hirschfeld et al., 2008). In the current study, the overlap condition contained nouns that were more likely to interfere with one another, which could have triggered conflict monitoring within the ACC. We did not find a similar phonological competition effect in the left VLPFC ROI. It is possible that this region would be more differentially engaged for the overlap and non-overlap conditions at a later stage of production than the one investigated here. Under overt production conditions, phonological competition could trigger cascading effects between regions involved in lexical selection, phonological planning and articulation (Peramunage, Blumstein, Myers, Goldrick, & Baese-Berk, 2011). Thus, activation could span from the dorsal BA 44/6 locus identified here to more anterior/ventral (BA 44/45) and more posterior (BA 6) regions.

The PPI analysis showed increased functional connectivity between ACC and the left cerebellum, right planum polare, and left frontal pole during the overlap versus non-overlap conditions. The cerebellar crus and frontal pole loci are consistent with a cognitive control network identified by intrinsic connectivity (Buckner, Krienen, Castellanos, Diaz, & Yeo, 2011; Yeo et al., 2011). The ACC and the frontal pole are part of a domain-general task-maintenance

network that shows sustained activity throughout a task epoch (Dosenbach et al., 2007). This network's involvement in sentence generation is inconsistent with results suggesting that it may be deactivated during speech (Garanmayeh et al., 2014). The results are more in line with findings from single word production, which suggest ACC involvement under linguistic competition (de Zubicaray et al., 2006; Hirschfeld et al., 2008; Piai et al., 2013), and with the theoretical proposal that the ACC may be involved in self-monitoring during speech (Nozari et al., 2011). The right and left planum polare are usually associated with auditory language comprehension. Thus, right planum polare involvement in the current study was not expected. However, it is worth noting that this region is anatomically close to the anterior operculum and insula, which are a part of the above-mentioned task maintenance network (Dosenbach et al., 2007). Further, there is growing evidence for bilateral coupling between production and comprehension systems (Silbert, Honey, Simony, Poeppel, & Hasson, 2014), which could be related to monitoring one's own speech (Yeo et al., 2011).

One limitation of this study arises from the use of a covert rather than overt sentence generation task to circumvent motion artifacts. However, we ensured that participants encoded all aspects of the scene by including a manual response phase. Further, we found significant activation differences between conditions *during the covert generation phase*. This suggests that participants were engaged in processes during this phase that differed between the conditions, and these differences were in accordance with our a priori hypotheses. A second limitation of the study is that the stimuli tapped lexical selection from a small set of familiarized nouns and verbs. Although we ensured that the scene stimuli did not repeat and that the task required the generation of new rather than memorized responses, the experimental task was

obviously less complex than real-life connected speech. Future studies involving overt production of a wider variety of sentences and words are needed to extend the current findings. Finally, although this study provides evidence that competition between co-activated words leads to cognitive control recruitment during sentence generation, it does not provide information about the particular time points or stages at which different regions are recruited. Techniques with greater temporal resolution are better suited to answer such questions.

To conclude, we examined the involvement of cognitive control during sentence generation using a covert scene description task. The ACC and posterior lateral frontal cortices were activated for sentence compared to word generation. Under phonological competition, ACC showed increased activation and connectivity within a task maintenance network. These findings are consistent with a role for domain-general cognitive control during sentence generation.

References

- Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., ... & Costa, A. (2011). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex*, bhr287.
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829-839.
- Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, 8(12), 539-546.

- Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C., & Yeo, B. T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(5), 2322-2345.
- Carter, C. S., & Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4), 367-379.
- Costa, A., & Caramazza, A. (2002). The production of noun phrases in English and Spanish: Implications for the scope of phonological encoding in speech production. *Journal of Memory and Language*, 46(1), 178-198.
- de Zubicaray, G., McMahon, K., Eastburn, M., & Pringle, A. (2006). Top-down influences on lexical selection during spoken word production: A 4T fMRI investigation of refractory effects in picture naming. *Human Brain Mapping*, 27(11), 864-873.
- Derrfuss, J., Brass, M., Neumann, J., & von Cramon, D. Y. (2005). Involvement of the inferior frontal junction in cognitive control: Meta-analyses of switching and Stroop studies. *Human brain mapping*, 25(1), 22-34.
- Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., ... & Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences*, 104(26), 11073-11078.
- Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Current Biology*, 22(21), 2059-2062.
- Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, 110(41), 16616-16621.

- Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, 16(5), 262-268.
- Geranmayeh, F., Wise, R. J., Mehta, A., & Leech, R. (2014). Overlapping networks engaged during spoken language production and its cognitive control. *Journal of Neuroscience*, 34(26), 8728-8740.
- Hagoort, P., & Indefrey, P. (2014). The neurobiology of language beyond single words. *Annual Review of Neuroscience*, 37, 347-62.
- Hartwigsen, G., Price, C. J., Baumgaertner, A., Geiss, G., Koehnke, M., Ulmer, S., & Siebner, H. R. (2010). The right posterior inferior frontal gyrus contributes to phonological word decisions in the healthy brain: evidence from dual-site TMS. *Neuropsychologia*, 48(10), 3155-3163.
- Hirschfeld, G., Jansma, B., Bölte, J., & Zwitserlood, P. (2008). Interference and facilitation in overt speech production investigated with event-related potentials. *Neuroreport*, 19(12), 1227-1230.
- Jaeger, T. F., Furth, K., & Hilliard, C. (2012a). Phonological overlap affects lexical selection during sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(5), 1439-1449.
- Jaeger, T. F., Furth, K., & Hilliard, C. (2012b). Incremental phonological encoding during unscripted sentence production. *Frontiers in Psychology*, 3, 481.
- January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of stroop and syntactic ambiguity resolution in Broca's area: implications for the neural basis of sentence processing. *Journal of Cognitive Neuroscience*, 21(12), 2434-2444.

- Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). Fsl. *Neuroimage*, 62(2), 782-790.
- McDermott, K. B., Petersen, S. E., Watson, J. M., & Ojemann, J. G. (2003). A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. *Neuropsychologia*, 41(3), 293-303.
- McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. *Neuroimage*, 61(4), 1277-1286.
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, 4(2), 61-64.
- Nelson, J. K., Reuter-Lorenz, P. A., Persson, J., Sylvester, C. Y. C., & Jonides, J. (2009). Mapping interference resolution across task domains: a shared control process in left inferior frontal gyrus. *Brain Research*, 1256, 92-100.
- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass*, 4(10), 906-924.
- Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology*, 63(1), 1-33.
- Nozari, N., & Thompson-Schill, S. L. (2013). More attention when speaking: does it help or does it hurt?. *Neuropsychologia*, 51(13), 2770-2780.

- O'Seaghdha, P. G., & Marin, J. W. (2000). Phonological competition and cooperation in form-related priming: Sequential and nonsequential processes in word production. *Journal of Experimental Psychology: Human Perception and Performance*, 26(1), 57-73.
- Peramunage, D., Blumstein, S. E., Myers, E. B., Goldrick, M., & Baese-Berk, M. (2011). Phonological neighborhood effects in spoken word production: An fMRI study. *Journal of cognitive neuroscience*, 23(3), 593-603.
- Piai, V., Roelofs, A., Acheson, D. J., & Takashima, A. (2013). Attention for speaking: domain-general control from the anterior cingulate cortex in spoken word production. *Frontiers in Human Neuroscience*, 7.
- Robinson, G., Shallice, T., & Cipolotti, L. (2005). A failure of high level verbal response selection in progressive dynamic aphasia. *Cognitive Neuropsychology*, 22(6), 661-694.
- Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences*, 106(1), 322-327.
- Schnur, T. T. (2011). Phonological planning during sentence production: Beyond the verb. *Frontiers in Psychology*, 2, 319. doi:10.3389/fpsyg.2011.00319
- Schwartz, M. F., & Hodgson, C. (2002). A new multiword naming deficit: Evidence and interpretation. *Cognitive Neuropsychology*, 19(3), 263-288.
- Segaert, K., Menenti, L., Weber, K., Petersson, K. M., & Hagoort, P. (2012). Shared syntax in language production and language comprehension—an fMRI study. *Cerebral Cortex*, 22(7), 1662-1670.

- Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Sciences*, 111(43), E4687-E4696.
- Snyder, H. R., Banich, M. T., & Munakata, Y. (2011). Choosing our words: Retrieval and selection processes recruit shared neural substrates in left ventrolateral prefrontal cortex. *Journal of Cognitive Neuroscience*, 23(11), 3470-3482.
- Thompson-Schill, S. L., Swick, D., Farah, M. J., D'Esposito, M., Kan, I. P., & Knight, R. T. (1998). Verb generation in patients with focal frontal lesions: A neuropsychological test of neuroimaging findings. *Proceedings of the National Academy of Sciences*, 95(26), 15855-15860.
- Thothathiri, M., Kim, A., Trueswell, J. C., & Thompson-Schill, S. L. (2012). Parametric effects of syntactic–semantic conflict in Broca's area during sentence processing. *Brain and Language*, 120(3), 259-264.
- Thothathiri, M., Gagliardi, M., & Schwartz, M. F. (2012). Subdivision of frontal cortex mechanisms for language production in aphasia. *Neuropsychologia*, 50(14), 3284-3294.
- Thothathiri, M., Schwartz, M. F., & Thompson-Schill, S. L. (2010). Selection for position: The role of left ventrolateral prefrontal cortex in sequencing language. *Brain and Language*, 113(1), 28-38.
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3), 1125-1165.

