NeuroImage 202 (2019) 116047

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Discourse management during speech perception: A functional magnetic resonance imaging (fMRI) study

Susanne Dietrich [a,*], Ingo Hertrich [b], Verena C. Seibold [a], Bettina Rolke [a]

[a] *Evolutionary Cognition, University of Tübingen, Germany*
[b] *Hertie Institute for Clinical Brain Research, University of Tübingen, Germany*

ARTICLE INFO

Keywords:
Monitoring
Presupposition
Reference
Evaluation
Accommodation
Context-dependent

ABSTRACT

Discourse structures enable us to generate expectations based upon linguistic material that has already been introduced. We investigated how the required cognitive operations such as reference processing, identification of critical items, and eventual handling of violations correlate with neuronal activity within the language network of the brain. To this end, we conducted a functional magnetic resonance imaging (fMRI) study in which we manipulated spoken discourse coherence by using presuppositions (PSPs) that either correspond or fail to correspond to items in preceding context sentences. Definite and indefinite determiners were used as PSP triggers, referring to (non-) uniqueness or (non-) existence of an item. Discourse adequacy was tested by means of a behavioral rating during fMRI. Discourse violations yielded bilateral hemodynamic activation within the inferior frontal gyrus (IFG), the inferior parietal lobe including the angular gyrus (IPL/AG), the pre-supplementary motor area (pre-SMA), and the basal ganglia (BG). These findings illuminate cognitive aspects of PSP processing: (1) a reference process requiring working memory (IFG), (2) retrieval and integration of semantic/pragmatic information (IPL/AG), (3) cognitive control of inconsistency management (pre-SMA/BG) in terms of "successful" comprehension despite PSP violations at the surface. These results provide the first fMRI evidence needed to develop a functional neuroanatomical model for context-dependent sentence comprehension based on the example of PSP processing.

1. Introduction

This study investigated which brain structures are involved in the processing of discourse coherence and incoherence. We manipulated discourse coherence by using presupposition (PSP) phrases, which adequately or inadequately refer to discourse content. PSPs inherit an implicit assumption about a common ground (context or knowledge) within a discourse. They are triggered by, for example, adverbs such as "again", which presuppose that something has already happened previously, verbs such as "win", which assume that somebody has played, or the definiteness of determiners such as "the", which presuppose the existence (Kirsten et al., 2014; Tiemann et al., 2011) or uniqueness of an item (Heim and Kratzer, 1998). For example, "Tina observed that *the* polar bear was very aggressive" presupposes the existence of a single (unique) polar bear within the context of the sentence. In general, PSP triggers are used to mark linguistic materials as referring to something "given" by means of linguistic or non-linguistic (e.g., facial expression) background information considered as common ground among participants in a conversation (Stalnaker, 2002). If a preceding context provides this information, the recipient obtains a coherent representation of the discourse (Garrod and Sanford, 1982) and successfully integrates the current semantic content into the context (Van Berkum et al., 2003). Thus, a PSP avoids redundant repetition of information by replacing referents, for example, with pronouns or by using determiners for referencing to particular known contextual items. In the sense of Grice (1975), PSPs can be used in order to optimize communication in order to keep a message informative, brief, and relevant (for a further discussion on presuppositions and Gricean reasoning see Schlenker, 2012).

Several cognitive processes are assumed to contribute to PSP processing. As a first step, the trigger-context reference has to be disambiguated and a relation between the trigger and the presupposed information has to be established. In case of a mismatch between the PSP trigger and the context, for example, a missing or inadequate referent, additional operations are assumed to be required, thereby further increasing the cognitive load in a recipient (Domaneschi and Di Paola, 2018). These operations could, for example, be rejecting the discourse

---

* Corresponding author.
E-mail address: s.dietrich@uni-tuebingen.de (S. Dietrich).

https://doi.org/10.1016/j.neuroimage.2019.116047
Received 20 December 2018; Received in revised form 9 July 2019; Accepted 22 July 2019

meaning in case of unacceptable errors, guessing the discourse meaning, or modifying the (assumed) context in order to obtain an acceptable discourse meaning. The latter process has been termed “accommodation” and means that even though a PSP was used in an inadequate manner, the violation can be repaired post hoc, by integrating the assumed meaning into a given context (Beaver and Zeevat, 2007; von Fintel, 2008; Simons, 2003). This process, however, presumably takes place in rather specific violations with a basically suitable context (e.g., “Inge had *never* bought red gloves until now. Today, *Susanne* bought red gloves *again* and put them on right away.” word-by-word translation from German to English), whereas other violations that leave the discourse un-interpretable (e.g., “*Susanne* had *never* bought red gloves until now. Today, *Susanne* bought red gloves *again* and put them on right away.”) are assumed to result in a falsification and rejection of the sentence meaning (Tiemann et al., 2011). All in all, even though the function of PSPs constitutes a fostering of discourse understanding, PSP usage increases cognitive processing load, and its misuse should result in comprehension difficulties.

The increased effort in PSP processing and, specifically, processing difficulties in case of violations of the discourse structure caused by an inadequate use of PSPs can be investigated by means of behavioral paradigms (e.g., self-paced reading tasks) or eye-tracking. Self-paced reading experiments have shown that processing of PSP triggers (i) differs from that of other sentence items not containing a PSP and (ii) starts quite early after the trigger has occurred (Tiemann et al., 2011). Further, self-paced reading experiments have shown that contextually implausible PSPs lead to longer reading times than plausible PSPs (Singh et al., 2016). Finally, the results of eye-tracking experiments indicate that the cognitive effort required for processing PSPs is influenced by whether the PSP trigger is presented in an embedded or unembedded environment. For example, Schwarz and Tiemann (2017) presented context sentences being either felicitous (e.g., “Tina went ice-skating for the first time, the weather was beautiful”) or infelicitous (e.g., “Tina wanted to go ice-skating for the first time, but the weather was miserable”) followed by target sentences including either an embedded PSP trigger (e.g., “This weekend, was Tina *again not* ice-skating”) or an unembedded trigger (e.g., “This weekend, was Tina *not again* ice-skating”). Reading times on the verb following the PSP trigger were found to be significantly higher when the trigger was unembedded (“not again”) under negation (infelicitous) than embedded or unembedded combined with a felicitous context. Furthermore, acceptability ratings for embedded triggers were higher than for unembedded triggers. The authors concluded that processing effort corresponds to the level of representational steps that have to be taken to arrive at a final interpretation: Subjects could interpret either that Tina had not been ice-skating before or that Tina had been ice-skating sometime recently. Furthermore, from the higher acceptability ratings in the latter context, the authors concluded that subjects prefer this kind of accommodation. Thus, subjects appeared to re-interpret the meaning of target sentence in a broader context in terms of “successful” comprehension despite PSP violations.

