

When Elephants Fly: Differential Sensitivity of Right and Left Inferior Frontal Gyri to Discourse and World Knowledge

Laura Menenti¹, Karl Magnus Petersson^{1,2}, René Scheeringa¹, and Peter Hagoort^{1,2}

Abstract

■ Both local discourse and world knowledge are known to influence sentence processing. We investigated how these two sources of information conspire in language comprehension. Two types of critical sentences, correct and world knowledge anomalies, were preceded by either a neutral or a local context. The latter made the world knowledge anomalies more acceptable or plausible. We predicted that the effect of world knowledge anomalies would be weaker for the local context. World knowledge effects have previously been observed in the left inferior frontal

region (Brodmann's area 45/47). In the current study, an effect of world knowledge was present in this region in the neutral context. We also observed an effect in the right inferior frontal gyrus, which was more sensitive to the discourse manipulation than the left inferior frontal gyrus. In addition, the left angular gyrus reacted strongly to the degree of discourse coherence between the context and critical sentence. Overall, both world knowledge and the discourse context affect the process of meaning unification, but do so by recruiting partly different sets of brain areas. ■

INTRODUCTION

“The elephant flies” is a sentence you can immediately understand, although you know that elephants typically do not fly. Maybe you thought of a circus traveling by airplane or maybe you thought of Dumbo. To comprehend a sentence, we rapidly make use of the knowledge we have about the world, and even if what we read or hear does not fit with what we typically take for granted we usually comprehend utterances as long as they are well-formed. Sentence comprehension can be further facilitated if the reader or hearer is provided an adequate context. If we first read the sentence, “Dumbo is a fantasy animal.” then “The elephant flies.” becomes a coherent sentence in its context. The question then is, how understanding of the sentence “The elephant flies” is modulated depending on what we just read or heard. When reading texts or listening to speech, we are constantly integrating the current word meaning with the preceding information. Grasping the neural implementation of contextual integration is an essential part of understanding the neural organization of language comprehension.

Modulations of sentence comprehension by discourse have previously been investigated with various techniques. In an eye-movement study, Duffy and Keir (2004) investigated the influence of disambiguating discourse on

processing of gender stereotypes. Sentences such as “The electrician taught herself. . .” evoke longer reading times of the word “herself” because electricians are stereotypically men and readers therefore take longer time integrating the incoming word “herself.” However, after a disambiguating context explaining that the electrician was a woman, the effect of stereotype anomaly on reading times at the critical word disappeared. Hald, Steenbeek-Planting, and Hagoort (2007) investigated modulation of world knowledge comprehension by discourse context with ERPs. Hagoort, Hald, Bastiaansen, and Petersson (2004) had already established that world knowledge anomalies, just like semantic anomalies, elicit an N400 effect, which indexes ease of semantic integration of content words. Hald et al. (2007) presented world knowledge sentences (correct or anomalies: “Dutch trains are *yellow/white* and blue.”) that were preceded by two types of contexts: (i) neutral contexts introduced the topic of the world knowledge sentence and were in line with the world knowledge of the listener/reader (e.g., “The Netherlands are famous for their designers. In addition, the Dutch railways have chosen a very conspicuous color scheme, which makes them recognizable everywhere.”); (ii) local contexts introduced information that made deviations from the default world knowledge more acceptable (e.g., “The coming world championships are one big national spectacle. The Dutch railways have painted the Dutch flag on their trains.”) They found that the N400 effect to world knowledge anomalies was weaker when the sentences were preceded by the local context,

¹Donders Institute for Brain, Cognition, and Behavior, Nijmegen, the Netherlands, ²Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands

thereby supporting the idea that ongoing discourse can modulate the integration of world knowledge information retrieved from long-term memory. Similarly, Nieuwland and Van Berkum (2006) investigated discourse modulation of animacy anomalies such as, “The peanut was in love.” Such animacy anomalies presented in isolation elicited a clear N400 effect compared to neutral sentences, as “The peanut was salted,” but when these sentences were preceded by a supporting discourse introducing an amorous peanut, the N400 effect not only disappeared, but was reversed: “The peanut was salted” now elicited a larger N400 than “The peanut was in love.” Taken together, this suggests that local discourse can override semantic and world knowledge anomalies and that the anomaly effect is thus inherently context-dependent.

To form a coherent representation of a multiword utterance, the reader must not only retrieve and integrate lexical information and world knowledge that is stored in long-term memory. She also has to form a representation of the ongoing discourse itself. Here it is important to make a distinction between three different types of discourse representation: the surface code, the text base, and the situation model (van Dijk & Kintsch, 1983; see also Graesser, Millis, and Zwaan, 1997, for an extensive and more recent review). The surface code is a literal representation of the exact wording and syntax of the text (“The animal tamer gently patted the elephant. The animal was afraid of entering the plane.”), the text base contains explicit text propositions in a stripped-down format that preserves meaning (e.g., *PAT* [*animal tamer, elephant*], *AFRAID* [*elephant*], *ENTER* [*elephant, plane*]), and the situation model is the content or the conceptual structure of the microworld that the text describes (*Circus traveling by airplane*). In the present article, by discourse representation, we mean the situation model level of representation, which determines how additional incoming information is interpreted.

A recurring issue in fMRI research on discourse comprehension is the involvement of the right hemisphere in language processing and more specifically in discourse processing. Jung-Beeman (2005; and Beeman 1998) proposed a division of labor between the right and left hemispheres in language processing. In this proposal, the right and left hemispheres perform similar functions in language processing but they encode information in a different way, with the right hemisphere performing coarser semantic coding than the left. In other words, it is suggested that semantic fields are more focused in the left hemisphere, whereas they are more diffuse in the right hemisphere. For example, the semantic field for the word *foot* might include *band, leg, toe* in the left hemisphere, but in addition, *bare, step, ball* in the right hemisphere. Reading *foot* will then activate a few related words quite strongly in the left hemisphere, and many words more weakly in the right. In addition, the right hemisphere has been shown to be involved in processing jokes (Bartolo, Benuzzi, Nocetti, Baraldi, & Nichelli, 2006) and metaphors (Stringaris et al.,

2006; but see Stringaris, Medford, Giampietro, Brammer, & David, 2007; see Mitchell & Crow, 2005, for a review). This has been taken to suggest that the right hemisphere is involved in processing nonliteral language. Recent evidence, however, seems to suggest that the division of labor is different, in that the left hemisphere is more sensitive to salient (dominant) meanings and direct semantic relationships, whereas the right hemisphere is more sensitive to nonsalient (novel and inferred) meanings and more remote relationships, irrespective of whether they are literal or not (Giora, 2007; Mashal, Faust, Hendler, & Jung-Beeman, 2007; Schmidt, DeBuse, & Seger, 2007).

