Phonological markers of information structure: An fMRI study

Tessa M. van Leeuwen [a,1], Monique J.A. Lamers [b,c], Karl Magnus Petersson [d,a], Carlos Gussenhoven [e], Toni Rietveld [e], Benedikt Poser [a,f], Peter Hagoort [a,d,*]

[a] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands
[b] Department of Language and Communication, VU University, Amsterdam, The Netherlands
[c] The Eargroup, Herentalsebaan 75, B-2100 Antwerp-Deurne, Belgium
[d] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[e] Department of Linguistics, Radboud University Nijmegen, Nijmegen, The Netherlands
[f] Erwin L. Hahn Institute for Magnetic Resonance Imaging, University Duisburg-Essen, Essen, Germany

ARTICLE INFO

ABSTRACT

In this fMRI study we investigate the neural correlates of information structure integration during sentence comprehension in Dutch. We looked into how prosodic cues (pitch accents) that signal the information status of constituents to the listener (new information) are combined with other types of information during the unification process. The difficulty of unifying the prosodic cues into overall sentence meaning was manipulated by constructing sentences in which the pitch accent *did* (focus-accent agreement), and sentences in which the pitch accent *did not* (focus-accent disagreement) match the expectations for focus constituents of the sentence. In case of a mismatch, the load on unification processes increases. Our results show two anatomically distinct effects of focus-accent disagreement, one located in the posterior left inferior frontal gyrus (LIFG, BA6/44), and one in the more anterior-ventral LIFG (BA 47/45). Our results confirm that information structure is taken into account during unification, and imply an important role for the LIFG in unification processes, in line with previous fMRI studies.

1. Introduction

In addition to the meaning of sentences as expressed by the words and the thematic relations that hold between them, languages typically also express ways in which sentences are related to the preceding discourse. One such aspect of linguistic meaning is known as 'information structure'. The information structure of a sentence essentially focuses the listener's attention on the crucial (new, focus) information in it. In English and Dutch, prosody plays a crucial role in marking information structure. For instance, in question–answer pairs, the new or relevant information in the answer will typically be pitch accented. After a question like 'What did Mary buy at the market?', the answer might be 'Mary bought VEGETABLES' (accented word in capitals). In this

case, vegetables is the focus constituent, which corresponds to the information provided for the Wh-element in the question.

There is no linguistic universal for signaling information structure. The way information structure is expressed varies across languages. In some languages it may impose syntactic locations for the focus constituent, in others focus-marking particles are used, or prosodic features like phrasing and accentuation (Gussenhoven, 2008; Kotschi, 2006; Miller, 2006). The language dependence of prosodic information marking was illustrated in an experiment with Dutch and Italian speakers by Swerts, Krahmer, and Avesani (2002). They showed that Italian speakers made no distinctions in the pronunciation of sentences like *C'é un triangolo nero* [It is a black triangle], regardless of whether the preceding context specified *triangolo* (*C'é un triangolo viola*) or *nero* (*C'é un rettangolo nero*) as the contrastive information. Dutch subjects, by contrast, produced different prosodic structures for such sentences, placing a pitch accent on just those words that represented the new information. Thus, when *Nu is er een rode driehoek te zien* [Now you see a red triangle] was followed by a picture of a black triangle, they marked the contrastive adjective *black* with a pitch accent: *En nu is er een ZWARTE driehoek te zien* [Now you see a BLACK triangle]. After *Nu is er een zwarte rechthoek te zien* [Now you see a black rectangle], however, a pitch accent was placed on the noun *triangle*: *En nu is er een zwarte DRIEHOEK te zien* [Now

* Correspondence to: Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands. Tel.: +31 24 3610648.
E-mail addresses: tesvlee@gmail.com (T.M. van Leeuwen), peter.hagoort@donders.ru.nl (P. Hagoort).
[1] Present address: Department of Neurophysiology, Max Planck Institute for Brain Research, Deutschordenstraße 46, 60528 Frankfurt am Main, Germany. Tel.: +49 69 96769 240.

you see a black TRIANGLE]. In addition, the authors showed that Dutch listeners were able to retrieve the correct context for each of these information structure prosodic markings well above chance, while Italian listeners were not. This example illustrates the saliency of prosodic marking of information structure within noun phrases for Dutch listeners. The prosodic marking guides the listener to the relevant and new information (the focus), and aids in the interpretation and integration of the context. Prosodic marking therefore plays an important role in the overall process of sentence comprehension, by signaling how the sentence relates to the preceding discourse and which information requires a deeper level of processing. It has been shown that in many cases listeners and readers do not always interpret sentences completely or even fully correctly. They extract the information that is needed in any given communicative situation, which is usually less than all that is provided in the input string. Ferreira, Bailey, and Ferraro (2002) introduced the phrase 'good-enough processing' to refer to the listeners' and readers' interpretation strategies. In this context, linguistic devices that mark information structure might help the listener/reader in allocating processing resources to the most relevant pieces of information in the linguistic input.

In this study we focus on the neuronal infrastructure that supports the extraction and integration of prosodic features that mark information structure. We would like to stress that prosodic accent is just one of the ways in which information structure can be conveyed; we are not aiming to elucidate the neural underpinnings of prosody *per se* as many studies have previously done (*e.g.*, Doherty, West, Dilley, Shattuck-Hufnagel, & Caplan, 2004; Hesling, Clément, Bordessoules, & Allard, 2005; Hesling, Dilharreguy, Clément, Bordessoules, & Allard, 2005; Humphries, Love, Swinney, & Hickok, 2005; Meyer, Alter, Friederici, Lohmann, & von Cramon, 2002; Meyer, Steinhauer, Alter, Friederici, & von Cramon, 2004; Plante, Creusere, & Sabin, 2002; Zhang, Shu, Zhou, Wang, & Li, 2010). Instead, we specifically aim to identify the neural substrates of the role of prosodic accent in information structure. Information structure is clearly a process that goes beyond the retrieval of single word information, such as the phonemic segments and the lexical stress pattern of a word. Information structure is inherently bound to multiword utterances and discourse, as the information structure of a sentence is always dependent on preceding sentences and discourse and information. As such it is part of what Hagoort (2005) has referred to as unification. Unification refers to the expressive power of human language, which is based on the possibility to combine elements from memory in endless, often novel ways. This process of deriving complex meaning from lexical building blocks is central to the human language faculty and has been found to recruit (left) frontal cortex structures (Hagoort, 2005; Hagoort, Baggio, & Willems, 2009; Snijders et al., 2009). For instance, during semantic unification single word meanings are integrated into an overall interpretation of the whole utterance.

To date, a few ERP studies have investigated the processing of information structure during language comprehension. Cowles, Kluender, Kutas, and Polinsky (2007) and Wang, Hagoort, and Yang (2009) characterized brain responses related to focus constituents. In a reading experiment, Cowles et al. (2007) found that if a contextually unexpected word was placed in focus position in an it-cleft construction ('*It was the rabbits that ate the lettuce*'), this leads to an N400 effect, illustrating the strong expectancy of a focus constituent created by the it-cleft. N400 effects are commonly associated with semantic anomalies or context violations in sentences (Kutas & Hillyard, 1980; Van Berkum, Zwitserlood, Hagoort, & Brown, 2003). A similar effect was found by Wang et al. (2009), who showed that if a semantic violation was in focus position, the semantic anomaly caused an N400 effect, whereas when the semantically inappropriate word was not in focus

position, a strongly reduced N400 effect was observed (semantic illusion). Differences in the processing of focus and non-focus constituents are also reported in studies that investigated auditory language processing. Irrespective of the pitch accent, focus constituents reveal a widely distributed late positivity (500–700 ms) in comparison to a non-focus constituent (Johnson, Breen, Clifton, & Morris Florak, 2003, see also Heim and Alter, 2006). These studies also revealed that prior context causes language users to generate expectations on the focus status of particular entities. Whereas a missing pitch accent on a focus constituent may result in an early anteriorly distributed negativity (100–500 ms), no effects are found on superfluous accents (*i.e.*, pitch accent on non-focus constituents) (*e.g.*, Heim & Alter 2006; Hruska & Alter, 2004; Toepel & Alter, 2004). These findings illustrate that information structure is actively used by the reader or listener during on-line sentence processing.

In a first fMRI study on information structure (Burholt Kristensen, Wang, Petersson, & Hagoort, 2012) we found common activations between a language task and a spatial attention task, indicating that focus markers such as pitch accents activate a domain general attention network, which is sensitive to semantic/pragmatic aspects of language. This indicated that attention and language comprehension are highly interactive, and that information structure might be used to recruit the contribution of attentional networks to process new information in the utterance more extensively. In the current fMRI study our goal was a different one. In this study we aim to functionally localize the brain regions that provide the required infrastructure for extracting information structure from pitch accent in Dutch. To this end, we manipulated the difficulty of incorporating the prosodic cues into the overall sentence interpretation, thereby increasing the unification load. In our study, a specific question was how the neural correlates of unification of prosodic marking relate to brain regions that are already known to be involved in various aspects of language comprehension, in particular those regions that are related to unification. Based on previous literature, unification is assumed to recruit the critical contribution of the left inferior frontal gyrus (LIFG) (Hagoort, 2005; Hagoort et al., 2009) along with temporal areas. For this reason, the LIFG is a region of special interest in our study. In particular, we were interested in the classical frontal language area known as Broca's area (which consists of Brodmann areas (BA) 45 and 44), and additionally in BA 47 and the ventral part of BA 6. The LIFG is activated during many language tasks (for reviews, see Bookheimer, 2002; Indefrey, 2004), and capable of holding information on-line over a relatively large time span of several seconds (Mesulam, 2002). Such working-memory capacity is a pre-requisite for unification operations, since all relevant information needs to remain available across utterances that often have a duration of a few seconds.