Another source of evidence pointing to processing difficulties in response to PSP violations are neurophysiological studies. In electrophysiological studies, it has been shown that enhanced processing load for the processing of PSPs (“*The* girl”) compared to assertions (“There was *a* girl”) is reflected in an enhanced posterior-parietal negativity after 400 ms (N400; Masia et al., 2017). Moreover, context-mismatching definite determiners presupposing the existence and uniqueness of an item not given in the context (Anderson and Holcomb, 2005; Heim, 1982; Krahmer, 1998) have been shown to elicit a N400/P600 complex immediately upon reading the PSP (Kirsten et al., 2014), indicating difficulties in word integration at a discourse-related semantic/pragmatic processing level (Van Berkum et al., 1999b). Moreover, the inadequate use of the definite determiner in the context of ambiguous anaphors has been found to elicit a frontal negativity after 400 ms (Nref) that is typically associated with referential ambiguity (Van Berkum et al., 1999a, 2003). Finally, in a magnetoencephalography (MEG) study addressing PSPs of determiner phrases in a spoken language mode, it has been shown that violations of discourse structure lead to an early latency effect on the M50 deflection as well as an increase in the amplitude of the M200 (Hertrich et al., 2015). These early modulations have been explained as being due to anticipatory top-down operations during auditory-phonetic processing in terms of sensory gating (M50) and phonological encoding (M200).

There have been very few fMRI studies on context-dependent speech processing, documenting even fewer neuroanatomical correlates of cognitive processes underlying PSP processing. One example is the study by Robertson et al. (2000). The authors showed right hemispheric inferior frontal activation during presentation of sentences containing definite determiners, while sentences containing indefinite determiners yielded left hemispheric inferior frontal activation. The results suggest that processing the definite determiner in the context of discourse coherence requires right-hemispheric pragmatic memory structures. Taken together, there are several theoretical assumptions concerning PSP processing as well as empirical evidence that PSP processing, especially processing of PSP violations, causes cognitive processing costs (for an overview see e.g., Schwarz, 2007, 2016). However, the relation between the theoretically proposed processes (e.g., accommodation), the contribution of different cognitive processes (e.g., reference process, falsification/verification, repair mechanisms) and the involvement of different brain areas (e.g., frontal vs. posterior-parietal, right hemisphere) in the processing of PSP are currently far from being deciphered.

The present study aims to shed more light on how spoken discourse is managed in the brain by investigating which neuroanatomical structures of the language network are functionally involved in PSP processing. We aimed to disentangle functional areas engaged in the cognitive processes underlying PSP processing such as reference processing, evaluation of plausibility, or handling of mismatch (accommodation or rejection). When considering the anatomical regions which might play a role in discourse processing, probable candidate regions are cortical regions surrounding the lower bank of the posterior sylvian fissure, that is, superior temporal gyrus (STG) and middle temporal gyrus (MTG) extending to posterior supra-marginal and angular gyrus (AG), as well as frontal areas comprising premotor, prefrontal, and inferior frontal cortex (Amunts et al., 2010). These areas have been found to be involved in processing of phonetic/phonological information as well as processing of more abstract representations of linguistic structures in terms of syntax and semantics (see Binder, 2017, for a review).

Our assumptions concerning the contribution of specific brain areas on PSP processing are based on the idea that specified modules of the brain contribute to speech processing (Friederici and Gierhan, 2013; Hickok, 2012; Hickok and Poeppel, 2007; Rodriquez-Fornells et al., 2009; Saur et al., 2010). Within this framework, speech signals are first perceived in the primary auditory cortex, are then categorized into phonemic features in the STG, and are later mapped to phonological representations in the anterior STG and the rolandic operculum. Further, abstract lexico-semantic representations are activated in the MTG and the temporal pole (Tp) which are integrated into a broader semantic/pragmatic context in the next step to foster discourse understanding. A well-suited candidate region for this last stage of representing conceptual meaning and semantic/pragmatic content is the AG localized in the posterior inferior parietal lobule (IPL). This area has been characterized as integrative for comprehension and reasoning, for example, when manipulating conceptual knowledge, reorienting the attentional system towards relevant information, retrieving facts for problem solving, or giving meaning to external events based on stored memories and prior experiences (see Seghier, 2013, for a review). Supporting this integrative view, the AG strongly interacts with left hemispheric MTG (Seghier, 2013), which has been suggested to store and provide semantic information and lexical features, too (Martin and Chao, 2001).

All these brain areas are functionally intertwined with the inferior frontal gyrus (IFG) in which speech functions are assumed to be integrated to understand higher-order language representations like sentences and discourses (Camos and Barrouillet, 2014). The left hemispheric (pre-) frontal cortex, including the IFG, has also been proposed to be involved in executive functions (Miller and Cohen, 2001) as well as in working memory functions, meaning that information is maintained, memory becomes updated, and conflicts are resolved (Baddeley, 2003; Zhang et al., 2003). Because speech processing constitutes active rehearsal of relevant information, executive functions such as detection of errors, encoding, storing, maintaining, and updating of working memory (Kessler and Oberauer, 2015; Morris and Jones, 1990) seem to be essential to perform the task in this study. The above-mentioned functions of the described areas argue for their essential contribution to discourse understanding and, thus, we assume that they will be found active during PSP processing.

Recently, and as an extension to these "classical language areas" described above, the pre-supplementary motor area (pre-SMA), a part of the medial frontal cortex, has been reported to be involved in various superordinate executive functions (Adank, 2012a, 2012b). The pre-SMA is linked by a large fiber bundle (the aslant tract) to the IFG (Kim et al., 2010). This fiber bundle has been found to be strongly marked in the language-dominant left hemisphere, but less so in the right hemisphere (Dick et al., 2014; Thiebaut de Schotten et al., 2012; Vassal et al., 2014). The strength of the connection in the language-dominant hemisphere supports the view that the pre-SMA might be important with respect to speech/language processing (Hertrich et al., 2016). Specifically, the pre-SMA contributes to fluent execution of speech and pre-articulatory repair, which becomes apparent in speech production (Alm, 2011). In addition, it seems to be involved in speech perception. For example, Dietrich et al. (2018) have shown that errors during speech perception increased under difficult perceptual circumstances when the pre-SMA was temporally impaired by transient virtual lesions. The authors proposed that the disturbed function of the pre-SMA led to an impairment of error management when speech perception became difficult. Finally, another function of the pre-SMA in discourse processing seems to be the temporal coordination between memory retrieval and the continuing speech sequence (Kotz and Schwartze, 2010). In sum, in order to guarantee fluent speech perception, including both error management and temporal coordination of memory processes, the pre-SMA is hypothesized to be yet another node within the network managing discourse processing.