A few studies indicate that regions in the right hemisphere are sensitive to discourse coherence. Kuperberg, Lakshmanan, Caplan, and Holcomb (2006) found that the right inferior frontal gyrus (RIFG) showed stronger responses to sentences unrelated to preceding two-sentence contexts than to related sentences. In line with this finding, St George, Kutas, Martinez, and Sereno (1999) found that untitled (and therefore less coherent) paragraphs elicited more activation than titled paragraphs in the right but not in the left hemisphere. Furthermore, Caplan and Dapretto (2001) found that such coherence violation effects depend on the precise type of anomaly. Their subjects read question–answer dialogues where the answer was on- or off-topic, versus logical or illogical. For example, for the question, “Do you believe in angels?” the on-topic answer was, “Yes, I have my own special angel,” whereas the off-topic answer was, “Yes, I like to go to camp.” For the question, “Do you like having fun?” the logical answer was, “Yes, because it makes me happy,” whereas the illogical answer was, “No, because it makes me happy.” They found that the effect of discourse anomalies was more left-lateralized for logicity and more right-lateralized for topicality.

Hagoort (2005) proposes that the left inferior frontal gyrus (LIFG) is an active workspace for language processing. This region is involved in unifying structured pieces of knowledge that are stored in long-term memory (Jackendoff, 2007; Vosse & Kempen, 2000). It also integrates linguistic with nonlinguistic information, including, for example, world knowledge (Hagoort et al., 2004) or gestures (Willems, Ozyurek, & Hagoort, 2007). Hagoort et al. (2004) compared integration of semantic and world knowledge. Semantic knowledge is knowledge about the words in our language. World knowledge, on the other hand, refers to our knowledge about the state of affairs in the world. An example will clarify the distinction: the sentence, “Dutch trains are sour,” violates our knowledge about the meaning of the word “sour,” which is not applicable to trains. On the other hand, the sentence, “Dutch trains are white,” is an interpretable sentence, but it just so happens that Dutch trains are yellow and, therefore, most Dutch people will know this sentence to be false. Hagoort et al. found that compared to correct sentences, the LIFG, specifically BA 45 and 47, showed stronger BOLD responses to both

semantic and world knowledge anomalies. They concluded that this result provided evidence for a role of the LIFG in the integration of world knowledge as well as semantic knowledge, in line with linguistic accounts that have argued for the absence of a clear distinction between semantic and world knowledge regarding lexical items (Jackendoff, 2002).

The aim of the present fMRI study is to investigate whether the LIFG unifies information from long-term memory and discourse. We investigated whether ongoing discourse could modulate the integration of world knowledge sentences in the LIFG. If a sentence that departs from our world knowledge (“The elephant flies”) is preceded by a discourse context that supports it (*Circus traveling by airplane*), unification should be easier. As described above, there are a few fMRI studies that shed light on how the brain processes multisentence utterances. Several of these have found that the right hemisphere is involved in discourse processing (see Mitchell & Crow, 2005; Bookheimer, 2002 for reviews). Because we expected right hemisphere involvement to be stronger in our study with multisentence texts than it was in Hagoort et al. (2004), where single sentences were presented, we investigated modulations of world knowledge unification not only in the left hemisphere regions where they found increased responses to anomalies but also in their right hemisphere counterparts. To avoid confusion, we do not think a true dichotomy exists between “correct” sentences and “anomalies.” Rather, there is a continuum from sentences that fit very well with our world knowledge to sentences that are odd in the light of what we know. However, for ease of reading, we use the terms “correct” and “anomaly” as shorthand.

Our paradigm was similar to the one used by Hald et al. (2007), with either neutral or local contexts preceding correct world knowledge sentences or anomalies. The local contexts were designed to make the world knowledge anomalies more plausible. The predictions were as follows: First, left BA 45 and 47 were expected to show a stronger response to world knowledge anomalies than to correct sentences in the neutral context. Second, the local context should facilitate integration of the anomaly. The effect of world knowledge should therefore be reduced in the local context compared to the neutral context. Third, because the right hemisphere is known to be involved in discourse processing, we predicted that the right BA 45 and 47 would also show modulation of world knowledge processing by discourse context. Finally, we investigated the effect of both world knowledge and discourse context at the whole-brain level to see which additional regions might be involved.

METHODS

Subjects

Thirty-two healthy right-handed Dutch native speakers with normal or corrected-to-normal vision (13 men,

19 women; mean age = 22 ± 3) participated in the experiment. All subjects had attended or were attending university education in the Netherlands. All subjects gave written informed consent prior to the experiment and received a fee for their participation.

Stimuli and Design

One hundred thirty-two sets of Dutch-language stimuli were created for this experiment. Like Hald et al. (2007), we used a 2×2 factorial design with the factors discourse context (neutral/local) and world knowledge (correct/anomaly). The contexts were either in accordance with world knowledge (neutral context), or contained a motivation for the world knowledge anomaly in the critical sentence, making the anomaly more acceptable (local context). The critical sentences were either in accordance with world knowledge or departed from world knowledge (anomaly). There was, however, one important difference compared to the ERP study by Hald et al. In the present study, the contexts were much more tightly matched: The two versions of a particular context had an identical syntactic structure, were matched on sentence and word length and on word frequency, and differed in as few words as possible. This was necessary due to the lower temporal resolution of fMRI compared to EEG. In an fMRI study, contexts differing as widely as they did in the ERP study could lead to strong but hard-to-interpret effects on BOLD responses to the critical sentences (see Table 1 for translated examples of our stimuli).

All contexts consisted of three sentences. The first sentence was equal between the two versions of the context and introduced the subject of the text. The other two sentences were matched on syntactic structure, word frequency, and number of words, and differed on as few words as possible. The three context sentences were followed by the critical sentence. This critical sentence had two versions, which only differed in one word, the critical word. The critical word never appeared in sentence-final position, and the contexts did not contain either of the critical words.