It has been suggested that, within the LIFG, a functionally defined gradient of language processing exists (Bookheimer, 2002; Bornkessel-Schlesewsky, Grewe, & Schlesewsky, 2012; Costafreda et al., 2006; Dapretto & Bookheimer, 1999; Friederici, Opitz, & Cramon, 2000; Gough, Nobre, & Devlin, 2005; Hagoort, 2005; Poldrack et al., 1999; Price, Moore, Humphreys, & Wise, 1997; Xiang, Fonteijn, Norris, & Hagoort, 2010), extending from the anterior-ventral to the more dorsal-posterior part of the LIFG. Although this gradient shows substantial overlap of activation patterns (Bookheimer, 2002; Petersson, Forkstam, & Ingvar, 2004), the above mentioned studies suggest that roughly speaking, the anterior-ventral part of the LIFG (BA 47 and 45) is involved in tasks that demand semantic processing, the more posterior-dorsal part (BA 44 and parts of BA 6) in phonological tasks, whereas syntactic processing is located in between the two extremes (BA 45 and 44). There is an ongoing discussion whether language-related activity in Broca's area reflects indeed purely language-specific processing

or more general executive and selection functions applied to the linguistic domain (*e.g.*, Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Thothathiri, Kim, Trueswell, & Thompson-Schill, 2012). Also, several studies have found overlapping activations within LIFG for different aspects of linguistic processing (Heim, Eickhoff, & Amunts, 2008), for instance evidence that BA45 is involved in both semantic and phonological processing (Heim, Eickhoff, & Amunts, 2009). Similarly, Roder, Stock, Neville, Bien, and Rosler (2002), investigating both syntactic and semantic processing, found that both aspects of language activated BA 44 and 45; in Rodd, Longe, Randall, and Tyler (2010), semantic ambiguity and syntactic ambiguity were compared and semantic ambiguity was also found to activate BA 44 in central LIFG; in Rodd, Johnsrude, and Davis (2012) a semantic ambiguity task led to similar ambiguity related effects in both anterior and posterior LIFG, and semantic manipulations were found to activate BA 44/posterior LIFG in more instances (McDermott, Petersen, Watson, & Ojemann, 2003; Zempleni, Renken, Hoeks, Hoogduin, & Stowe, 2007; Zhu et al., 2013). However, a recent meta-analysis (Hagoort & Indefrey, 2014) clearly confirms the general tendency for semantic aspects of language to activate BA 45/47 and syntactic processing to activate BA44/45.

Support for functionally distinct subregions in LIFG comes from numerous studies that have used paradigms with a higher unification load for specific aspects of language processing. Studies with a higher semantic unification load have used either semantic or pragmatic anomalies, or semantic ambiguities (*e.g.*, Davis et al., 2007; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Kuperberg, Sitnikova, & Lakshmanan, 2008; Rodd, Davis, & Johnsrude, 2005; Ruschemeyer, Zysset, & Friederici, 2006). In these semantic unification studies, reviewed in Hagoort et al. (2009), Brodmann areas 45 and 47 in the LIFG were activated most consistently. For syntactic unification, for example Snijders et al. (2009) exploited syntactic ambiguity in an ROI approach to confirm the involvement of the central LIFG around MNI coordinates [−44 19 14] in syntactic unification. In other studies reporting BA 44 activation, the difficulty of syntactic integration, argument hierarchy, and/or syntactic working memory load was manipulated (Bornkessel, Zysset, Friederici, von Cramon, & Schlesewsky, 2005; Fiebach, Schlesewsky, Lohmann, von Cramon, & Friederici, 2005; Friederici et al., 2000; Grewe et al., 2005; Kristensen, Engberg-Pedersen, Nielsen, & Wallentin, 2013).

More support for a functional gradient in LIFG stems from the analysis of resting state data in a study by Xiang et al. (2010). In this study, seed regions in three parts of the LIFG were defined; pars opercularis (BA 44), pars triangularis (BA 45), and pars orbitalis (BA 47). Of these seed regions, the functional connectivity patterns to the rest of the brain were explored. Although overlap exists, a clear topographical functional connectivity pattern in the left middle frontal, parietal, and temporal areas was revealed for the three seed regions: BA 44 correlated mainly with areas involved in phonological and syntactic processing (posterior superior temporal gyrus); BA 45 mainly correlated with brain areas implicated in syntactic processing (posterior middle temporal gyrus) and also extended to brain regions for phonological and semantic processing; BA 47 mainly correlated with brain regions involved in semantic processing (posterior inferior temporal gyrus). This study directly supports the existence of a functional gradient in LIFG. In our current study we were interested to see whether or not specific sub-regions of the LIFG would be involved in extracting information structure from pitch accents in Dutch.

In contrast to semantic and syntactic unification, phonological unification has, as far as we know, not yet been addressed specifically. Many studies, however, have implicated BA 44 and BA 6 in tasks involving phonology *per se*. Gough et al. (2005), for

instance, used a double dissociation paradigm to show that differential parts of the LIFG are involved in semantic and phonological processing. They applied Transcranial Magnetic Stimulation (TMS) over either the anterior or the posterior LIFG, and found that performance on a synonym judgment task was affected when TMS was applied over the anterior LIFG, while the performance on a homophone judgment task was affected when TMS was applied over the posterior LIFG (BA 44/6). A similar preference of the posterior LIFG for tasks involving phonological decisions was established by Devlin, Matthews, and Rushworth (2003), who found increased activation for phonological decisions (number of syllable judgment) relative to semantic (man-made or not) decisions in the posterior LIFG. For phonological versus semantic decisions on the same material, Price et al. (1997) found a left precentral activation close to the posterior LIFG for the phonological task. Studies involving rhyming tasks have also reported effects in BA 44/posterior LIFG (*e.g.*, Booth et al., 2006; Burton, Locasto, Krebs-Noble, & Gullapalli, 2005). As these studies concerned phonological tasks that targeted single word processing, it cannot be claimed that phonological unification at the sentence level was at stake in these experiments. It is clear though, that the posterior LIFG is important for phonological processing.

The study that we present here deviates from previous studies not only in that it focuses on sentence level processing, but also in investigating suprasegmental cues instead of segmental phonological features. This is inherent to the relevance of the phonological cues for information structure. To be able to identify the unification process related to prosodic marking of information structure in the brain, we constructed discourses in Dutch consisting of pairs of sentences which described colored shapes that were subsequently presented on the screen. The second sentence contained new information in either the adjective (color) or in the noun (shape) of a noun phrase; this new information was introduced by the colored shapes present on the screen. This new information is indicated by capitalization of the word, and it was always marked by means of a pitch accent in the recorded utterances. In (ii), triangle represents the new information relative to (i):

(i) *Now you see a blue circle*
(ii) *And now you see a blue TRIANGLE*
    We describe this situation as a *focus-accent agreement*, as the pitch accent is placed correctly on the word that represents the focus constituent (*i.e.*, new information as conveyed by the accompanying pictures). In our study, we included sentence-pairs such as (i) and (ii), with focus-accent agreement, as the condition for which unification of prosodic marking was relatively easy. To construct sentences for which unification of prosodic marking was more difficult, we assumed that a pitch accent placed on a word that provides old information is incompatible with the expected focus marking. A contextually incompatible pitch accent would be harder to integrate with the lexical information during the unification process, because the prosodic marking is not highlighting the most relevant, new information, but instead old and hence less relevant information.
    Under our assumption, an incompatible pitch accent will lead to more demands on the neuronal infrastructure for the unification of prosodic information that marks information structure. To create stimuli in which an incompatible pitch accent is present, we included sentence-pairs like (i) and (ii), except that in this condition the pitch accent in (ii) is not placed on the noun *triangle*, but on the adjective *blue* (*And now you see a BLUE triangle*). As blue is not the new information in the sentence, and therefore not the focus constituent, this accent is incorrectly placed; there is *focus-accent disagreement*. Similarly, (iii) and (iv) illustrate focus-accent agreement created by an accent on the adjective. In this case, the focus-accent disagreement is created

by a version of (iv) which has the pitch accent on the noun *circle*
(*And now you see a red CIRCLE*).

(iii) *Now you see a blue circle*

(iv) *And now you see a RED circle*

In our design, focus-accent disagreement (FAD) is the condition in which there is a *high* unification load for prosodic marking and focus-accent agreement (FAA) is the condition with a *low* unification load. By contrasting these conditions we were able to identify the neural correlates of unification processes related to prosodic marking.