To investigate the contribution of the described candidate areas/network in processing discourse incoherencies, we measured hemodynamic responses by means of fMRI while listeners heard pairs of context and test sentences comprising either matching or mismatching samples. Specifically, the test sentences included definite determiners ("the"), presupposing uniqueness and existence (Heim, 1982; Anderson and Holcomb, 2005; Krahmer, 1998) or indefinite determiners ("a"), suggesting non-uniqueness or novelty of an item (Alonso-Ovalle et al., 2011; Chemla, 2008; Kirsten et al., 2014; Napoli, 2013). The mismatching conditions using the definite determiner were constructed by using it when many different discourse entities had been introduced in the context (uniqueness violation), or by using it to refer to a negated discourse entity (existence violation). Likewise, the mismatching conditions using the indefinite determiner were constructed by using it with a contextually-introduced unique protagonist (violation of the anti-uniqueness PSP and the novelty violation). We focused primarily on the violation of discourse structure, here termed the mismatching effect. Our idea was that differences between processing a coherent discourse structure and processing an incoherent one should expose the brain structures that are specifically engaged in processing discourses in general.

Even though we assumed that participants would notice each of these violations, we expected that they would experience these violations in a graded manner. Specifically, violations of (non-) uniqueness might mainly affect the experience of morpho-syntactic accuracy. Thus, although errors might be salient (mismatch in number), plausibility of interpretation (i.e., comprehension) is unlikely to be endangered, or, at least, a meaningful interpretation might be accommodated. In contrast, violation of the existence PSP represents an irreparable error because the trigger refers to an explicitly negated item in the context. Moreover, we assume that the violation of novelty might be perceived as inaccurate because the trigger presupposes a non-existing entity, but one entity had already been mentioned in the context. In this situation, the common ground (context) of the discourse must be extended beyond the local sentence context and several potential interpretations must be considered by the listener to reach a plausible discourse interpretation. This, however, might be possible under these additional processing constraints.

In sum, we hypothesize that several cognitive processes are involved in processing a discourse containing PSPs and in reaching a plausible interpretation of the discourse. According to the literature mentioned above, these cognitive processes should show up in the activation of distinct brain areas: verbal working memory processes are expected to be relevant for binding a PSP trigger to previously presented information (initial reference process). These working memory processes should activate the IFG. Further, identification of a plausible referent and its contextual integration is expected to be a function of the IPL/AG (advanced reference). Finally, in order to obtain a coherent and intelligible representation of the respective sentence (i.e., comprehension), reference processes and the evaluation of plausibility have to be temporally coordinated by a superordinate structure, which we assume is represented by the pre-SMA. Since these three regions are thought to be part of a network relevant to discourse processing, we expect (H1) that all three regions (at least in the left hemisphere) will be activated during processing of both coherent and violated discourse structures (independent of the experimental condition). Considering the mismatching effect (mismatching vs. matching sentence pairs), we expect different activation patterns across experimental conditions, since – as mentioned above – discourse violations are expected to be managed differentially. Generally, we expect mismatching effects within IFG across all conditions (H2) due to increased cognitive effort in the initial part of the reference process, during the detection of the PSP trigger. We further expect (H3) mismatching effects within the IPL/AG across all conditions, because the protagonist triggered by the PSP seems not to be plausible across all conditions in the context and, thus, higher effort is required for contextual integration. Further (H4), we expect significant mismatching effects within the pre-SMA, but only under conditions allowing for re-analysis of discourse violation. This is because we assume that the pre-SMA controls the executive system (comprehension) adjusting for the long-lasting process of re-interpretation (re-reference, evaluation of plausibility). To give an overview of the results and the contributions of particular brain areas in discourse processing, we will summarize our understanding of discourse processing by means of a neuroanatomical model.

2. Methods

2.1. Participants

Twenty adult volunteers participated in the experiment (7 males; mean age = 31.2, SD = 11.58 years). With one exception, all participants were right-handed (Edinburgh handedness inventory, mean laterality index: 82.9, SD = 38.10) native German speakers. None of them had any signs of neurological or psychiatric disorders (information had to be drawn from personal interviews). Participants were compensated with 10 Euro for 60 min. All subjects provided written informed consent prior to the fMRI measurements, and the experimental procedures were approved by the ethics committee of the Medical Faculty of the University of Tübingen.

2.2. Stimulus materials

Materials comprised pairs of spoken German context and test sentences that were presented via headphones and had been used as stimuli in a previous study (see Hertrich et al., 2015). We employed two different types of subsets, requiring either processing of the uniqueness PSP of the definite determiner (subset 1) or processing of the existence PSP (subset 2).

In subset 1 (uniqueness = U), the context sentence introduced either a single item (uniqueness) indicated by a singular noun phrase or multiple items (non-uniqueness) indicated by the plural combined with a quantifier.

(1) Manuel saw a television program on ZDF about *a* species of dolphin in the Pacific Ocean.
(2) Manuel saw a television program on ZDF about *several* species of dolphin in the Pacific Ocean.

The subsequently presented test sentence always contained a corresponding singular phrase starting with either a definite (D) or an indefinite (I) determiner.

(3) Manuel noticed that *the* species of dolphin was especially shy.
(4) Manuel noticed that *a* species of dolphin was especially shy.

This resulted in four possible combinations. There were two matching conditions: singular context (1) paired with definiteness (3), presupposing uniqueness and plural context (2) paired with indefiniteness (4), presupposing non-uniqueness. Likewise, there were two mismatching conditions: plural (2) paired with definiteness (3), violating uniqueness and singular (1) paired with indefiniteness (4), violating non-uniqueness.

In subset 2 (existence = E) the context introduced either a single item or – by negation – a non-existing item.

(1) Tina has *a* swimming badge.
(2) Tina has *no* swimming badge.

The test sentence either referred to an exisiting item by the use of the definite determiner (existence PSP) or introduced a new item by using the indefinite determiner in combination with an existence-creating verb such as “build” or “buy” (novelty PSP).

(3) When Tina showed *the* swimming badge in the swimming course, she got cramps in her calves.
(4) When Tina earned *a* swimming badge in the swimming course, she got cramps in her calves.