Two hundred twenty-five sets of stimuli were pre-tested with 75 subjects in total. Nineteen subjects were given the single critical sentences only, with no prior context. The sentences were truncated before the critical word and subjects were asked to complete them to make a complete, grammatical sentence. The participants were instructed to fill in the first thing that came to mind, and to keep completions as short as possible. Thirty-one subjects were given a list of items consisting of half neutral and half local contexts in a random order, again followed by the truncated critical sentences. Stimuli were selected for the experiment if none of the subjects completed the critical sentence with the world knowledge anomaly in the single sentence condition. After the neutral context, not more than two subjects should fill in the critical word constituting the anomaly,

Table 1. Example Stimuli

| | <i>Examples</i> |
|--------------------|---|
| Neutral contexts | <p>(1) Carl Barks wrote many Donald Duck stories and invented Duckburg. In his early sketches we see Huey, Dewey and Louie as young well-behaved boys with hats and scarves. They often go out to help old ladies.</p> <p>(2) More and more lamp posts are placed in the Netherlands. This way it is easier to see the road. This is nice for drivers.</p> <p>(3) There now is a vaccination against measles. The disease is more frequent the lower the age of patients is. They are more vulnerable to contagion.</p> |
| Local contexts | <p>(1) Carl Barks wrote many Donald Duck stories and invented Duckburg. In his early sketches we see Huey, Dewey and Louie as young bad boys with striped sweaters and masks. They often go out to rob old ladies.</p> <p>(2) More and more lamp posts are placed in the Netherlands. This way it is harder to see the night sky. This is sad for astronomers.</p> <p>(3) There now is a vaccination against measles. The disease is more dangerous the higher the age of patients is. They are more vulnerable to complications.</p> |
| Critical sentences | <p>(1) Donald Duck's nephews are <u>boy scouts/thieves</u> and very smart.</p> <p>(2) With the lights on you can see <u>more/less</u> at night.</p> <p>(3) Measles is a disease that especially affects <u>children/elderly</u>.</p> |

Neutral or local contexts were followed by one out of two critical sentences, which only differed on one word (underlined).

whereas after the local context, at least three subjects should respond with the anomaly. The difference between the cloze probabilities of the critical word constituting the world knowledge anomaly after the local and neutral contexts was required to be at least 20%. Paired-sample *t* tests showed that the two contexts differed significantly in terms of cloze probability of the correct critical word [$t(133) = 18.3, p < .001$] and of the anomalies [$t(133) = -27.2, p < .001$]. The characteristics of the selected items are reported in Table 2. The cloze probability of the correct critical word is higher after the neutral context than in the single sentence, and lower after the local context. The cloze probability of the anomaly is higher after the local context than after the neutral context, or in the single sentence. Local contexts do not succeed in completely overriding world knowledge, shown by the residual cloze probability of the

correct word in the local context condition (34%). In the experiment, four different lists each contained one version of every item. The items were presented in four different pseudorandom orders. No more than three items of the same condition were presented in a row.

Procedure

Each participant saw 132 trials (33/condition). The experimental session was split in four subsequent runs of ~15 min. To ensure that the subjects remained attentive, they were given a content question about the preceding text on 10% of the trials (e.g., “Who wrote the Donald Duck stories?”), and two possible answers from which to choose with a button press (“Carl Barks/Walt Disney”). After two runs participants underwent the anatomical scan.

A trial consisted of the following elements: The discourse contexts were presented sentence-by-sentence. Each sentence was presented for a total duration of 900 msec + (250 msec times the number of words). The interstimulus interval (ISI) between sentences was 750 msec and, after the context had been presented, an ISI between 2500 and 3500 msec occurred in which a blank screen was presented. Subsequently, the critical sentence was presented word-by-word, with 500 msec per word and a 100-msec ISI between words. The intertrial interval (ITI) between the critical sentence and the next discourse context was jittered between 4000 and 6000 msec. When there was a question, it was presented 2000 msec after the last word of the critical

Table 2. Results of the Pretest for the Selected Items

| <i>Condition</i> | <i>Cloze Probability</i> | |
|------------------|--------------------------|----------------|
| | <i>Correct</i> | <i>Anomaly</i> |
| Neutral context | 77% | 1% |
| Local context | 34% | 43% |
| Single sentence | 59% | 0% |

Reported are the cloze probabilities of correct completions and anomalies after the truncated critical sentence (*Measles is a disease that especially affects _____*). The truncated critical sentence was presented after the neutral context, the local context, or in isolation.

sentence. The two possible answers were presented below the sentence 500 msec after its onset and stayed on screen for 3500 msec. At the end of this question–answer period, there was a normal ITI of 4000–6000 msec before the next context was presented.

Scanning Parameters

Subjects were scanned with a Siemens 3-T Tim-Trio MRI scanner, using a 12-channel surface coil. The TR was 1.86 sec, each volume contained 28 slices of 3 mm thickness with a slice gap of 10%. The voxel size was $3 \times 3 \times 3 \text{ mm}^3$, the field of view was 224 mm. We kept the TR shorter than usual to allow for subsequent dynamic causal modeling analyses, which require a shorter TR. The number of slices, therefore, did not allow acquisition of a full brain volume in most subjects. The experimenter always made sure that the entire temporal and frontal lobes were scanned, where the activations of interest were expected. However, this implied that in many subjects data from a part of the parietal lobe were not acquired. A whole-brain, high-resolution, structural T1-weighted MP-RAGE sequence was also performed to characterize the subjects' anatomy (TR = 2.30 sec, 192 slices with voxel size of 1 mm^3 , FOV = 256 mm).

Data Analysis

The data were preprocessed using SPM5 (Friston et al., 1995). Images were realigned to correct for movement artifacts, were slice-timing corrected, and the mean functional image for each subject was coregistered to the EPI template provided by SPM5 and to the subjects' anatomical image. Finally, images were anatomically normalized to the MNI space and spatially filtered with an isotropic 3-D Gaussian kernel (FWHM = 10 mm). Data were analyzed using the general linear model and statistical parametric mapping (Friston et al., 1995). In a 2×2 design, with the factors context (neutral/local) and world knowledge (correct/anomaly), all four regressors were modeled from the onset of the critical word to the end of the sentence and were subsequently convolved with the canonical hemodynamic response function provided by SPM5. Additional regressors were included for the discourse context, the first part of the critical sentence, the intertrial and interstimulus intervals, and finally, six movement parameters. At the single-subject level, the four conditions of interest were contrasted separately against the first part of the critical sentences of all conditions, which formed a high-level baseline. The four resulting contrast images were included in a second-level random effects analysis as well as the factor subject to account for between-subject variability. For the ROI analysis small-volume correction was used for two 10-mm spheres centered around the local maxima for the world knowledge effect reported by Hagoort et al. (2004) (MNI coordinates: left BA 47 [−48 30 −13]

and left BA 45 [−44 30 10]) and their right homotopic regions (see Figure 1 for the location of the ROIs). For both the ROI and the whole-brain analysis, images were thresholded at $t = 3.16$ ($p < .001$, uncorrected). The cluster size was used as the test statistic and only clusters significant at $p < .05$, corrected for multiple nonindependent comparisons, are reported. Local maxima are also reported for all clusters with their respective family-wise error (FWE) corrected p values.