Keeping in mind the evidence for a strong role of the LIFG in unification processes, we specifically tested the hypothesis that the LIFG is involved in the unification of prosodic marking. As the posterior part of the LIFG is implied in tasks involving phonological decisions, our hypothesis was further specified to the posterior LIFG as the most likely region to be involved.

2. Materials and methods

2.1. Participants

Sixteen volunteers aged 19–35 (8 men) were recruited from the Radboud University Nijmegen. All participants were native speakers of Dutch, right-handed, and reported no history of neurological or psychiatric disease, all of which was assessed by an extensive screening questionnaire prior to participation. All participants had normal or corrected to normal vision and reported normal hearing. Informed consent was obtained from participants prior to scanning. The study was approved by the local ethics committee, in accordance with the declaration of Helsinki.

2.2. Materials

We constructed 144 short, two-sentence Dutch discourses in which the first sentence had the structure *Now you see a* [*Adjective*] [*Noun*], and the second *And now you see a* [*Adjective*] [*Noun*]. Each Adjective–Noun combination described an object, which varied in color and shape. The first sentence and object of each discourse provided the context (*e.g.*, '*Now you see the green triangle*') and the second object varied from the first in either shape or color (*e.g.*, '*And now you see the blue triangle*'), such that either the shape noun or the color adjective was the focus constituent (new information) of the second sentences (target sentences). In 50% of the target sentences, the expected prosodic accent for new information was placed correctly, namely on the focus constituent of the sentence (shape noun or color adjective). In these sentences the non-focus word was not accented. In the other 50% of second sentences, the accent was placed incorrectly on the non-focus word in the sentences (shape noun or color adjective). In both the correctly and the incorrectly accented sets of sentences, the word *now* also had a pitch accent, which increased the naturalness of the discourses. We refer to the first condition as the focus-accent agreement condition (FAA), and to the second condition as the focus-accent disagreement condition (FAD).

Thirty-six items were constructed for each of the four experimental conditions, namely focus-accent agreement discourses for color and for shape change, and focus-accent disagreement discourses for color and for shape change; examples are provided in Table 1. All stimuli were presented using Presentation software (version 9.70, *Neurobehavioral Systems Inc.*, http://www.neurobs.com/). Objects were constructed in Presentation and consisted of arrows, squares, circles and triangles (Wingdings, font size 65) in the colors blue, green, yellow and brown. Stimuli were presented against a black full screen background. The display screen in the scanner measured 48×36.5 cm^2 and

was placed 60 cm from the participant (controlled by a Dell Pentium IV Windows XP computer, display mode 640×480 pixels at 60 Hz).

A male, native speaker of Dutch recorded all first sentences with pitch accents on both the adjective and the noun, while all second sentences were recorded twice, once with a pitch accent on the adjective but not on the noun, and once with a pitch accent on the noun, but not on the adjective. All pitch accents were realized as pitch peaks, H∗L in the ToDI system (for details see Gussenhoven, Rietveld, Kerkhoff, & Terken, 2003), which means that the pitch first rose (H=high tone) and then fell (L=low tone) to create the accent (∗). All sentences began and ended in low pitch, %L and L% in the ToDI system, where % indicates the start and end of an intonational phrase, respectively (Gussenhoven et al., 2003). We used a Sony digital audio data recorder at a 16 kHz audio sampling rate, edited the stimuli in Praat (Boersma & Weenink, 2004), and converted the stimuli to stereo using the Windows XP Sound Recorder. The first sentences of the pairs had an average duration of 1431 ± 54 ms; target sentences with the accent on the adjective had an average duration of 1455 ± 53 ms, while target sentences with an accent on the noun had an average duration of 1418 ± 52 ms. The same target sentences, either with a pitch accent on the noun or on the adjective, were used for both the focus-accent agreement and the focus-accent disagreement condition (the information provided in the first sentence determined whether there was focus-accent agreement or not). Therefore there was no difference in sentence duration between the FAA and FAD experimental conditions.

All sentences were grammatically correct. To control for unspecific language processing effects that might contaminate the results, the first sentences of all discourses (collapsed across conditions) were used as a baseline condition. In these sentences, all pitch accents were neutral and although unification processes will happen at the sentence level, suprasegmental unification effects were not expected for these sentences, rendering them suitable as a baseline condition.

A reversed speech condition in which no unification processes were possible was added as an additional baseline and consisted of 36 regular items (9 from each condition) for which the sentences were played backwards (for reversed speech results see Supplementary materials). Finally, we added 28 catch trials (15% of all stimuli). In the catch trials, one of the sentences did not correctly describe the object on the screen (*e.g.*, a green triangle would be displayed, and the mismatching sentence would either be '*And now you see a blue triangle*' or '*And now you see a green square*'). Participants had to press a button on recognizing these mismatches. This task was chosen such that it allowed us to ascertain that the participants were integrating the sentences with the visual information at all times, while not introducing any confounds of the task with the focus-accent agreement manipulation. Integration of the visual information with the sentences was crucial for the establishment of the new information in the sentence, as the new information was first conveyed by the pictures and then the pitch accent in the second sentence would either be in agreement with this focus constituent or be in disagreement.

2.3. Procedure

Prior to the experiment, all stimuli were pseudo-randomized (maximum of two repetitions for any stimulus characteristic) and randomly assigned to one of four runs of equal length. For each run, there was a minimum occurrence of 6 items and a maximum occurrence of 12 items per condition. Seven catch trials were inserted into each run at random locations, and all runs started with one start-up trial. During the first start-up trial subjects indicated whether the sentences were clearly audible or not. Prior to scanning, participants completed a practice set of 12 items containing all conditions. Participants were instructed to listen to the sentences and to respond only when the sentence did not match the object that was displayed. Response devices were the keyboard for the practice session and an MR-scanner compatible Lumitouch response box for the actual experiment. In the scanner, stimuli were presented to both ears through headphones.

2.4. Stimulus presentation

Each single trial started with a variable jitter period (0, 500, 1000 or 1500 ms) prior to stimulus onset, to allow variable sampling of the BOLD response. During

Table 1
Example stimuli for all four experimental conditions.

Condition	Example discourse	English translation
FA agreement; shape change (focus)	S1: Nu is de blauwe cirkel te zien.	S1: Now you see the blue circle.
	S2: En nu is de blauwe DRIEHOEK te zien.	S2: And now you see the blue TRIANGLE.
FA agreement; color change (focus)	S1: Nu is de blauwe cirkel te zien.	S1: Now you see the blue circle.
	S2: En nu is de GROENE cirkel te zien.	S2: And now you see the GREEN circle.
FA disagreement; shape change (focus)	S1: Nu is de blauwe cirkel te zien.	S1: Now you see the blue circle.
	S2: En nu is de BLAUWE driehoek te zien.	S2: And now you see the BLUE triangle.
FA disagreement; color change (focus)	S1: Nu is de blauwe cirkel te zien.	S1: Now you see the blue circle.
	S2: En nu is de groene CIRKEL te zien.	S2: And now you see the green CIRCLE.

Capitals denote the placing of the pitch accent. S1: first sentence; S2: second (target) sentence. FA: focus-accent.
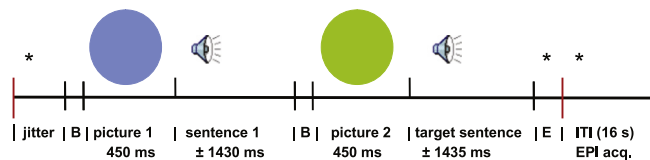
Fig. 1. Time course of a single trial. One trial lasted 6 s (silent period of 3 TRs). Jitter: 0, 500, 1000 or 1500 ms randomly assigned jitter period at start of trial. (B) 200 ms Blank screen. (E) Variable time at the end of a trial. ITI: inter-trial interval (16 s scanning period), during which EPI images were acquired. Red lines mark the end and beginning of image acquisition. Pictures remained on the screen until the sentences were completed. All stimulus presentation occurred during the 6 s silent periods. A fixation cross was displayed when no other stimuli were presented. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the jitter period a fixation cross was displayed. Then, the first picture was displayed and the sentence followed after this picture had been on the screen for 450 ms (see Fig. 1 for the set-up of one trial). The picture remained on the screen until the auditory sentence was completed, and was followed by a blank screen for 200 ms. Next, the second picture was presented, and the target sentence followed when this second picture had been on the screen for 450 ms. At the end of the target sentence, the picture disappeared and a fixation cross was displayed until the next trial. The inter-trial-interval was 16 s. Please note that the BOLD responses to the critical adjectives and nouns in the first sentences (baseline sentences) and the critical adjectives and nouns in the second sentences do not overlap in time. These critical words were always separated by minimally ~1800 ms, which is enough separation in time to sample the peaks of the BOLD response for each critical word separately (Huettel, 2012). This 1800 ms included 200 ms for the blank screen between stimuli and 450 ms for the presentation of the picture, as well as the time from the last possible critical word (noun) in the first sentence to the end of that sentence (± 1430–850 ms) and the time to the first possible critical word (adjective) in the target sentences (± 550 ms).