There were again four possible combinations. Matching conditions combined an existing item (1) with the definite determiner (3) or a non-existing item (2) with the indefinite determiner (4). In the non-matching conditions, the definite determiner (3) referred either to a non-existing item (2), or the indefinite determiner (4) was paired with an already introduced item (1).

The sentence material comprised 160 pairs of sentences including two subsets (uniqueness and existence) of 20 quadruplets each. Sixteen practice sentence pairs, similar to the test sentences, were constructed to familiarize participants with the experimental procedure. Each sentence pair was fluently uttered by a female speaker with consistent prosodic modulation, in order to obtain a natural speech flow without any discontinuity effects (pitch height, voice quality, speech rate; see Hertrich et al., 2015). The duration of the sentence pairs varied between 7 and 10 s. Evaluation of speech rate and consonant articulation (voice onset time) did not show any significant differences between context-matching and context-violating cognates (see Hertrich et al., 2015). The sentence pairs were presented via headphones (Sennheiser HD 570, binaural stimulus application, adapted for MRI measurements by removal of their permanent magnets).

2.3. Procedure

The course of a single trial is depicted in Fig. 1. During a trial, subjects directed their eyes to a red fixation cross in front of them while listening to the context sentence, which was followed by the test sentence. Participants were instructed to listen carefully to the sentences and to respond concerning the pragmatic adequateness of test sentences relative to context sentences when the fixation cross had switched to a green color (about 2 s after test sentence offset). They rated their impression by means of a four-point scale (1 = bad, 2 = somewhat bad, 3 = somewhat good, 4 = good). This task was introduced to keep the participants attentive while listening and to obtain a measure of the subjective coherence or incoherence of the sentence pairs. Behavioral responses were recorded by means of button presses on a MRI-compatible device. Buttons were pressed with the index and middle finger of the subjects’ left hand (this was also true for the left-handed subject): (1) double click of finger one = “bad”, (2) single click of finger one = “somewhat bad”, (3) single click of finger two = “somewhat good”, and (4) double click of finger two = “good”. The assignment of the two fingers to the rating scale was balanced across participants. Maximum response time amounted to 4 s.

2.4. MRI data acquisition

The functional imaging sessions were organized in 10 runs of 16 sentence pairs each in order to provide short pauses for the subjects. Additionally, 20 silent baseline intervals (null events, scanner noise) were added to the stimulus materials in order to obtain unmodeled event types. Hemodynamic responses were modeled at PSP onset (epoch-related with a duration of 2.5 s, Fig. 1). The test materials were distributed across the ten runs (16 sentence pairs and 2 null events per run) and presented in a pseudo-randomized order, which was constrained in order to prevent more than one sentence pair of a quadruplet being presented within the same run. Each run comprised eight sentence pairs of each subset, uniqueness (U) and existence (E), with an equal number of indefinite (I) and definite (D) determiners and an equal number of matching and mismatching items. Thus, the following four conditions occurred, each in a matching and mismatching manner: uniqueness definite (UD), uniqueness indefinite (UI), existence definite (ED), existence indefinite (EI). Mean ITI was 4 s (following the response interval). In order to increase the temporal resolution of hemodynamic responses with respect to the scan repetition time, a jitter (nine steps by 178 ms) was induced, shifting stimulus presentation. Because the headphones provided sufficient dampening of environmental scanner noise, it was not necessary to provide the subjects with earplugs during the experiment. Prior to the experimental functional imaging session, all subjects performed a practice run (16 practice sentence pairs plus 2 null events) in order to get acquainted with the task. Loudness of stimulus presentation

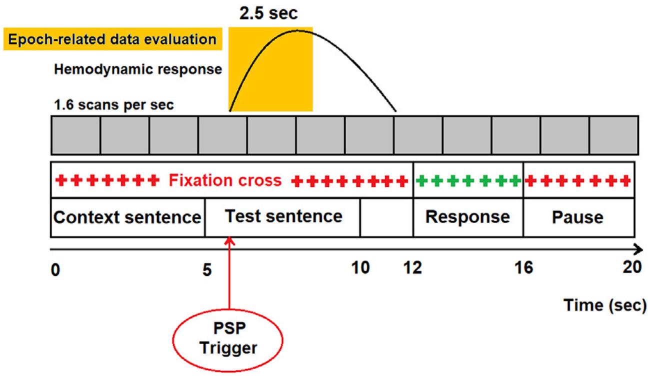

Fig. 1. Time course of stimulus presentation and fMRI data recording. The arrow indicates the onset of the presupposition (PSP) trigger within the test sentence. The duration of the sentence pairs varied between 7 and 10 s. Thus, the time point of the PSP onsets differed between sentence pairs. FMRI data were continuously recorded with a repetition time of 1.6 s. Hemodynamic responses were modelled at PSP onset (epoch-related with a duration of 2.5 s).

was adjusted until the subjects reported that they could comfortably listen to the sentences.

The experiment was run on a 3 T MRI system (PRISMA, Siemens, Erlangen, Germany), using an echo-planar imaging sequence (echo-time = 30 ms, 64×64 matrix with a resolution of $3 \times 3\,\text{mm}^2$, 27 axial slices across the whole brain volume, repetition time = 1.6 s, slice thickness = 4 mm, flip angle = 90°, 235 scans per run). The scanner generated a constant background noise throughout fMRI measurements, serving as the baseline condition of the experimental design (null event). Anatomical images required for the localization of the hemodynamic responses were obtained by means of a GRAPPA sequence (T1-weighted images, repetition time = 2.3 s, echo-time = 2.92 ms, flip angle = 8°, slice thickness = 1 mm, resolution = $1 \times 1\,\text{mm}^2$) of a bi-commissural (anterior-posterior commissure) orientation.

2.5. Data analyses

Analysis of behavioral data included calculation of individual mean scores on the rating scale for each of the eight conditions. A repeated measure analysis of variance (rmANOVA) was applied with respect to the within subject factors CONGRUENCE (match/mismatch), SUBSET (uniqueness/existence), and DEFINITENESS (definite/indefinite determiner). For post-hoc pairwise comparisons of significant interactions, we performed pairwise *t*-tests with *p*-values corrected according to Bonferroni Holm. In order to investigate acceptability ratings associated with potential repair mechanisms in mismatching conditions, we calculated medians of rating data of each mismatch condition and grouped participants into ten low-raters (values < median) and ten high-raters (values > median). Further, ratings of each subgroup (high-, low-rater) were averaged in order to test whether the mean is different from the value of two in the four-point scale (i.e., “somewhat bad”) indicating either “bad” or “somewhat good” judgements (one-sample *t*-test, Bonferroni Holm corrected, eight tests).