RESULTS

ROI Analysis

In a small-volume corrected analysis for both ROIs in the LIFG, clusters were found that showed a significant effect of world knowledge in the neutral context. Significant clusters were also present in both RIFG ROIs. For the local context, no clusters were found that showed a significant world knowledge effect in three of the four regions (left BA 45, right BA 45, and right BA 47). In the left BA 47, a cluster was present showing a residual effect of world knowledge in the local context, although it was smaller than in the neutral context. We predicted that the effect of world knowledge would be weakened by the local context because the world knowledge anomaly would be easier to integrate after such a context. Clusters showing the predicted interaction between world knowledge and context were found in right BA 47 and BA 45, but in neither of the left ROIs (see Table 3 for all cluster coordinates and p values).

For the average beta values in each ROI, the interaction between context and world knowledge was significant in both right-hemisphere ROIs, but in neither of the left-hemisphere ROIs, although it showed a trend

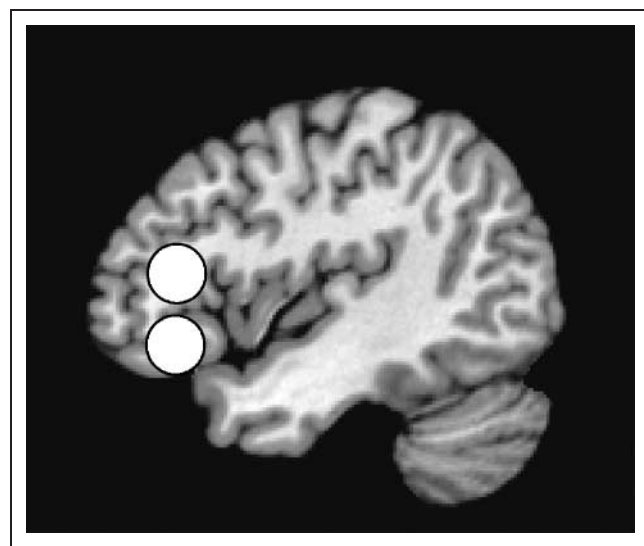


Figure 1. ROIs in the current study. Two 10-mm spheres centered at MNI coordinates [−44, 30, 10] (BA 45, top) and [−48, 30, −13] (BA 47, bottom), and their right hemisphere equivalents.

Table 3. Results for the ROI Analyses

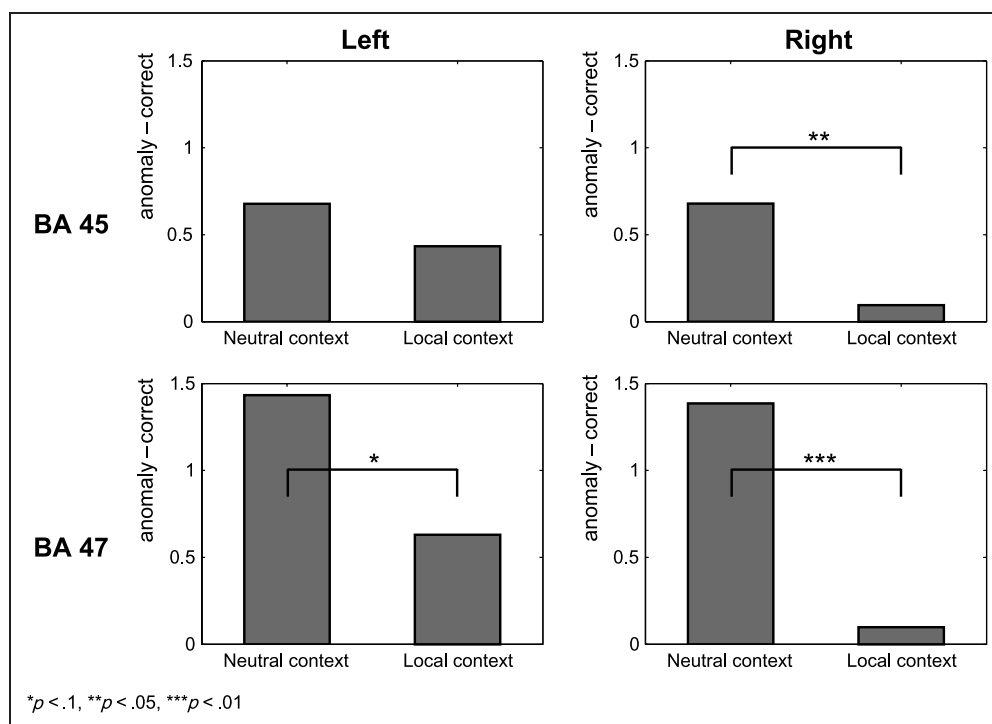
| ROI | Cluster Size (<i>kE</i>) | <i>p</i> (Cluster) | <i>t</i> (Voxel) | <i>p</i> _{FWE} (Voxel) | <i>x</i> | <i>y</i> | <i>z</i> |
|--|----------------------------|--------------------|------------------|---------------------------------|----------|----------|----------|
| <i>World Knowledge Effect in Neutral Context</i> | | | | | | | |
| L BA 45 | 103 | .003 | 4.300 | .001 | -48 | 24 | 16 |
| L BA 47 | 300 | .000 | 5.330 | .000 | -46 | 34 | -16 |
| R BA 45 | 271 | .000 | 5.630 | .000 | 50 | 28 | 16 |
| R BA 47 | 307 | .000 | 5.800 | .000 | 44 | 38 | -12 |
| <i>World Knowledge Effect in Local Context</i> | | | | | | | |
| L BA 47 | 25 | .015 | 4.270 | .000 | -48 | 38 | -16 |
| <i>Interaction World Knowledge × Context</i> | | | | | | | |
| R BA 47 | 1 | .032 | 3.360 | .022 | 48 | 36 | -4 |
| R BA 45 | 5 | .026 | 3.410 | .019 | 50 | 32 | 4 |
| | 8 | .024 | 3.300 | .025 | 50 | 30 | 12 |
| | 1 | .032 | 3.260 | .028 | 40 | 34 | 2 |

Small-volume correction was performed for the four ROIs located in bilateral BA 45 and 47. Reported are all clusters significant at cluster level with $p < .05$, corrected for multiple comparisons. The MNI coordinates of the maxima of all clusters and their respective voxelwise t and p values (FWE corrected for multiple nonindependent comparisons) are also reported.

toward an interaction in left BA 47 [left BA 45: $F(1, 31) < 1$; left BA 47: $F(1, 31) = 3.4, p < .074$; right BA 45: $F(1, 31) = 6.3, p < .017$; right BA 47: $F(1, 31) = 9.5, p < .008$]. Figure 2 shows the difference between correct sentences and anomalies in all ROIs, clearly showing that the world knowledge effect is more reduced in the RIFG than in the LIFG.