2.5. Image acquisition

MR data were acquired with a 3.0 T Siemens Magnetom Trio MR scanner. A single shot gradient echo-planar imaging (EPI) sequence was used to acquire functional MR images (36 slices, TE=35 ms, TR=2000 ms, flip angle=90°, 224 mm FOV, 64×64 matrix, 3.5×3.5 mm^2 voxel size, 3.0 mm slice thickness, 0.3 mm slice gap). Rapid gradient switching within the EPI sequence produces loud noise which potentially could impair the perception of the prosodic manipulations during stimulus presentation. We therefore adjusted the vendor supplied EPI sequence in such a way that stimulus presentation could occur while the scanner was not acquiring images for the duration of 3 TRs (6 s silent period). After each stimulus presentation, 8 consecutive volumes were acquired to read out the BOLD response (16 s scanning period). This interleaved silent steady state (ISSS) imaging with a slow event-related design yields a better signal-to-noise ratio (SNR) than conventional sparse imaging (Schwarzbauer, Davis, Rodd, & Johnsrude, 2006). During the silent period, steady state longitudinal magnetization was maintained by silent slice-selective excitation pulses, to avoid T1-related signal decay at the onset of each scanning period. Additional scanner noise reduction was obtained by switching off the fat suppression function (Fat Saturation) for the entire sequence. A pilot EPI run was performed to test the SNR of the sequence without fat suppression and no significant negative effects were found (the TR was significantly reduced without fat suppression). Furthermore the excitation pulse length was doubled to reduce the amplitude of the corresponding slice-selective gradient, and the slice refocusing gradient was played out over the entire available time before the readout rather than at maximum possible amplitude (for technical details we refer the reader to de Zwart, van Gelderen, Kellman, & Duyn, 2002).

The experiment consisted of four functional runs of approximately 20 min each. After the completion of two runs, a high-resolution T1-weighted structural image (MPRAGE, TE=3.93 ms, TR=2300 ms, 256 mm FOV, 256×256 matrix, 1 mm^3 isotropic resolution) was acquired for coregistration of the functional images to anatomical images. Participants were allowed to take a 10 min break outside the scanner before completing the last two runs. Atlas-based registration (AutoAlign, Siemens; Van der Kouwe et al., 2005) was applied for all EPI runs to ensure the same slice positions across all four sessions, before and after the break (AutoAlign corrects for differences in participant positioning between scans). During each functional session 432 whole-brain volumes were acquired, each consisting of 36 slices oriented along the AC–PC plane. The first four volumes of each session were discarded due to transient T1 saturation.

2.6. Data analysis

Data were preprocessed and analyzed with SPM2/SPM5 (Wellcome Department of Imaging Neuroscience, http://www.fil.ion.ucl.ac.uk/spm/software/). To

correct for head motion, all functional MR images were spatially realigned to the first image for each subject using a set of six rigid body transformations for each image. Functional images were then coregistered to the anatomical images and slice timing correction was applied. Normalization to the standard EPI template of SPM2 was performed to allow for group comparisons and the images were resampled to a $2 \times 2 \times 2$ mm^3 resolution. Finally all images were spatially smoothed using a 10 mm FWHM isotropic Gaussian filter.

Prior to model construction, the time points of all event onsets as acquired from our Presentation logfiles had to be adjusted. Since no MR images were acquired during the silent periods (3 TRs) in which the stimuli were presented, the total added time of the acquired and later concatenated MR images was less than the total experimental time as recorded in the logfiles. The event onsets obtained from the logfiles were therefore corrected for the number of silent periods that preceded the stimulus onset, to again align the presentation time of the stimuli with the time course of the acquired MR images. All events were included in the design matrix and the BOLD signal was modeled by the canonical haemodynamic response function (HRF) and its temporal derivative. The undershoot of the modeled HRF was re-parameterized to assure that it had returned to baseline after 16 s; we only acquired 8 MR images (16 s) after each stimulus presentation and later concatenated the MR images in time. By re-parameterizing the HRF we were able to compensate for the 6 s of stimulus presentation in which the BOLD response returned to baseline, but we were unable to sample this. Parameter estimates were obtained for each condition and each participant to allow second-level random effects analysis of between-subject variability. The six motion parameters obtained during preprocessing were included in the model as covariates. A high-pass filter (128 s cut-off) was used to remove low-frequency drifts.

The SPM [T] maps were thresholded corresponding to $P_{uncorrected} < 0.001$ and an extent threshold ≥ 10 voxels. Cluster-level inference was applied ($P_{FWE} < 0.05$). Coordinates of peak activity are reported in MNI coordinates in the order [x, y, z]. Corresponding brain regions and Brodmann areas were retrieved from the Talairach Daemon database server (Lancaster et al., 1997) and verified using the Anatomy Toolbox of SPM5 (Eickhoff et al., 2005).

Because we were particularly interested in the LIFG, we took on a region of interest approach. Because we were additionally hoping to see which subregion of the LIFG would be involved in our phonological task, three different Regions of Interest (ROIs) were specified in the LIFG; one ROI was located in the anterior-ventral part, one in the posterior-dorsal part of LIFG, and one ROI was located centrally in between the two extremes. Roughly speaking, the posterior-dorsal ROI comprises part of Brodmann areas (BA) 44 and 6, the anterior-ventral ROI includes BA 47 and parts of BA 45, and the more 'central' ROI comprises BA 44 and 45, although considerable overlap exists between the three ROIs. ROIs were defined as 15 mm spheres around the following centers of mass (MNI coordinates): [x, y, z]= [−46, 10, 29] (posterior-dorsal), [−44, 19, 14] (central), and [−42, 26, 6] (anterior-ventral), see also Fig. 2A. Centers of mass for each ROI were taken from Petersson et al. (2004), who used the studies reported in a meta-analysis by Bookheimer (2002) to determine subregions within the LIFG based on functional distinctions in the tasks used (semantic versus syntactic versus phonological aspects of language, when moving from anterior-ventral to posterior-dorsal LIFG). Although there was substantial variability in the tasks as well as spatial variability in the reported activation sites, a rough posterior-dorsal to anterior-ventral functional distinction could be made in the LIFG. Since Petersson et al. (2004) showed that there was rather high spatial variability in the exact location of activation, each of the three ROIs in our study was defined as a sphere with 15 mm radius.

To test for significantly activated sites within the ROIs, we used Small Volume Correction as implemented in SPM5, ensuring the results are corrected for multiple comparisons within each ROI. Results that reach a cluster level significance level of $P_{FWE} < 0.05$ are reported. Additional statistical analyses were performed with SPSS (SPSS 14.0 for Windows).

3. Results

The behavioral results (catch trial detection rates and reaction times) are summarized in the Supplementary materials.

3.1. ROI analysis

The main contrast, which we used to investigate the effects of focus-accent disagreement on unification processes, was the contrast between the focus-accent disagreement conditions (difficult to integrate) and the focus-accent agreement conditions (relatively easy to integrate). We conducted ROI analyses in the LIFG for this contrast. The results of the region of interest analyses are listed in Table 2 and shown in Fig. 2A. For the posterior-dorsal ROI, there was a significant cluster of activation (p=0.040 corrected cluster level) with its peak activation at coordinates [−46, −4, 28]. The
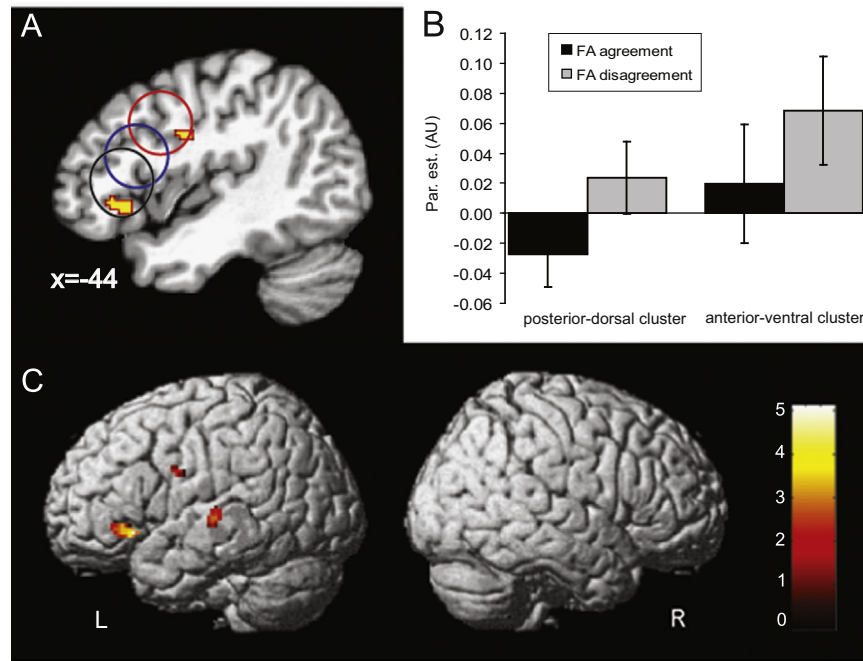
Fig. 2. Processing of focus-accent disagreement. (A) Sagittal plane through the LIFG showing the positions of the posterior-dorsal (red), central(blue) and anterior-ventral (black) ROIs, with significant activations derived from the ROI analyses (see Table 2). (B) Plots of the average parameter estimates for the FA agreement and FA disagreement conditions within the cluster at [−46, −4, 28] in the posterior-dorsal ROI, and the cluster at [−32, 28, −4] in the anterior-ventral ROI (see Table 2). Values are normalized with respect to the baseline stimuli (normal sentences). Error bars depict the SEM. (C) Left hemisphere (L) and right hemisphere (R) surface plots of whole brain activation map (random effects, 16 subjects, $P_{uncorrected} < 0.001$, extent threshold of 10 voxels). For coordinates of local maxima see Table 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Results of ROI analyses for the contrast of focus-accent disagreement versus focus-accent agreement. We tested for significant (P_{FWE} cluster level corrected for multiple comparisons) clusters within each 15 mm spherical ROI. The ROIs were centered on [−46, 10, 29], [−42, 26, 6] and [−44, 19, 14] for the posterior-dorsal, anterior-ventral, and central ROI, respectively (no suprathreshold clusters were found for the central ROI). Whole brain threshold was set to $P_{uncorrected} < 0.001$ and extent threshold 10 voxels. The p-value (cluster-level FWE-corrected), T-value and MNI coordinates of the local maxima are listed, as well as the size of the corresponding cluster (k). L=left, R=right, IFG=inferior frontal gyrus, PrecG=precentral gyrus. Brodmann areas (BA) are in parentheses.