Preprocessing of the fMRI data encompassed slice time and motion correction, normalization to the Montreal Neurological Institute (MNI) template space, and smoothing by means of an 8 mm full-width half maximum Gaussian kernel (SPM8 software package; http://www.fil.ion.ucl.ac.uk/spm). For the sake of statistical analysis, the blood oxygen level-dependent (BOLD) responses were modeled by means of a prototypical hemodynamic function within the context of a general linear model (event durations 2.5 s). Any low-frequency temporal drifts were removed using a 128 s high-pass filter.

The analysis of the functional imaging data encompassed the following steps:

(i) In order to estimate which brain areas were generally involved in auditory language processing at the level of discourse understanding, we computed a whole-brain analysis based on the average effect of all conditions (UD, UI, EI, ED of both, match and mismatch) contrasted each to the baseline (null event).

(ii) In a second step, we investigated effects of the experimental manipulations (UD, UI, ED, EI) on brain activation by means of a factorial three-way rmANOVA including the within-subject factors CONGRUENCE (match/mismatch), SUBSET (uniqueness/existence), and DEFINITENESS (definite/indefinite). Again, we used the contrast between events and baseline (null event) in a whole-brain analysis. In order to show the direction of the effects, we computed SPM *T*-contrasts between the two levels of each factor.

(iii) Third, we performed regions of interest (ROIs) analyses with the goal of depicting the activation pattern across experimental conditions on the brain regions hypothesized to be relevant for speech (IFG, IPL, pre-SMA). The most common approach for exploratory ROI analyses is to create small spheres at the peaks of activation clusters (Poldrack, 2007). Friston et al. (2006), however, highly recommended embedding the ROI localization in factorial designs that includes orthogonal comparisons of interest, and Poldrack and Mumford (2009) suggested alpha adjusting when ROI selection is performed on the basis of whole brain data. We followed these recommendations, using an orthogonal factor design and correcting for multiple comparisons (FWE). Moreover, by using the conjunction of the three main effects, any bias of particular main effects was removed. The location of the relevant clusters was validated by anatomical masks resulting from AAL2 atlas information (Tzourio-Mazoyer et al., 2002). Peak-coordinates from hemodynamic responses inside these masks were taken to define the center of spherical ROIs using a 6 mm radius around the peak-coordinate. Since we hypothesized increased cognitive load in processing incoherent discourse structures, we focused on mismatching effects by subtracting matching conditions from mismatching conditions. To investigate the contribution of different ROIs to the mismatching effect, one-sample *t*-tests were applied testing whether activation was significantly different from zero. Bonferroni Holm corrections for four tests were applied separately for each ROI. Further, we investigated effects of the experimental conditions (UD, UI, ED, EI) on the mismatching effect by means of a factorial two-way rmANOVA including the within-subject factors SUBSET (uniqueness/existence) and DEFINITENESS (definite/indefinite). Again, we used the contrast between events and baseline (null event) as the dependent variable.

(iv) To gain an insight into brain areas responsible for violation repair, we conducted an additional analysis and correlated (using Pearson’s *r*) acceptability ratings of each of the four experimental conditions (UD, UI, ED, EI) with the hemodynamic mismatching effects of the left and right hemispheres, using an adjusted alpha level of $p < .006$ for the first step (2 hemispheres x 4 conditions). In order to make the rating scale (1 = “unacceptable”) comparable to the scale of hemodynamic responses (0 = no percent signal change) the logarithm of the subjects’ mean rating across trials for each mismatch condition was used. Based on significant Pearson’s correlation, whole-brain covariance analyses were conducted for those experimental conditions showing an effect in at least one of the ROIs. For these analyses, log-transformed acceptability ratings were co-varied with the hemodynamic mismatching responses.

3. Results

3.1. Behavioral data

Fig. 2 displays the acceptability ratings for each experimental condition, and Table 1 lists the statistical effects. As expected, violations of discourse coherence (mismatching conditions) yielded significantly lower acceptability ratings compared to matching pairs (CONGRUENCE, $F(1, 19) = 367.47$, $p < .0001$, $\eta_p^2 = 0.95$). This congruence main effect was accompanied by several two- and three-way interactions (Table 1). Since we were primarily interested in the mismatching effect, the most important interactive pattern (CONGRUENCE × SUBSET × DEFINITENESS) was revealed by the comparison of the four different mismatching conditions (alpha adjusted $p < .008$ for the first step). Acceptability in the mismatching existence definite condition (ED) was lower than in mismatching uniqueness definite (UD, $T = 3.44$, $p_{\text{Bonf}} = .011$), mismatching uniqueness indefinite (UI, $T = 4.42$, $p_{\text{Bonf}} = .002$), and mismatching existence indefinite (EI, $T = 5.96$, $p_{\text{Bonf.}} < 0.001$), but no further differences were found (all $p > .09$). The grouping in high- and low-raters and the comparison to the rating value of two (comparison of the four different mismatching conditions of the two groups, alpha adjusted $p < .006$ for the first step), showed that low-raters ($n = 10$) judged all mismatching conditions as being unacceptable (all $p_{\text{Bonf.}} < 0.0001$), while high-raters ($n = 10$) evaluated the (non-) uniqueness indefinite and definite conditions (UI, UD) as “less unacceptable” ($p > .10$). Furthermore, high-raters judged the indefinite determiner of the existence subset (EI) as more than “less unacceptable” ($p_{\text{Bonf}} = .025$) going on the scale in the direction of

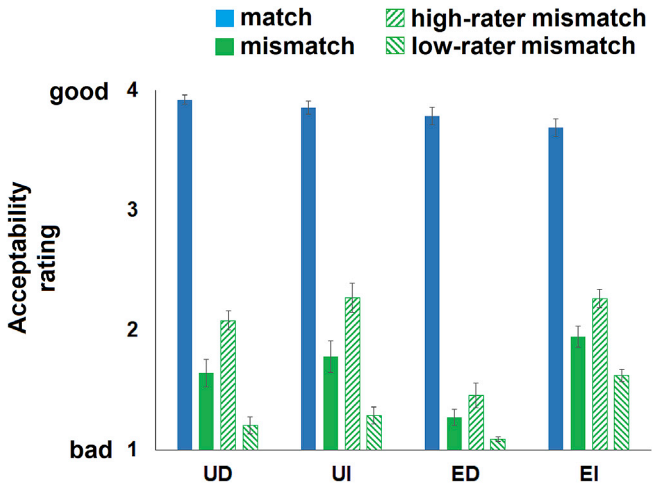

Fig. 2. Acceptability ratings for each experimental condition averaged across subjects and separately for matching (blue) and mismatching (green) conditions. In the mismatching condition, mean rating values and standard error of the mean are additionally split into high- and low-raters. UD = definite determiner in the uniqueness subset, UI = indefinite determiner in the uniqueness subset, ED = definite determiner in the existence subset, EI = indefinite determiner in the existence subset.