To further specify the interaction between context and world knowledge in the RIFG, we extracted the average beta values for the clusters showing a significant interaction in the small-volume corrected analysis in the two right-hemisphere ROIs. We inserted these in a $2 \times 2 \times 2$ (Brain region \times Context \times World knowledge) repeated measures general linear model. A significant main effect

Figure 2. Difference bar graphs between correct and anomalous sentences in neutral and local contexts, in four ROIs (right/left BA 45 and 47). Displayed are the differences in average beta values between correct sentences and anomalies for both types of contexts for the entire ROIs.



was found for world knowledge [$F(1, 31) = 17.86, p < .0001$], which as expected was qualified by a Context \times World knowledge interaction [$F(1, 31) = 14.92, p < .001$], but also by an Area \times World knowledge interaction [$F(1, 31) = 6.27, p < .018$], with a stronger effect of world knowledge in BA 45 than in BA 47. For the Context \times World knowledge interaction, in which we were mainly interested, we tested for the simple effects of both world knowledge and context. The effect of world knowledge was significant within the neutral ($F = 31.9, p < .0001$) but not the local context ($F < 1$). The effect of context was significant for both correct sentences ($F = 11.9, p < .002$) and world knowledge anomalies ($F = 4.3, p < .047$). The effect was reversed in the latter condition: The beta values were higher for world knowledge anomalies in the neutral than in the local context. This suggests that the local context facilitated integration of the world knowledge anomalies.

We did not find the interaction between context and world knowledge in the small-volume corrected data in LIFG. Because, however, we found a trend toward an interaction in the mean beta values for left BA 47, we performed the same simple effects tests outlined above for the average beta values of this ROI. The effect of world knowledge was significant for the neutral context ($F = 20.770, p < .0001$) and showed a trend toward significance in the local context ($F = 3.572, p < .068$). The effect of context was only significant for the correct sentences ($F = 7.925, p < .008$; anomalies: $F < 1$). The trend toward an interaction effect in left BA 47 is therefore characterized by a reduced effect of world knowledge in the local context. Figure 3 shows the differences between local and neutral

contexts for the two world knowledge conditions in all ROIs, clearly showing that the context effect is reversed in the right- but not in the left-hemisphere ROIs.

Whole-brain Analysis

To further investigate the modulation of world knowledge effects by preceding discourse, we also performed a whole-brain analysis. Results for all following comparisons are reported in Table 4. Clusters are reported if significant at $p < .05$ at cluster-level, corrected for multiple nonindependent comparisons.

World Knowledge Effect in the Neutral Context

Seven clusters showed a stronger response to world knowledge anomalies than to correct sentences. These were bilateral inferior frontal (BA 45 and 47), right middle temporal (BA 21), the bilateral angular gyri (BA 39/40), the right caudate nucleus, and a cluster spreading across the bilateral superior frontal gyri (BA 9). All clusters are shown in Figure 4. In the whole-brain analysis, there was no significant effect of world knowledge in the local context.

Main Effect of Context

One cluster showed a marginally significant effect at the cluster level (but highly significant at the voxel level). This cluster was located in the left middle temporal gyrus and showed a stronger response to local than to

Figure 3. Difference bar graphs between neutral and local contexts for correct and anomalous sentences, in four ROIs (right/left BA 45 and 47). Displayed are the differences in average beta values between neutral and local context contexts for each type of sentence.