ROI	Location	Cluster level p-value[*] (FWE-corr.)	k	MNI coordinates (x, y, z)	T-value
Posterior-dorsal	L PrecG/IFG (BA 6/44/9)	0.040	23	−46, −4, 28	4.59
Anterior-ventral	L IFG (BA 47)	0.002	165	−32, 28, −4	5.05
	L IFG (BA 45)			−50, 22, −6	3.80

[*] Please note that when applying Bonferroni correction for multiple testing in 3 ROIs, the p-values are p=0.12 for the posterior-dorsal cluster and p=0.006 for the anterior-ventral cluster.

cluster of activation corresponded to the location of BA 6 in the posterior LIFG, close to the precentral gyrus. For the anterior-ventral ROI, there was a significant (p=0.002) cluster of activation with two local maxima located at coordinates [−32, 28, −4] and [−50, 22, −6]. This cluster was located in BA 47 and BA 45 in the LIFG. For the centrally located ROI, there were no supra-threshold activations.

We set out to further characterize the effects of focus-accent agreement and disagreement within the two clusters obtained from the ROI analysis. We calculated the parameter estimates within each cluster for both of the experimental conditions and for the baseline. Only voxels that belonged to the actual clusters of activation were included (not the entire ROI spheres were used). The parameter estimate values were normalized with respect to the baseline, which

consisted of normal sentences (the first sentences of all trials, collapsed across conditions), by subtracting the parameter estimate values of the baseline activation from the parameter estimate values for the FAA and FD conditions. Reversed speech was not used as a baseline, since the contrast between normal sentences and reversed speech did not reveal a larger effect in language related areas for normal sentences (see Section 3.2.2, Fig. 3, and Table 4 for details). In Fig. 2B, the averaged parameter estimates across all subjects are plotted for the cluster in the posterior-dorsal LIFG, with a maximum at [−46, −4, 28], and for the cluster in the anterior-ventral LIFG, with maxima at [−34, 28, −4] and [−50, 22, −6]. Subjecting the parameter estimate values for each cluster to statistical analyses did not yield any evidence for an interaction between cluster location and focus-accent agreement ($F(1,15) < 1$). Please see the Supplementary material for detailed statistical analyses.

3.2. Whole brain analysis

3.2.1. Focus-accent disagreement and focus-accent agreement conditions

To check if there were any effects of focus-accent disagreement in other brain regions than the LIFG, we conducted a whole brain analysis for the contrast of focus-accent disagreement versus focus-accent agreement. The results are presented in Fig. 2C and Table 3. The main clusters that were activated more strongly for the FAD condition is located in the LIFG and the left superior temporal gyrus, but none of the clusters reached significance at the $P_{FWE} < 0.05$ cluster level (family-wise error corrected for multiple comparisons).

We also investigated whether the processing of focus-accent agreement recruited any different brain regions than processing of focus-accent disagreement. We therefore investigated the contrast of focus-accent agreement versus focus-accent disagreement (FAA versus FAD). First we performed an ROI analysis in the LIFG analogous to the ROI analysis for the FAD versus FAA contrast, which did not yield any significant effects. We then performed a
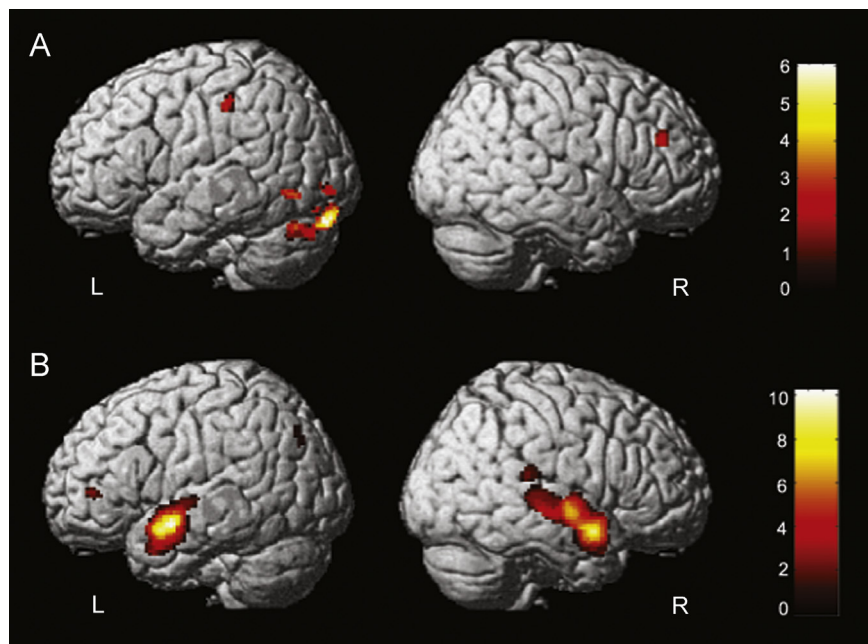
Fig. 3. Surface plots of contrasts involving reversed speech. Random effects analysis, 16 subjects, whole brain analysis thresholded at $P_{\text{uncorrected}} < 0.001$ and an extent threshold of 10 voxels. Coordinates of local maxima are listed in Table 4. Left (L) and right (R) hemisphere surface plots of (A) the baseline (normal sentences) condition versus the reversed speech condition and (B) the reversed speech versus the baseline condition.

Table 3

Results of whole brain analyses. For both contrasts, clusters from the whole brain analysis at $P_{\text{uncorrected}} < 0.001$ and an extent threshold of 10 voxels are reported. Cluster size (k), T-value and MNI coordinates of the local maxima are listed. Brodmann areas (BA) are in parentheses.

Contrast	Location	k	MNI coordinates			T-value
			x	y	z	
FAD > FAA	L STG (BA 22/41)	57	−60	−24	4	5.10
	L IFG (BA 47)	181	−32	28	−4	5.05
	L IFG (BA 45)		−50	22	−6	3.80
	L PrecG/IFG (BA 6/44/9)	27	−46	−4	28	4.59
	R STG (BA 22)	13	46	−26	−10	4.15
FAA > FAD	R precuneus (BA 7/31)	256	16	−54	42	5.77*
	L precuneus (BA 7)		−16	−48	46	4.06
	Precuneus/cingulate (BA 7/31)		0	−52	40	4.06
	R SFG (BA 10)	36	16	62	26	5.02
	L caudate	26	−14	8	16	4.66
	R STG/MTG (BA 39)	67	56	−64	22	4.50
	L MFG (BA 9)	40	−26	40	38	4.47
	R MFG (BA 8/9)	80	30	30	44	4.38

Abbreviations: FAD=focus-accent disagreement, FAA=focus-accent agreement, L=left, R=right, STG=superior temporal gyrus, IFG=inferior frontal gyrus, PrecG=precentral gyrus, SFG=superior frontal gyrus, MTG=middle temporal gyrus, MFG=middle frontal gyrus.

* Clusters significant at cluster level $P_{\text{FWE}} < 0.05$.

whole brain analysis to look for other regions that were possibly more active in the focus-accent agreement condition. The results are summarized in Table 3. The only region that was significantly (at cluster level $P_{\text{FWE}} < 0.05$) more activated for focus-accent agreement compared to focus-accent disagreement was the precuneus, bilaterally, located in the parietal lobes (BA 7/31). Fig. S1 provides details of the activation levels of the precuneus.