"somewhat good". By contrast, high-raters judged the definite determiner of the existence subset (ED) as clearly not "less unacceptable" ($p_{Bonf} = .002$) going on the scale in the direction of "bad".

3.2. FMRI data

3.2.1. Whole-brain analyses

Average effect of conditions Hemodynamic responses ($p < .001$, $k = 10$, FWE corr.) averaged across mismatching and matching conditions (UD, UI, ED, EI versus null events) emerged within primary auditory areas of both hemispheres as well as adjacent structures of the superior/middle temporal cortex and temporal poles (see Fig. 3, Table 2). Additionally, participants showed activation within left- and right-hemispheric inferior frontal and premotor regions, the left pre-SMA, the right (predominantly) cerebellum, and the left IPL.

Main effects of experimental factors The rmANOVA revealed main effects of CONGRUENCE in activations of the left medial frontal gyrus including pre-SMA, bilateral inferior frontal (IFG), and right hemisphere IPL ($p < .05$, $k = 10$, FWE corr.; Table 3a). Post-hoc SPM T-contrasts ($p < .05$, $k = 10$, FWE corr.) showed a mismatching effect (mismatching – matching) in the left pre-SMA and bilateral IFG (Fig. 4a, Table 4a), whereas the left gyrus rectus showed a stronger activation in the matching than the mismatching condition (Fig. 4a, Table 4b). The main effect of the factor SUBSET (uniqueness/existence) was represented by activations within bilateral superior temporal gyrus (STG) extending to the temporal pole, the left insula extending to IFG, the right rolandic operculum and the IFG pars opercularis, the left SMA proper and postcentral regions, and right precentral gyrus ($p < .001$, $k = 10$, FWE corr.; Table 3b). Post-hoc SPM T-contrasts ($p < .001$, $k = 10$, FWE corr.) indicated that frontal, premotor, and motor activation clusters mainly resulted from uniqueness conditions compared to the existence condition, whereas the inverse comparison (existence vs. uniqueness) yielded bilateral STG activation (temporal pole; Fig. 4b, Table 4c, d). The main effect of the factor DEFINITENESS (indefinite/definite) comprised activation within the bilateral IFG, IPL, left pre-SMA, and middle frontal gyrus ($p < .0001$, $k = 33$, FWE corr. at cluster-level; Table 3c). Post-hoc SPM T-contrasts ($p < .0001$, $k = 35$, FWE corr. at cluster-level) indicated that all activation clusters resulted from the indefinite conditions (Fig. 4c, Table 4e).

3.2.2. ROI analyses

The conjunction analysis ($p < .005$, $k = 8$ uncorrected) resulted in activations within bilateral pre-SMA, bilateral IPL overlapping the angular gyrus (AG), bilateral basal ganglia (BG), including the right putamen, left pallidum, and bilateral anterior insula activation extending to the IFG pars triangularis (left) and pars opercularis (right; Fig. 5a, Table 5a). To localize ROI center-coordinates within relevant regions, anatomical masks of SMA, IFG, and IPL of each hemisphere were applied to the conjunction analysis clusters (Table 5b). Investigations of the hemodynamic responses in the ROIs averaged over hemispheres concerning the experimental factors revealed a stronger mismatching effect in the uniqueness subset compared to the existence subset within all three ROIs (SUBSET, IFG: $F(1, 19) = 7.27$, $p = .014$, $\eta_p^2 = 0.28$; IPL (AG): $F(1, 19) = 6.15$, $p = .023$, $\eta_p^2 = 0.25$; pre-SMA: $F(1, 19) = 8.15$, $p = .010$, $\eta_p^2 = 0.30$). Within pre-SMA, this main effect, was qualified by the interaction with DEFINITENESS (SUBSET × DEFINITENESS, $F(1, 19) = 6.34$, $p = .021$, $\eta_p^2 = 0.25$).

Table 1
Repeated measures ANOVA testing the effects of CONGRUENCE (match [M]/mismatch [MM]), SUBSET (uniqueness [U]/existence [E]), and DEFINITENESS (definite [D]/indefinite [I]) on acceptability ratings. Significant post hoc t-tests (Bonferroni Holm corrected) are also listed. Note: The three-way interaction was resolved by means of two sample t-tests in order to compare mismatch conditions with each other.

Main effects/interactions	F(1, 19)	p-value	η_p^2	Post hoc	p-value
CONGRUENCE	367.47	0.000	.95		
SUBSET	4.30	0.052	.19		
DEFINITENESS	30.09	0.000	.61		
CONGRUENCE × SUBSET	0.24	0.633	.01		
CONGRUENCE × DEFINITENESS	39.65	0.000	.68	M: D > I MM: I > D I: M > MM D: M > MM	0.013 0.000 0.000 0.000
SUBSET × DEFINITENESS	13.04	0.002	.41	ED < UD ED < EI	0.001 0.000
CONGRUENCE × SUBSET × DEFINITENESS	8.78	0.008	.32	MM: ED < EI MM: ED < UI MM: ED < UD	0.000 0.001 0.011

Note: η_p^2 = partial eta square of main effects and interactions (ANOVA).

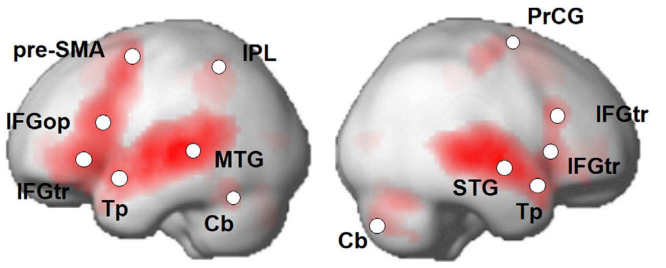

Fig. 3. Descriptive whole-brain analysis showing the average activation across all conditions (mismatch and match) marked in red ($p < .001$, $k = 10$, FWE corr.). Listening to the sentences elicited activation of the classical language areas (dots) such as bilateral inferior frontal gyrus pars opercularis (IFGop) and pars triangularis (IFGtr), superior temporal gyrus (STG), and middle temporal gyrus (MTG) extending to the temporal poles (Tp). In addition, pre-supplementary motor area (pre-SMA), precentral gyrus (PrCG), inferior parietal lobe (IPL), and cerebellum (Cb) were activated.