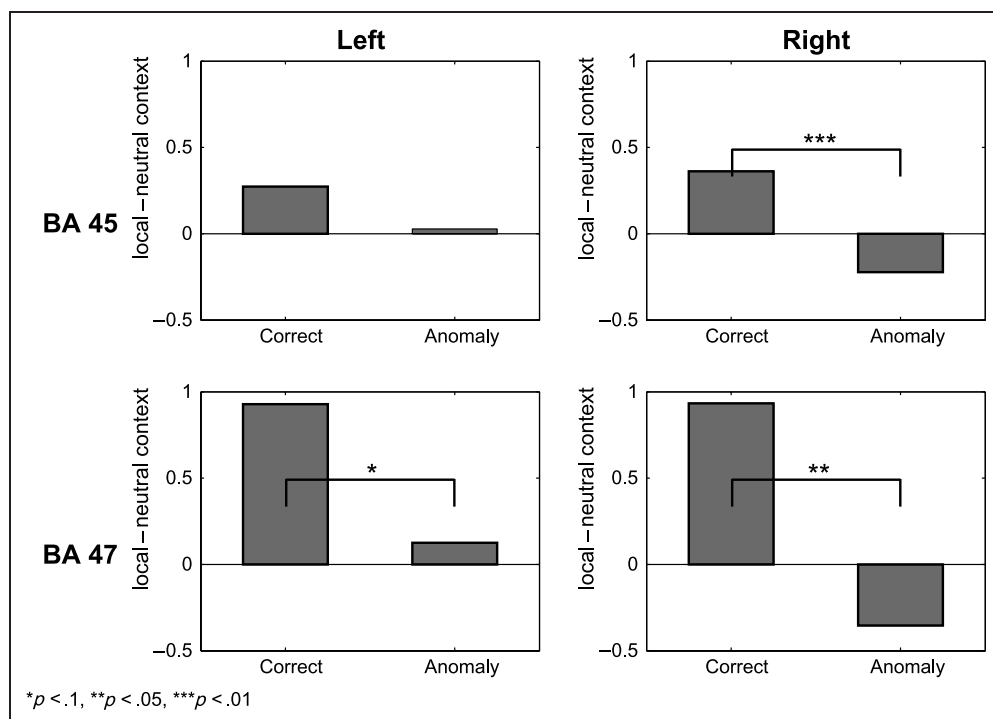


Table 4. Results for All Tested Contrasts in Whole-brain Analysis

| Effect | <i>x</i> | <i>y</i> | <i>z</i> | Label | Cluster Size (<i>kE</i>) | <i>p</i> (Clusters) | <i>t</i> | <i>p</i> _{FWE} (Voxel) |
|---|--------------------------------|----------|----------|------------|----------------------------|---------------------|----------|---------------------------------|
| World knowledge effect in neutral context | 46 | 42 | -14 | R BA 47 | 2412 | .000 | 6.52 | .000 |
| | -46 | 34 | -16 | L BA 47 | 1502 | .000 | 5.33 | .004 |
| | 14 | 54 | 24 | R BA 9 | 1612 | .000 | 4.94 | .020 |
| | 64 | -42 | -8 | R BA 21 | 944 | .000 | 4.87 | .034 |
| | 56 | -56 | 26 | R BA 39/40 | 347 | .023 | 4.87 | .026 |
| | -58 | -64 | 24 | L BA 39/40 | 815 | .000 | 5.74 | .001 |
| | 12 | 4 | 10 | R caudate | 394 | .015 | 4.19 | .237 |
| World knowledge effect in local context | <i>No significant clusters</i> | | | | | | | |
| Main effect context (local > neutral) | -60 | -22 | 20 | L BA 21 | 266 | .054 | 5.46 | .003 |
| Interaction World knowledge × Context | -64 | -48 | 28 | BA 39/40 | 375 | .018 | 4.83 | .030 |

Listed are clusters with a cluster size (*kE*) significant at $p < .05$, corrected for multiple comparisons. The local maxima of all clusters and their respective voxelwise *t* and *p* values (FWE corrected for multiple nonindependent comparisons) are also reported.

neutral contexts. Simple main effects analyses revealed that for this cluster the effect was independent of whether the final sentence was correct or constituted a world knowledge anomaly (correct: $F = 19.26$, $p < .001$, anomaly: $F = 11.69$, $p < .002$).

Interaction World Knowledge and Context

This comparison resulted in one significant cluster located in the left angular gyrus (see Figure 5A). Two clusters showed a marginally significant response. These were the right hemisphere equivalent of the cluster in the left angular gyrus as well as a cluster in the RIFG. This last result is similar to the effect observed in the ROI analysis. Comparison of the beta values in the left angular gyrus revealed a crossover interaction between context and world knowledge, suggesting that this region is predominantly sensitive to the coherence between critical sentence and context (see Figure 5B).

DISCUSSION

The first main result of this study was to replicate the world knowledge effect of Hagoort et al. (2004): Sentences with world knowledge anomalies triggered significantly greater activation of the left inferior frontal gyrus (LIFG; BA 45 and 47) compared to correct sentences, particularly for the neutral context. A similar effect was observed in the RIFG (BA 45 and 47; cf. Figure 2).

The second main finding of this study relates to the modulation of the world knowledge effect by context. More specifically, we predicted that the effect of world knowledge would be reduced in the local context compared to the neutral context. Consistent with this, there was a trend toward an interaction between context and world knowledge in the LIFG (in BA 47, but not in BA 45), suggesting that this region supports on-line unification of linguistic and extralinguistic information (cf. Figure 2). The residual effect of world knowledge in the local context observed in the LIFG suggests that the ongoing

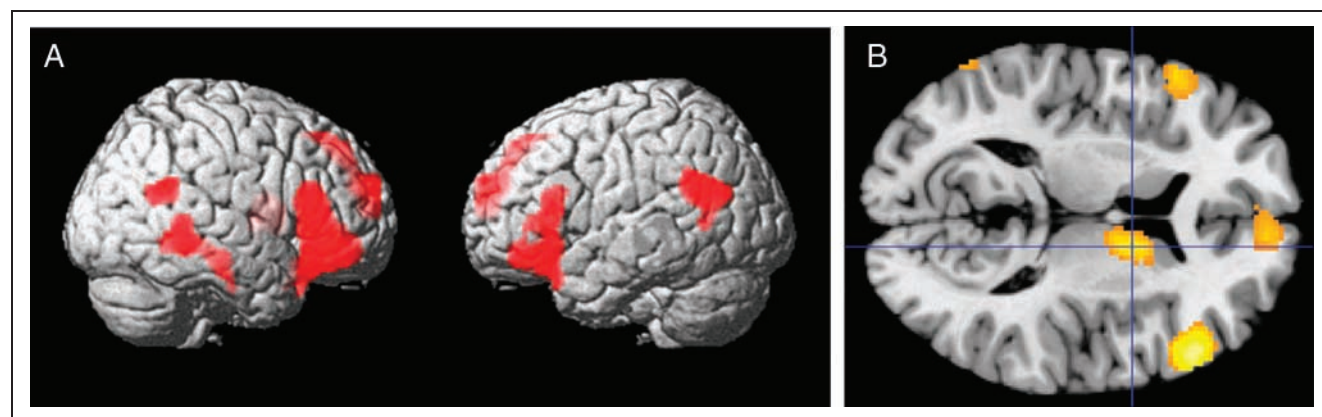
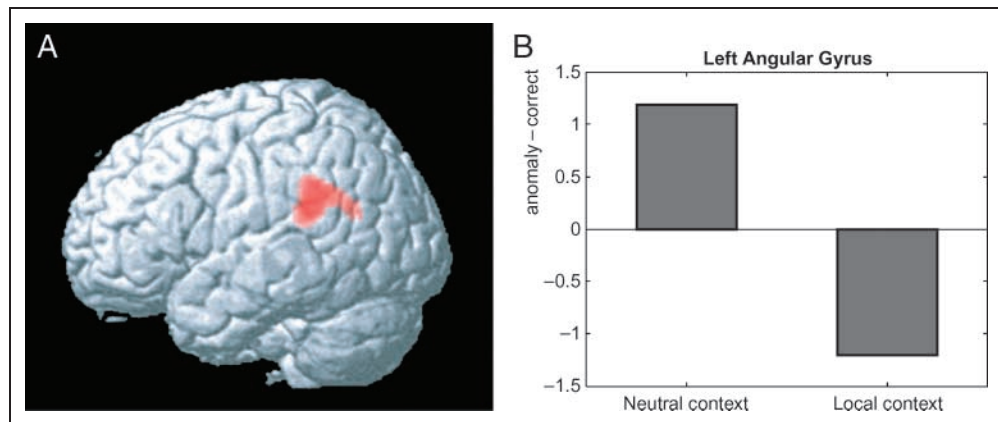


Figure 4. Effect of world knowledge in the neutral context. The map is thresholded at $p < .001$, uncorrected; only clusters with $p < .05$, corrected at cluster level, are shown. (A) Significant clusters rendered onto the cortical surface. (B) Transversal slice at $z = 10$ showing the significant cluster in the caudate nucleus.