3.2.2. Processing of normal sentences and reversed speech

Reversed speech was included in our design to serve as a baseline, as we hypothesized that there would be no unification

processes possible for reversed speech. Reversed speech has been used extensively as a baseline condition for auditory language studies (*e.g.*, Price et al., 1996; Roder et al., 2002; Strand, Forssberg, Klingberg, & Norrelgen, 2008); for a review discussing the use of reversed speech as a baseline condition see Stoppelman, Harpaz, and Ben-Shachar (2013). We investigated the pattern of activation for this condition. First, we contrasted the normal sentence condition versus the reversed speech condition, to identify areas involved in regular sentence processing. All analyses were done at a whole brain threshold of $P < 0.001_{\text{unc}}$ with an extent threshold of 10 voxels. In this contrast however, there were no significant activations within any of the three language ROIs in the LIFG. The whole brain analysis (see Table 4 and Fig. 3A) revealed significant clusters of activation (at cluster level $P_{\text{FWE}} < 0.05$) in the left lingual gyrus (BA 18) and bilaterally in the cingulate cortex/supplementary motor area (BA 32/6), but not in areas that are specialized in language processing. We then looked at the effects of reversed speech by contrasting the reversed speech condition against the normal sentence condition. There were no significant effects within any of the ROIs in the LIFG. The whole brain analysis (see Table 4 and Fig. 3B) revealed significant clusters (cluster level $P_{\text{FWE}} < 0.05$) in left middle/superior temporal gyrus (BA 21/38) and in right middle/superior temporal gyrus (BA 21/22).

4. Discussion

In this study we investigated the neural correlates of the extraction and integration of information structure during sentence comprehension in Dutch. We were specifically interested in how the prosodic cues (pitch accents) that signal the listener about focus constituents are combined with other types of information (*e.g.*, lexical items) during the unification process. The difficulty of integrating the prosodic cues into an overall sentence meaning was manipulated by constructing sentences in which the pitch accent *did* (focus-accent agreement), and sentences in which the pitch accent *did not* (focus-accent disagreement) match the expected focus constituents of the sentence.

Table 4

Results of whole brain analysis for contrasts involving reversed speech. For both contrasts, clusters from the whole brain analysis at $P_{uncorrected} < 0.001$ and an extent threshold of 10 voxels are reported. Cluster size (k), T-value and coordinates of the local maxima are listed. Brodmann areas (BA) are in parentheses.

Contrast	Location	k	MNI coordinates			T-value
			x	y	z	
NS > RS	L lingual gyrus (BA 18)	189	−22	−90	−16	6.15
	R MeFG/SMA (BA 32/6)	506	10	10	48	5.22*
	L SMA/cingulate cortex (BA6/32)		−4	4	48	4.90
	Cingulate cortex (BA 32/24)		0	10	40	4.89
	L IOG (BA 17/18)	156	−24	−88	−2	5.21
	L MOG (BA 18)		−26	−80	2	5.14
	L MOG (BA 18)		−38	−78	−2	3.80
	L cerebellum	89	−36	−68	−26	5.10
	L cerebellum		−36	−76	−30	4.32
	L IPL (BA 40)	93	−40	−32	44	4.72
	R anterior cingulate cortex (BA 32)	47	8	40	14	4.69
	R MFG (BA 10/46)	59	42	40	24	4.64
	L ITG (BA 37)	61	−50	−70	−6	4.47
RS > NS	L MTG/STG (BA 21/38)	1143	−48	0	−16	9.92*
	R STG/MTG (BA 38/21)	1528	46	4	−18	8.57*
	R MTG/STG (BA 21/22)		64	−8	−4	8.08
	R STG/MTG (BA 22/21)		70	−30	2	5.20
	R STG (BA 42/22)	20	68	−32	18	4.75
	L IFG/MFG (BA 46/10)	22	−46	44	6	4.24
	L amygdala	17	−22	−6	−20	4.17
	L precuneus/angular gyrus (BA 19/39)	21	−46	−74	30	4.03
	L angular gyrus/precuneus (BA39/7)		−40	−74	38	3.83

Abbreviations: NS = normal sentences, RS = reverse speech condition, L = left, R = right, MeFG = medial frontal gyrus, SMA = supplementary motor area, MOG = middle occipital gyrus, IOG = inferior occipital gyrus, IPL = inferior parietal lobule, MFG = middle frontal gyrus, ITG = inferior temporal gyrus, MTG = middle temporal gyrus, STG = superior temporal gyrus, IFG = inferior frontal gyrus.

* Clusters significant at cluster level $P_{FWE} < 0.05$.

Based on the previous literature, the LIFG was a region of special interest in our study, and we were additionally interested to see which subregion of the LIFG would be involved in phonological unification. Our results show two anatomically distinct effects of focus-accent disagreement in the LIFG; one is located in the anterior-ventral part of the LIFG (BA 47/45), and another effect is located in the posterior-dorsal LIFG (BA 6/44).

In the focus-accent disagreement condition, the prosodic cues and expectations about the information status of constituents do not match and are therefore harder to combine into the overall interpretation of the sentence; there is a higher unification load. Our results confirm that more processing resources in the brain are needed in the focus-accent disagreement condition, and that this recruits the LIFG. That the LIFG is activated in cases of higher unification load is a finding that is in line with previous fMRI studies (*e.g.*, Burholt Kristensen et al., 2012; Hagoort et al., 2004; Willems, Ozyurek, & Hagoort, 2007).

The fact that we see two anatomically distinct effects in the LIFG is especially interesting. One effect is located in BA 6 and BA 44, in the posterior LIFG, and is not so strong; another effect is located more anteriorly in BA 47/45. Both LIFG subregions are involved in the specific unification processes taking place in our paradigm, in which phonological cues indicate semantically relevant focus constituents of the sentence: the phonological manipulation has consequences for the semantic processing of the sentence. We can discuss whether the two subregions we see activated might be subserving exactly the same language processes. Considering previous literature, the effect in the posterior LIFG is consistent with studies in which posterior LIFG is found to be active during phonological processing, mainly in tasks targeting single word phonology (*e.g.*, Devlin et al., 2003; Gough et al., 2005). In our stimuli, it is not simply the phonological properties of single words that are manipulated, but prosodic cues that have an implication for the information structure of the sentence. Information structure has been shown to be relevant during on-line sentence processing (Cowles et al., 2007). We can show that the posterior LIFG, along with more anterior parts, is involved in solving this suprasegmental phonological task. On the basis of our results, we can tentatively conclude that the prosodic features in language that are relevant for extracting information structure recruit the posterior LIFG.

The other, clearly distinct effect that we find in the LIFG is located in a more ventral subregion, in BA 47/45. When we look at the pattern of activations that we find for the FAD versus FAA manipulation in the whole brain analysis, we also see that the superior temporal gyrus (STG) is activated. Although this activation does not reach our whole brain threshold for significance, activity in STG is seen often alongside activation in the ventral LIFG in tasks involving semantic unification (for a review, see Hagoort et al., 2009). The ventral part of the LIFG has been implicated in many studies in which semantics were manipulated. Several previous studies have found effects in BA 45/47 for semantic mismatches (Willems et al., 2007) or ambiguous stimuli (*e.g.*, Hagoort et al., 2004; Rodd et al., 2005, for a review see Hagoort et al., 2009), and the anterior-ventral LIFG was found to be involved in semantic tasks most often in a meta-analysis by Bookheimer (2002). It is to be expected that semantic processing is affected in our paradigm. During the presentation of the stimuli, expectations about the information structure of the upcoming sentence are built up as soon as the second picture is displayed on the screen (see Fig. 1), which is prior to the moment when the target sentence is actually heard. Thus, the context of the target sentence and the expectations about its information structure are already built up at the moment when the (wrong) pitch accent is presented. When there is a mismatch between the pitch accent and the expected focus of the sentence, additional processing is required to (re)interpret the sentence correctly, as previous expectations are not met. This expectation violation can be regarded as a semantic mismatch, analogous perhaps to an N400 effect that occurs when expectations are violated (Hagoort et al.,

2009). Altogether, previous literature suggests that the activation in ventral LIFG could be interpreted as a consequence of the increase in semantic unification load that occurs in our paradigm.

When interpreting our results we have to keep in mind that in our study, prosodic cues and expectation violations are strongly intertwined, and since these aspects were manipulated together we cannot disentangle them in our design. Our data are consistent with an interpretation of involvement of the posterior LIFG in phonological unification and the more ventral parts of LIFG in semantic unification. As such, they are consistent with the existence of a functional gradient of language processing in LIFG. However, our paradigm cannot provide direct support for a functional distinction within LIFG. It could also be that both subregions are involved in both aspects of unification in our task.

Alternative accounts of the organization of LIFG point to the role that the LIFG plays in executive functions and selection processes that are related to language processing (Novick, Trueswell, & Thompson-Schill, 2005; Thompson-Schill et al., 1997; Thothathiri et al., 2012). Thompson-Schill et al. (1997) have argued that the LIFG performs more general selection processes that lead to the selection of the correct alternative among competing sources of information. For instance in Thothathiri et al. (2012) the authors parametrically manipulated syntactic–semantic conflict and find evidence that LIFG activation (BA44/45) is increased for increasing conflict, in line with a role for LIFG in mediating between conflicting representations. Also, Novick, Kan, Trueswell, and Thompson-Schill (2009) find evidence that patients with focal lesions to BA44/45 are impaired in situations where syntactic, conceptual and/or semantic information compete and that they are not able to resolve the task in these conflicting situations. In our task, one could argue that in the process of unification in the FAD condition, selection processes are crucial since the correct alternative of multiple possible semantic interpretations of the sentences have to be selected from the conflicting sources of context (picture) and phonological (accent) information. However, as we have argued elsewhere (Hagoort et al., 2009) selection is part of a unification process, since the outcome of selection is a unified interpretation of the input string. In this context, LIFG is not claimed to be language-specific, but operations of selection and unification get their domain specificity from the interaction between LIFG and areas that subserve the domain-specific knowledge.