Figure 1. Experimental conditions. (A) In the Overlap condition, participants described scenes involving nouns with similar phonological onsets (“The surfer chased the surgeon”). (B) In the Non-Overlap condition, participants described scenes involving nouns with dissimilar phonological onsets (“The surfer chased the gymnast”). (C) In the Baseline condition, participants described scenes involving a single noun (“clown”).

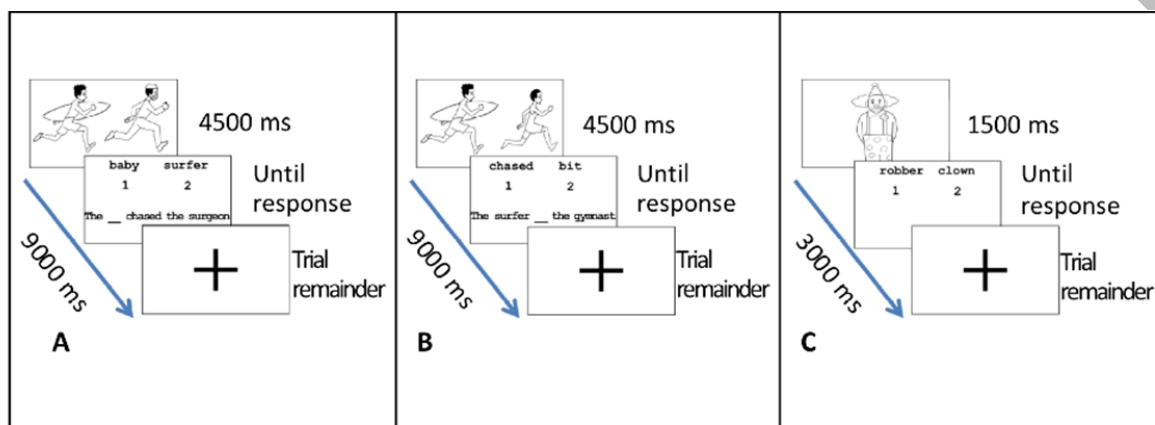


Figure 2: Overlap versus Non-overlap comparison within cognitive control related ROIs. Error bars denote corrected Cousineau confidence intervals (Morey, 2008).