Table 2

Hemodynamic responses for the average effect of conditions versus null events ($p < .001$, $k = 10$, family wise error (FWE) corrected at cluster- and peak-level. Regions were labeled using the AAL2 atlas. Montreal Neurological Institute (MNI) coordinates (x, y, z) in the left (L) and right (R) hemisphere.

Region Label	Extent	*T*-value	MNI Coordinates x	y	z
Temporal_Sup_R	1735	403.109	63	−9	−3
Temporal_Pole_Sup_R	1735	231.323	54	12	−18
Frontal_Inf_Tri_R	1735	55.118	60	21	6
Temporal_Mid_L	3584	358.081	−60	−30	6
Temporal_Sup_L	3584	352.653	−57	−12	−3
Temporal_Pole_Sup_L	3584	217.417	−54	6	−12
Supp_Motor_Area_L	512	136.015	−6	6	54
Cerebelum_Crus2_R	499	114.517	15	−78	−39
Cerebelum_6_R	499	73.626	30	−63	−27
Cerebelum_8_R	499	65.034	30	−60	−48
Cingulate_Ant_L	542	95.413	−9	42	−3
Cerebelum_Crus1_L	38	61.780	−48	−60	−27
Parietal_Inf_L	258	59.856	−33	−51	39
Cingulate_Mid_R	106	57.102	6	−36	45
Precentral_R	157	55.164	36	−12	66
Cerebelum_6_L	68	51.552	−24	−54	−24
Cerebelum_Crus2_L	18	44.721	−15	−78	−42
Calcarine_L	61	40.439	0	−84	9
Cerebelum_Crus1_L	17	37.763	−9	−78	−27

Abbreviations: Ant = anterior, Crus (lat.: limb) = part of cerebellum, Inf = inferior, Mid = middle, Sup = superior, Supp = supplementary, Tri = pars triangularis.

Table 3

Hemodynamic responses for the main effects (rmANOVA) of the factors congruence (match/mismatch, subset (uniqueness/existence), and definiteness (definite/indefinite). Regions were labeled using the AAL2 atlas. Montreal Neurological Institute (MNI) coordinates (x, y, z) in the left (L) and right (R) hemisphere.

Region Label	Extent	*F*-value	MNI Coordinates x	y	z
a) CONGRUENCE ($p < .05$, $k = 10$, FWE corr.)					
Frontal_Sup_Medial_L	1128	71.645	−3	30	39
Frontal_Mid_2_R	1128	53.919	39	15	42
Insula_L	196	62.329	−33	18	−6
Frontal_Inf_Tri_R	226	54.294	45	21	3
Parietal_Inf_R	135	46.280	51	−51	42
Frontal_Inf_Orb_2_R	13	28.998	45	45	−12
Frontal_Mid_2_L	67	28.570	−51	15	42
Cingulate_Mid_R	10	27.533	6	−18	30
Frontal_Med_Orb_L	11	24.748	−6	42	−12
b) SUBSET ($p < .00001$, $k = 10$, FWE corr.)					
Temporal_Sup_R	579	224.375	60	−9	−6
Temporal_Sup_L	509	202.362	−54	−12	0
Heschl_L	509	167.035	−42	−27	9
Rolandic_Oper_R	369	111.897	42	6	12
Pallidum_R	369	66.050	24	0	0
Insula_L	356	105.644	−45	6	3
Precentral_R	704	83.105	39	−15	51
SupraMarginal_R	704	75.753	60	−36	27
Supp_Motor_Area_L	456	81.668	−6	−6	54
Cingulate_Mid_R	456	71.613	9	3	39
Cerebelum_4_5_L	59	73.204	−15	−51	−21
Occipital_Mid_L	88	67.227	−39	−69	6
Parietal_Inf_L	197	65.293	−48	−33	42
Postcentral_L	197	56.764	−57	−21	27
Cingulate_Mid_R	33	57.258	12	−24	45
Cingulate_Mid_L	13	52.919	−15	−27	42
c) DEFINITENESS ($p < .0001$, $k = 33$, FWE cluster corr.)					
Frontal_Inf_Orb_2_L	356	33.247	−39	21	−6
Frontal_Inf_Oper_L	356	27.360	−48	18	12
Frontal_Inf_Orb_2_L	356	19.157	−45	42	−9
Frontal_Sup_Medial_L	281	28.532	−3	21	42
Supp_Motor_Area_L	281	26.660	−6	9	63
Frontal_Inf_Tri_R	236	27.844	51	24	−3
Precuneus_R	44	23.936	9	−66	42
Temporal_Sup_R	37	20.447	54	−27	0
Frontal_Mid_2_R	57	20.291	39	12	45
Parietal_Inf_R	33	18.735	48	−54	48

Abbreviations: Inf = inferior, Sup = superior, Mid = middle, Orb = pars orbitalis, Oper = pars opercularis, Supp = supplementary, Med = medial, Tri = pars triangularis.

We hypothesized that the three ROIs are part of a network relevant to discourse processing. To investigate the activation pattern resulting from each of the matching (M) and mismatching (MM) conditions for each PSP condition (UD, UI, ED, EI) in the three ROIs of each hemisphere, we performed *T*-tests testing the effect against zero (corrected, $p < .006$ for the first testing). We expected that all three regions would be significantly activated during coherent as well as violated discourse structures, at least within the language-dominant left hemisphere (H1). Within all three left-hemispheric ROIs, we found significant activation for all (M and MM) conditions (all $p_{\mathrm{Bonf}} < .027$). Similarly, within the right-hemispheric IFG, all conditions (M and MM) showed significant activation (all $p_{\mathrm{Bonf}} < .002$). In contrast, in the right-hemispheric IPL/AG, only the mismatching uniqueness definite and indefinite conditions showed significant activation (UI: $p_{\mathrm{Bonf}} = .015$, UD: $p_{\mathrm{Bonf}} = .003$). Within right-hemispheric pre-SMA, four of the eight conditions showed significant activation (uniqueness definite mismatching, UD: $p_{\mathrm{Bonf}} < .0001$; uniqueness indefinite matching, UI: $p_{\mathrm{Bonf}} = .013$; uniqueness indefinite mismatching, UI: $p_{\mathrm{Bonf}} < .001$; existence indefinite mismatching, EI: $p_{\mathrm{Bonf}} < .001$).