The BOLD effects we find in the LIFG are not very strong; especially the cluster in the phonological ROI is not large, which can be accounted for by a number of different factors. First of all, in our design, the focus-accent disagreement stimuli are rather frequent, as they make up 50% of all focus-accent items and 35% of the total number of experimental items when the reversed speech and catch trial items are included. The high frequency of the disagreement items might attenuate the increased BOLD response to these stimuli, as adaptation and familiarization to the manipulation is likely to take place. The fact that only four different colors and shapes were used throughout the experiment will have added to the repetitive nature of the stimuli.

Another factor that may have contributed to the relatively small disagreement effects is the long inter stimulus interval (ISI) of our slow event-related design. This slow design was necessary because we adopted a sparse imaging design to allow for stimulus presentation without interference from the scanner noise. In a sparse design no images are acquired during stimulus presentation and the BOLD response is allowed to return to baseline before the next stimulus is presented, so that it can be fully sampled (Schwarzbauer et al., 2006). This leads to a reduction in power compared to a conventional event-related design, which may make it more difficult to detect BOLD effects in the brain.

For the focus-accent agreement condition, we reported increased activity in the precuneus. The precuneus is an important area in the so-called default mode network (DMN) (Raichle et al.,

2001). This network is most strongly activated in the absence of external stimulation, and it has been suggested that it is related to internally driven mental activity. The stronger the processing requirements of stimulus events are, the larger the deactivation of the DMN, although the amount of deactivation in different areas that make up the DMN may depend on the precise task requirements at hand (Seghier & Price, 2012). In the current study, we see a stronger deactivation in the precuneus for the FAD condition than for the FAA condition (illustrated in Fig. S2 in the Supplementary material), which is in line with previously reported deactivation of precuneus in semantic processing and perceptual language processing (Binder & Desai, 2011; Seghier & Price, 2012).

Reversed speech was originally included in our experimental design to serve as a baseline condition. Previous literature on reversed speech (Stoppelman et al., 2013) led us to expect additional activity in language related areas for the comparison of normal versus reversed speech, since for normal speech, language comprehension processes are taking place. However, we found no evidence of such effects. The contrast of normal versus reversed speech did not yield strong activity in frontal or temporal areas that are normally involved in language comprehension. Instead, reversed speech led to strong activation of the superior and middle temporal gyri (Fig. 3 and Table 4). Several properties of our study may have caused the reversed speech stimuli to fail as a proper baseline condition. The first aspect is the multimodal nature of our experimental design. During all trials, pictures were present on the screen, which were being described by the auditory sentences. The relevance of the visual content for the auditory sentences was stressed even more because of the behavioral task, *i.e.*, the detection of mismatching auditory and visual information. Our experiment, therefore, cannot be regarded as a purely linguistic experiment where only language stimuli are present; instead, a visual context has to be matched with the presentation of the auditory stimuli. The second relevant aspect of our experimental design is the highly repetitive nature of the stimuli which generates, together with the visual displays, a very high predictability for the auditory stimuli. In all of the normal speech trials the auditory sentences described the picture that was displayed on the screen, and there were only four different shapes and four different colors present in the pictures. Importantly, the pictures were always displayed prior to the presentation of the auditory sentences, in order to shape a context for the auditory information. Hence, a fair amount of repetition suppression might have occurred. Our experimental design was slow by nature, the stimuli were repetitive and highly predictable from the visual context, and the reversed speech trials were relatively infrequent (~20% of trials) and therefore may have strongly drawn the attention of the participants. Taking into account these factors, we propose that the pictures on the screen may have invited the participants to strongly engage their attention in order to understand the contents of the reversed speech stimuli and/or to recognize words and/or speech sounds. After all, the sentences were describing the picture on the screen. This situation is very different from other language studies in which reversed speech was presented as a baseline stimulus without accompanying visual stimuli. Moreover, in these other experiments, the auditory stimulus content was much more varied throughout the experiments, *i.e.*, the linguistic content of the sentences in those experiments was less predictable prior to hearing the actual auditory stimulus (Stoppelman et al., 2013).

The enhanced temporal activity for reversed speech points to auditory processing of the stimuli. In line with our findings, a recent meta-analysis finds more activation in auditory cortex when participants listen to relatively meaningless sounds compared to meaningful speech (Skipper, in press). Also, a recent study measuring brain activity with intracranial electrocorticography

reports that reversed speech trials activated temporal and frontal cortices more strongly than normal speech trials (Brown et al., 2012). The authors suggest that increased attention for reversed speech may have yielded this pattern of enhanced engagement of frontal and temporal regions. Their findings illustrate, similar to our results, that reversed speech may not be suitable as a baseline condition in all situations.

In summary, this study shows that two anatomically distinct subregions in the LIFG are recruited during the processing of phonological markers that indicate focus constituents of the linguistic input as more (*i.e.*, new information) or less (*i.e.*, old information) semantically relevant. Our results are consistent with a crucial role for the LIFG in suprasegmental phonological unification processes.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuropsychologia. 2014.03.017.

References

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, 15, 527–536.

Boersma, P., & Weenink, D. (2004). *Praat: Doing phonetics by computer.* ⟨http://www.praat.org⟩.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience*, 25, 151–188.

Booth, J. R., Lu, D., Burman, D. D., Chou, T. L., Jin, Z., Peng, D. L., et al. (2006). Specialization of phonological and semantic processing in Chinese word reading. *Brain Research*, 1071, 197–207.

Bornkessel-Schlesewsky, I., Grewe, T., & Schlesewsky, M. (2012). Prominence vs. aboutness in sequencing: A functional distinction within the left inferior frontal gyrus. *Brain and Language*, 120, 96–107.

Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage*, 26, 221–233.

Brown, E. C., Muzik, O., Rothermel, R., Matsuzaki, N., Juhasz, C., Shah, A. K., et al. (2012). Evaluating reverse speech as a control task with language-related gamma activity on electrocorticography. *NeuroImage*, 60, 2335–2345.

Burholt Kristensen, L., Wang, L., Petersson, K. M., & Hagoort, P. (2012). The interface between language and attention: Prosodic focus marking recruits a general attention network in spoken language comprehension. *Cerebral Cortex*, ⟨http://dx.doi.org/10.1093/cercor/bhs1164⟩.

Burton, M. W., Locasto, P. C., Krebs-Noble, D., & Gullapalli, R. P. (2005). A systematic investigation of the functional neuroanatomy of auditory and visual phonological processing. *NeuroImage*, 26, 647–661.

Costafreda, S. G., Fu, C. H. Y., Lee, L., Everitt, B., Brammer, M. J., & David, A. S. (2006). A systematic review and quantitative appraisal of fMRI studies of verbal fluency: Role of the left inferior frontal gyrus. *Human Brain Mapping*, 27, 799–810.

Cowles, H. W., Kluender, R., Kutas, M., & Polinsky, M. (2007). Violations of information structure: an electrophysiological study of answers to wh-questions. *Brain and Language*, 102, 228–242.

Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron*, 24, 427–432.

Davis, M. H., Coleman, M. R., Absalom, A. R., Rodd, J. M., Johnsrude, I. S., Matta, B. F., et al. (2007). Dissociating speech perception and comprehension at reduced levels of awareness. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 16032–16037.

de Zwart, J. A., van Gelderen, P., Kellman, P., & Duyn, J. H. (2002). Reduction of gradient acoustic noise in MRI using SENSE-EPI. *NeuroImage*, 16, 1151–1155.

Devlin, J. T., Matthews, P. M., & Rushworth, M. F. (2003). Semantic processing in the left inferior prefrontal cortex: A combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, 15, 71–84.

Doherty, C., West, W., Dilley, L., Shattuck-Hufnagel, S., & Caplan, D. (2004). Question/statement judgments: An fMRI study of intonation processing. *Human Brain Mapping*, 23, 85–98.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, 25, 1325–1335.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11, 11–15.

Fiebach, C. J., Schlesewsky, M., Lohmann, G., von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Mapping*, 24, 79–91.

Friederici, A. D., Opitz, B., & von Cramon, D. Y. (2000). Segregating semantic and syntactic aspects of processing in the human brain: An fMRI investigation of different word types. *Cerebral Cortex*, 10, 698–705.

Gough, P. M., Nobre, A. C., & Devlin, J. T. (2005). Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. *Journal of Neuroscience*, 25, 8010–8016.

Grewe, T., Bornkessel, I., Zysset, S., Wiese, R., von Cramon, D. Y., & Schlesewsky, M. (2005). The emergence of the unmarked: A new perspective on the language-specific function of Broca's area. *Human Brain Mapping*, 26, 178–190.

Gussenhoven, C. (2008). Notions and subnotions in information structure. *Acta Linguistica Hungarica*, 55, 381–395.