Figure 5. Interaction of world knowledge and context in the angular gyrus. (A) Location of the cluster showing a significant interaction. (B) Difference bar graph for the beta values in this cluster.



discourse (i.e., the local context) does not fully neutralize the world knowledge effect during comprehension of the critical sentence but the trend toward an interaction between context and world knowledge in the left BA 47 indicates that the context does affect processing of the world knowledge sentences in this area at least to some extent.

In contrast to the LIFG, in the RIFG, the interaction between world knowledge and context was significant, consistent with our third prediction that the right hemisphere is involved in discourse processing. Indeed, specification of the interaction in the RIFG revealed that after the local context, the BOLD response to anomalies was weaker in the local than in the neutral context. This suggests that the RIFG is more sensitive to the discourse modulation than the LIFG.

These results are in line with several studies showing a strong involvement of the right hemisphere in discourse processing (for a review, see Bookheimer, 2002). The RIFG, therefore, also seems to play an important role in semantic unification, in addition to the LIFG (an effect already noticed in the data of Hagoort et al., 2004, albeit nonsignificant). Moreover, the RIFG might be relatively more involved than the LIFG in forming an integrated representation of ongoing discourse. Our results are also in line with the idea that the left hemisphere is more sensitive to dominant semantic relationships, whereas the right hemisphere is more sensitive to novel, inferred semantic relationships (Giora, 2007; see Introduction). The LIFG continues to detect the discrepancy in the sentence, “Measles is a disease that especially affects elderly,” due to the strong association between *measles* and *children*, whereas the RIFG has inferred the novel association between *measles* and *elderly* from the context (*There now is a vaccination against measles. The disease is more dangerous the higher the age of patients is. They are more vulnerable to complications.*)

What the precise division of labor is between the LIFG and the RIFG remains to be explored. The fact that we did not observe a significant interaction between context and world knowledge in the LIFG suggests that this

region remains sensitive to semantic unification of incoming information with prior world knowledge, even if preceding discourse context overrides this knowledge. The interpretation of the world knowledge effect in the RIFG is less straightforward: The interaction between context and world knowledge and the absence of a world knowledge effect in the local context suggest that this area is predominantly sensitive to the congruency between discourse context and target sentence. However, because the effect of world knowledge is not completely reversed in the local context, the region is also sensitive to information from long-term memory.

The main new finding in this study is that semantic unification at the level investigated here is more bilateral than was earlier observed. Clearly, the RIFG is involved in forming a representation of the ongoing discourse. This result fits with previous claims that the right hemisphere is involved in processing multisentence texts. As outlined in the Introduction, the right hemisphere is sensitive to text coherence (Kuperberg et al., 2006; Caplan & Dapretto, 2001; St George et al., 1999). Long, Baynes, and Prat (2005) used divided visual field experiments to determine what kind of discourse representation each hemisphere contributes to. They concluded that only the left hemisphere forms a propositional representation (the textbase in the description of van Dijk & Kintsch, 1983; see Introduction), evidenced by stronger priming effects for words previously presented in the same propositions than in different propositions when presented in the right visual field. The right hemisphere shows no differential priming for words in same or different propositions and, therefore, does not appear to form a propositional representation. However, because it does show stronger priming for words in same passages than in different passages, it seems to form a more general discourse representation. We suggest that this general discourse representation is a situation model. Because the target sentences in our paradigm contain distinct propositions from the preceding passages, they cause no local anomalies in the propositional structure. However, at a more global level, they violate the content of the

preceding discourse. Our results therefore suggest that the RIFG is more sensitive to discourse anomalies than the LIFG.

In addition to the ROI analysis, we also performed a whole-brain analysis to further investigate effects of world knowledge and the modulations of these effects by context. We observed a main effect of context in the left middle temporal gyrus. Note, however, that this effect occurs at the critical part of the final sentence and is independent of the final sentence type which the subjects read. The contexts were tightly matched on word frequency, number of words, and syntactic structure, although the local contexts were, by definition, more novel than the neutral ones, which may have increased the overall integration load, irrespective of the type of sentence that had to be integrated. For the effect of world knowledge in the neutral context, we observed several significant clusters, located in the IFG bilaterally, the bilateral medial superior frontal gyri, the angular gyrus bilaterally, the middle temporal gyrus, and the right caudate nucleus. Compared to Hagoort et al. (2004), there were three modifications in the current study. First, we scanned 32 subjects instead of 14, leading to greater statistical power in our analyses. Second, we used a more high-level baseline, which may have reduced the noise in the data, thereby increasing sensitivity. Finally and crucially, our world knowledge sentences were preceded by contexts. This meant that a stronger response to world knowledge anomalies than to correct sentences is very likely induced by a mismatch between the critical sentence and the preceding context. The world knowledge anomalies in this condition do not only violate world knowledge, but also the content of the neutral discourse context. In the local context, it is the correct sentence that violates the previous discourse. In regions where congruency between the previous context and the critical sentence is what matters, an interaction between world knowledge and context would be expected. This interaction was present in the left angular gyrus (and was present just below threshold in the right angular gyrus). These regions therefore appear to be more sensitive to the coherence between the discourse context and the critical sentence than to world knowledge per se.

Different functions have been hypothesized for the angular gyrus. For example, it has been implicated in visual word form recognition (Deleon et al., 2007). Our critical words were designed to have a higher cloze probability in precisely the two conditions that show the lowest activation in the left angular gyrus. These highly constraining contexts may have led to top-down influences on word form retrieval, which would explain the interaction effect.