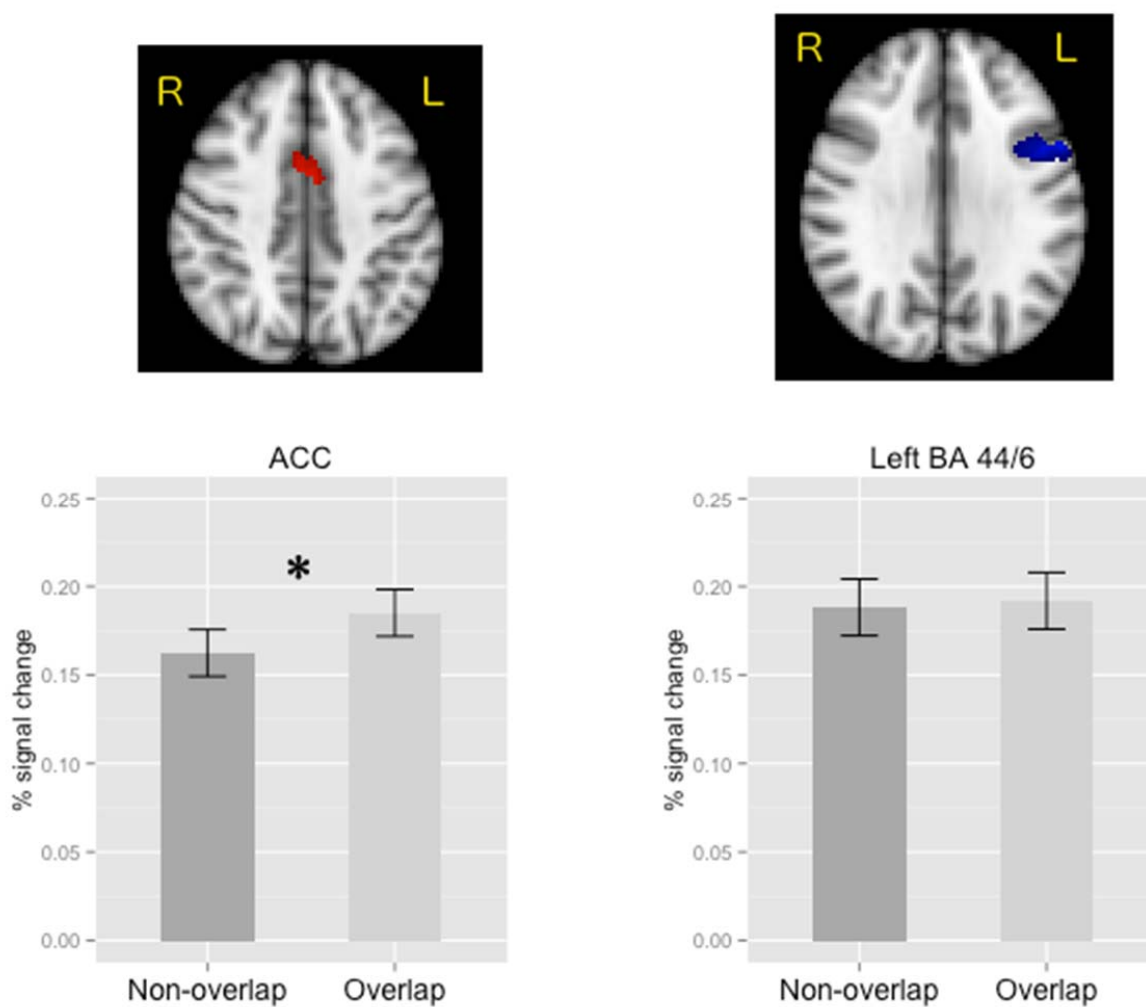


Figure 3: Clusters showing increased functional connectivity with the ACC ROI during overlap compared to non-overlap.

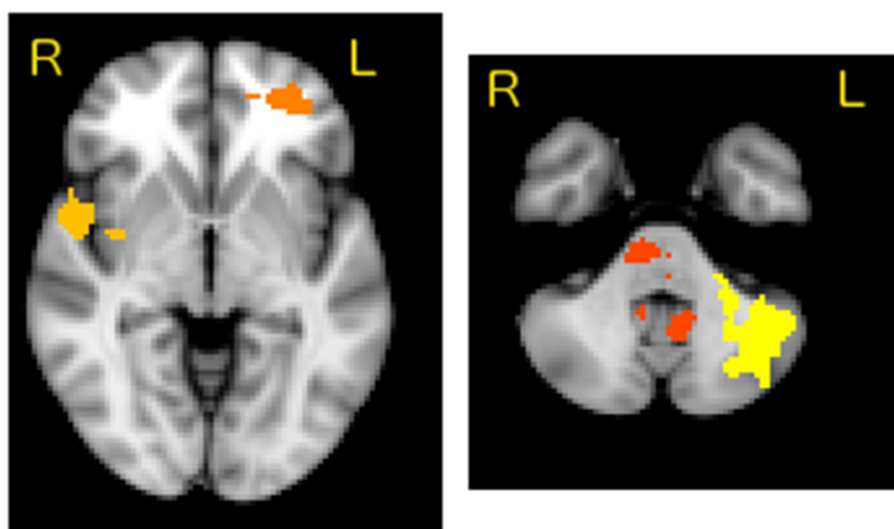


Table 1: Whole brain results. Sentence generation > Word generation

# Voxels in cluster	Region	BA	Max Z	x	y	z
6353	<i>Bilateral Occipital Cortex</i>	7/19	5.1	-18	-78	46
			4.39	-16	-82	50
			4.35	20	-84	42
1510	<i>Paracingulate/Cingulate Cortex</i>	6/24	3.91	10	10	50
			3.72	8	14	36
			3.63	6	8	48
1062	<i>Left Precentral/Prefrontal Cortex</i>	6/44	4.13	-46	0	44
			4.03	-44	10	26
			3.89	-54	-2	42
848	<i>Right Precentral/Prefrontal Cortex</i>	6/44	3.49	40	-6	52
			3.41	28	-6	56
			3.33	54	6	40

Z>2.3, cluster p<.05. 3 local maxima shown within each cluster. BA=Brodman Area

Table 2: PPI results for the ACC ROI

# Voxels in cluster	Region	BA	Max Z	x	y	z
793	<i>Cerebellum (Left Crus I/II)</i>		3.54	-48	-64	-34
			3.41	-32	-62	-40
			3.38	-38	-60	-40
330	<i>Right Planum Polare</i>	22/38	3.69	54	0	-8
			3.46	48	6	-8
			3.39	56	4	-4
301	<i>Left Frontal Pole</i>	47	3.34	-30	48	-2
			3.28	-34	50	0
			3.27	-32	46	2
286	<i>Cerebellum (Left IX)/Brain Stem</i>		3.35	-6	-54	-38
			3.24	-6	-58	-36
			3.05	-8	-40	-44

Z>2.3, cluster p<.05. 3 local maxima shown within each cluster. BA=Brodman Area