Gussenhoven, C., Rietveld, T., Kerkhoff, J., & Terken, J. (2003). *Transcription of Dutch Intonation; Courseware.* ⟨http://todi.let.ru.nl⟩.

Hagoort, P. (2005). On Broca, brain and binding: A new framework. *Trends in Cognitive Sciences*, 9, 416–423.

Hagoort, P., Baggio, G., & Willems, R. (2009). Semantic unification. In: M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (4th ed.). MIT Press USA.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304, 438–441.

Hagoort, P., & Indefrey, P. (2014). The neurobiology of language beyond single words. *Annual Review Neuroscience*, 37(1).

Heim, S., & Alter, K. (2006). Prosodic pitch accents in language comprehension and production: ERP data and acoustic analyses. *Acta Neurobiologiae Experimentalis*, 66, 55–68.

Heim, S., Eickhoff, S. B., & Amunts, K. (2008). Specialisation in Broca's region for semantic, phonological, and syntactic fluency? *NeuroImage*, 40, 1362–1368.

Heim, S., Eickhoff, S. B., & Amunts, K. (2009). Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modelling. *NeuroImage*, 48, 616–624.

Hesling, I., Clément, S., Bordessoules, M., & Allard, M. (2005). Cerebral mechanisms of prosodic integration: Evidence from connected speech. *NeuroImage*, 24, 937–947.

Hesling, I., Dilharreguy, B., Clément, S., Bordessoules, M., & Allard, M. (2005). Cerebral mechanisms of prosodic sensory integration using low-frequency bands of connected speech. *Human Brain Mapping*, 26, 157–169.

Hruska, C., & Alter, K. (2004). How prosody can influence sentence perception. In: A. Steube (Ed.), *Information structure: Theoretical and empirical evidence.* (pp. 211–226). Berlin: Walter de Gruyter.

Huettel, S. A. (2012). Event-related fMRI in cognition. *NeuroImage*, 62, 1152–1156.

Humphries, C., Love, T., Swinney, D., & Hickok, G. (2005). Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Human Brain Mapping*, 26, 128–138.

Indefrey, P. (2004). Hirnaktivierungen bei syntaktischer Sprachverarbeitung: eine Meta-analyse. In: H. M. Mueller, & S. Rickheit (Eds.), *Neurokognition der Sprache* (pp. 31–50). Tübingen: Stauffenberg Verlag.

Johnson, S. M., Breen, M., Clifton, C. J. R., & Morris Florak, J. (2003). ERP investigation of prosodic and semantic focus. In *Paper/poster presented at the cognitive neuroscience*, New York City.

Kotschi, T. (2006). Information structure in spoken discourse. In: K. Brown (Ed.), *Encyclopedia of language & linguistics* (pp. 677–683). Oxford: Elsevier.

Kristensen, L. B., Engberg-Pedersen, E., Nielsen, A. H., & Wallentin, M. (2013). The influence of context on word order processing – An fMRI study. *Journal of Neurolinguistics*, 26, 73–88.

Kuperberg, G. R., Sitnikova, T., & Lakshmanan, B. M. (2008). Neuroanatomical distinctions within the semantic system during sentence comprehension: Evidence from functional magnetic resonance imaging. *NeuroImage*, 40, 367–388.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203–205.

Lancaster, J. L., Rainey, L. H., Summerlin, J. L., Freitas, C. S., Fox, P. T., Evans, A. C., et al. (1997). Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. *Human Brain Mapping*, 5, 238–242.

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, 293–303.

Mesulam, M. -M. (2002). The human frontal lobes: Transcending the default mode through contingent encoding. In: D. T. Stuss, & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 8–31). Oxford: Oxford University Press.

Meyer, M., Alter, K., Friederici, A., Lohmann, G., & von Cramon, D. (2002). Functional MRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping*, 17, 73–88.

Meyer, M., Steinhauer, K., Alter, K., Friederici, A., & von Cramon, D. (2004). Brain activity varies with modulation of dynamic pitch variance in sentence melody. *Brain and Language*, 89, 277–289.

Miller, J. (2006). Focus. In: K. Brown (Ed.), *Encyclopedia of language & linguistics* (pp. 511–518). Oxford: Elsevier.

Novick, J., Trueswell, J., & Thompson-Schill, S. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 263–281.

Novick, J. M., Kan, I. P., Trueswell, J. C., & Thompson-Schill, S. L. (2009). A case for conflict across multiple domains: Memory and language impairments

following damage to ventrolateral prefrontal cortex. *Cognitive Neuropsychology*, *26*, 527–567.

Petersson, K. M., Forkstam, C., & Ingvar, M. (2004). Artificial syntactic violations activate Broca's region. *Cognitive Science*, *28*, 383–407.

Plante, E., Creusere, M., & Sabin, C. (2002). Dissociating sentential prosody from sentence processing: Activation interacts with task demands. *NeuroImage*, *17*, 401–410.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage*, *10*, 15–35.

Price, C. J., Moore, C. J., Humphreys, G. W., & Wise, R. J. S. (1997). Segregating semantic from phonological processes during reading. *Journal of Cognitive Neuroscience*, *9*, 727–733.

Price, C. J., Wise, R. J. S., Warburton, E. A., Moore, C. J., Howard, D., Patterson, K., et al. (1996). Hearing and saying: The functional neuro-anatomy of auditory word processing. *Brain*, *119*, 919–931.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 676–682.

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, *15*, 1261–1269.

Rodd, J. M., Johnsrude, I. S., & Davis, M. H. (2012). Dissociating frontotemporal contributions to semantic ambiguity resolution in spoken sentences. *Cerebral Cortex*, *22*, 1761–1773.

Rodd, J. M., Longe, O. A., Randall, B., & Tyler, L. K. (2010). The functional organisation of the fronto-temporal language system: Evidence from syntactic and semantic ambiguity. *Neuropsychologia*, *48*, 1324–1335.

Roder, B., Stock, O., Neville, H., Bien, S., & Rosler, F. (2002). Brain Activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: A functional magnetic resonance imaging study. *NeuroImage*, *15*, 1003–1014.

Ruschemeyer, S. -A., Zysset, S., & Friederici, A. D. (2006). Native and non-native reading of sentences: An fMRI experiment. *NeuroImage*, *31*, 354–365.

Schwarzbauer, C., Davis, M. H., Rodd, J. M., & Johnsrude, I. S. (2006). Interleaved silent steady state (ISSS) imaging: A new sparse imaging method applied to auditory fMRI. *NeuroImage*, *29*, 774–782.

Seghier, M. L., & Price, C. J. (2012). Functional heterogeneity within the default mode network during semantic processing and speech production. *Frontiers in Psychology*, http://dx.doi.org/10.3389/fpsyg.2012.00281.

Skipper, J. (2014). Echoes of the spoken past: How auditory cortex "hears" context during real-world speech processing. *Philosophical Transactions of the Royal Society B: Biological Sciences* (in press).

Snijders, T. M., Vosse, T., Kempen, G., Van Berkum, J. J. A., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: An fMRI study using word-category ambiguity. *Cerebral Cortex*, *19*, 1493–1503.

Stoppelman, N., Harpaz, T., & Ben-Shachar, M. (2013). Do not throw out the baby with the bath water: Choosing an effective baseline for a functional localizer of speech processing. *Brain and Behavior*, *3*, 211–222.

Strand, F., Forssberg, H., Klingberg, T., & Norrelgen, F. (2008). Phonological working memory with auditory presentation of pseudo-words – An event related fMRI study. *Brain Research*, *1212*, 48–54.

Swerts, M., Krahmer, E., & Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics*, *30*, 629–654.

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

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*, 259–264.

Toepel, U., & Alter, K. (2004). On the independence of information structure processing from prosody. In: A. Steube (Ed.), *Information structure: Theoretical and empirical evidence*. (pp. 227–240). Berlin: Walter de Gruyter.

Van Berkum, J. J. A., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research*, *17*, 701–718.

Van der Kouwe, A. J., Benner, T., Fischl, B., Schmitt, F., Salat, D. H., Harder, M., et al. (2005). On-line automatic slice positioning for brain MR imaging. *NeuroImage*, *27*, 222–230.

Wang, L., Hagoort, P., & Yang, Y. F. (2009). Semantic illusion depends on information structure: ERP evidence. *Brain Research*, *1282*, 50–56.

Willems, R. M., Ozyurek, A., & Hagoort, P. (2007). When language meets action: The neural integration of gesture and speech. *Cerebral Cortex*, *17*, 2322–2333.

Xiang, H. D., Fonteijn, H. M., Norris, D. G., & Hagoort, P. (2010). Topographical functional connectivity pattern in the Perisylvian language networks. *Cerebral Cortex*, *20*, 549–560.

Zempleni, M. Z., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *NeuroImage*, *34*, 1270–1279.

Zhang, L., Shu, H., Zhou, F., Wang, X., & Li, P. (2010). Common and distinct neural substrates for the perception of speech rhythm and intonation. *Human Brain Mapping*, *31*, 1106–1116.

Zhu, Z. D., Feng, G. Y., Zhang, J. X., Li, G. C., Li, H., & Wang, S. P. (2013). The role of the left prefrontal cortex in sentence-level semantic integration. *NeuroImage*, *76*, 325–331.