Price (2000) argues that the angular gyrus is not specific to visual word forms but rather is engaged when semantic associations are made. Mechelli, Josephs, Lambon Ralph, McClelland, and Price (2007) lend empirical support to the suggestion that the left angular gyrus is sensitive to

semantic associations between words by showing that the left angular gyrus shows a stronger BOLD response to semantically related word pairs than to unrelated and phonologically related word pairs. Humphries, Binder, Medler, and Liebenthal (2007) propose a similar but slightly more complex role of the left angular gyrus in semantic integration. In a study which compared congruent sentences, sentences with unrelated words and pseudo-word sentences, and reshuffled word list versions of all these, they found stronger activation in the left angular gyrus for congruent sentences than for all other conditions. They therefore suggested that this area is involved in integrating incoming semantic information with an ongoing context. At first sight, our results do not match those of Humphries et al. (2007), as they find weaker activations in the random sentences (which would seem the hardest to integrate) than in the congruent sentence. Like us, Ni et al. (2000) also found stronger activations in the angular gyrus in sentences containing semantic anomalies than congruent sentences. One possibility is that the random sentences in the Humphries et al. study fail to generate semantic integration in the first place, whereas sentences with an anomaly (present study, Ni et al., 2000) do generate a context in which the word meaning embodying the anomaly is then harder to integrate.

The inferior parietal region has also been implicated in the phonological loop component of working memory (Baddeley, 2003). The phonological loop is a component of working memory that temporarily stores sound and language. It consists of two components, a phonological store and an articulatory rehearsal process. Baddeley (2003) and others have suggested that the inferior parietal region is involved in the storage component. This might explain the interaction between context and world knowledge that we observed for this region. It is possible that when subjects encounter a sentence that does not easily combine with the information that they just read, they retrieve the surface form and check whether their interpretation was correct. However, because our experiment was not designed to distinguish between the different possibilities outlined above, we will not draw strong conclusions about the role of the left angular gyrus at this point.

In conclusion, we suggest that the RIFG is specifically involved in forming a general representation (a situation model) of ongoing discourse and in unifying discourse information with previously stored knowledge in long-term memory. Based on our findings, we suggest that both the LIFG and the RIFG are recruited in unification, but with a division of labor when it comes to discourse comprehension. What precisely distinguishes the different types of context that recruit the LIFG and the RIFG, respectively, for unification is a question for further research.

Reprint requests should be sent to Laura Menenti, Donders Centre for Cognitive Neuroimaging, P.O. Box 9101, 6500 HB Nijmegen, the Netherlands, or via e-mail: Laura.Menenti@fcdonders.ru.nl.

REFERENCES

- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829–836.
- Bartolo, A., Benuzzi, F., Nocetti, L., Baraldi, P., & Nichelli, P. (2006). Humor comprehension and appreciation: An fMRI study. *Journal of Cognitive Neuroscience*, 18, 1789–1798.
- Beeman, M. (1998). Coarse semantic coding and discourse comprehension. In M. Beeman & C. Chiarello (Eds.), *Right hemisphere language comprehension: Perspectives from cognitive neuroscience* (pp. 255–284). Mahwah, NJ: Erlbaum.
- Bookheimer, S. Y. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neurosciences*, 25, 151–188.
- Caplan, R., & Dapretto, M. (2001). Making sense during conversation: An fMRI study. *NeuroReport*, 12, 3625–3632.
- Deleon, J., Gottesman, R. F., Kleinman, J. T., Newhart, M., Davis, C., Heidler-Gary, J., et al. (2007). Neural regions essential for distinct cognitive processes underlying picture naming. *Brain*, 130, 1408–1422.
- Duffy, S. A., & Keir, J. A. (2004). Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition*, 32, 551–559.
- Friston, K., Holmes, A., Worsley, K., Poline, J.-B., Frith, C., & Frackowiak, R. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2, 189–210.
- Giora, R. (2007). Is metaphor special? *Brain and Language*, 100, 111–114.
- Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology*, 48, 163–189.
- Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, 9, 416–423.
- Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304, 438–441.
- Hald, L., Steenbeek-Planting, E. G., & Hagoort, P. (2007). The interaction of discourse context and world knowledge in online sentence comprehension. Evidence from the N400. *Brain Research*, 1146, 210–218.
- Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2007). Time course of semantic processes during sentence comprehension: An fMRI study. *Neuroimage*, 36, 924–932.
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
- Jackendoff, R. (2007). A parallel architecture perspective on language processing. *Brain Research*, 1146, 2–22.
- Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, 9, 512–518.
- Kuperberg, G. R., Lakshmanan, B. M., Caplan, D. N., & Holcomb, P. J. (2006). Making sense of discourse: An fMRI study of causal inferring across sentences. *Neuroimage*, 33, 343–361.
- Long, D. L., Baynes, K., & Prat, C. S. (2005). The propositional structure of discourse in the two cerebral hemispheres. *Brain and Language*, 95, 383–394.
- Mashal, N., Faust, M., Hendler, T., & Jung-Beeman, M. (2007). An fMRI investigation of the neural correlates underlying the processing of novel metaphoric expressions. *Brain and Language*, 100, 115–126.
- Mechelli, A., Josephs, O., Lambon Ralph, M. A., McClelland, J. L., & Price, C. J. (2007). Dissociating stimulus-driven semantic and phonological effects during reading and naming. *Human Brain Mapping*, 28, 205–217.
- Mitchell, R. L. C., & Crow, T. J. (2005). Right hemisphere language functions and schizophrenia: The forgotten hemisphere? *Brain*, 128, 963–978.
- Ni, W., Constable, R. T., Mencl, W. E., Pugh, K. R., Fulbright, R. K., Shaywitz, S. E., et al. (2000). An event-related neuroimaging study distinguishing form and content in sentence processing. *Journal of Cognitive Neuroscience*, 12, 120–133.
- Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, 18, 1098–1111.
- Price, C. J. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy*, 197, 335–359.
- Schmidt, G. L., DeBuse, C. J., & Seger, C. A. (2007). Right hemisphere metaphor processing? Characterizing the lateralization of semantic processes. *Brain and Language*, 100, 127–141.
- St George, M., Kutas, M., Martinez, A., & Sereno, M. I. (1999). Semantic integration in reading: Engagement of the right hemisphere during discourse processing. *Brain*, 122, 1317–1325.
- Stringaris, A. K., Medford, N., Giora, R., Giampietro, V. C., Brammer, M. J., & David, A. S. (2006). How metaphors influence semantic relatedness judgments: The role of the right frontal cortex. *Neuroimage*, 33, 784–793.
- Stringaris, A. K., Medford, N. C., Giampietro, V., Brammer, M. J., & David, A. S. (2007). Deriving meaning: Distinct neural mechanisms for metaphoric, literal, and non-meaningful sentences. *Brain and Language*, 100, 150–162.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, 75, 105–143.
- Willems, R. M., Ozyurek, A., & Hagoort, P. (2007). When language meets action: The neural integration of gesture and speech. *Cerebral Cortex*, 17, 2322–2333.