To investigate the mismatching effect for each condition (UD, UI, ED, EI) in the three ROIs (IFG, IPL/AG, pre-SMA) averaged over hemispheres, we performed (corrected) *T*-tests testing the mismatching effect against zero (Fig. 5b). Generally, we expected significant mismatching effects within IFG across all conditions (H2) since in all violation conditions the initiation of the reference process may not run smoothly. Indeed, within IFG there was a mismatching effect for each condition (uniqueness definite, UD: $T = 3.61$, $p_{\mathrm{Bonf}} = .005$; uniqueness indefinite, UI: $T = 3.85$, $p_{\mathrm{Bonf}} = .004$; existence definite, ED: $T = 2.16$, $p_{\mathrm{Bonf}} = .043$; existence indefinite, EI: $T = 2.59$, $p_{\mathrm{Bonf}} = .036$; Fig. 5b left). We further expected significant mismatching effects within IPL/AG across all conditions (H3), since the referent triggered by the PSP seems not to be plausible across all these conditions and, thus, higher effort to contextual integration should be required. Within IPL (AG) there was a mismatching effect for all conditions except for the indefinite determiner of the existence subset (uniqueness definite, UD: $T = 3.44$, $p_{\mathrm{Bonf}} = .008$; uniqueness indefinite, UI: $T = 4.05$, $p_{\mathrm{Bonf}} = .003$; existence definite, ED: $T = 2.78$, $p_{\mathrm{Bonf}} = .024$; existence indefinite, EI: $T = 0.87$, $p_{\mathrm{Bonf}} = .39$; Fig. 5b middle). We further expected mismatching effects within the pre-SMA, but only under conditions allowing for re-analysis of discourse violation, since the pre-SMA presumably controls the executive system (comprehension) adjusting for the long-lasting processes of re-interpretation (re-reference, evaluation of plausibility) (H4). Within the pre-SMA, a mismatching effect was present for all conditions except the definite determiner of the existence subset (uniqueness definite, UD: $T = 4.05$, $p_{\mathrm{Bonf}} = .003$; uniqueness indefinite, UI: $T = 3.55$, $p_{\mathrm{Bonf}} = .006$; existence definite, ED: $T = -0.06$, $p_{\mathrm{Bonf}} = .96$; existence indefinite, EI: $T = 3.39$, $p_{\mathrm{Bonf}} = .006$; Fig. 5b right).

3.2.3. Correlation and covariance analyses

The Pearson's correlation analyses between acceptability ratings and the hemodynamic mismatching effect revealed a positive correlation for the existence indefinite condition (EI) within the right pre-SMA (Pearson's $r = 0.60$ $p_{\mathrm{Bonf}} = .043$; Fig. 6a). The remaining conditions (UD, UI, ED) did not show any correlation within one of the ROIs (all $p > .10$). Thus, we restricted the whole-brain covariance analysis to the existence indefinite condition (EI). The analysis ($p < .005$, $k = 511$, corrected) showed strong covariance within bilateral BG, specifically, the caudate

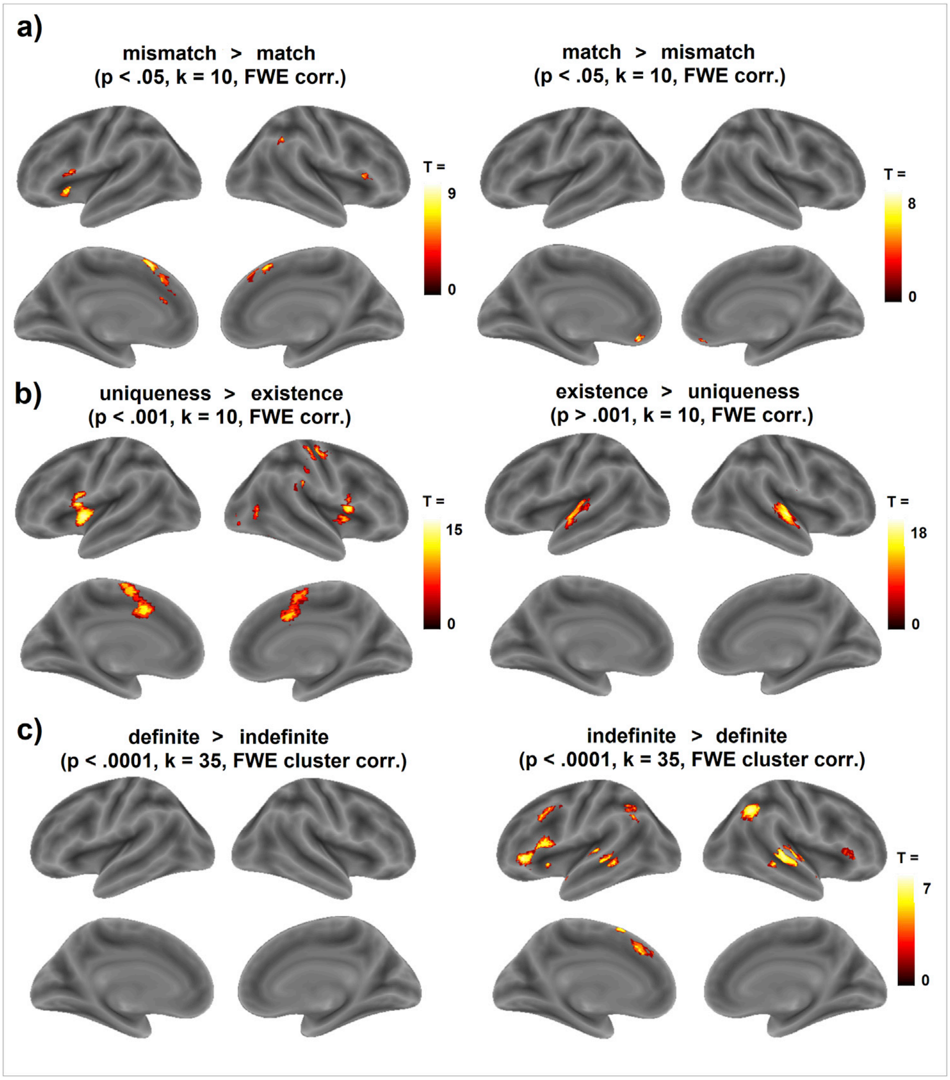

Fig. 4. Post-hoc SPM *T*-contrasts between the two levels of each factor from the whole brain ANOVA. Post hoc of the factors (a) CONGRUENCE (mismatch/match), (b) SUBSET (uniqueness/existence), and (c) DEFINITENESS (definite/indefinite). Thresholds were family wise error (FWE) corrected at cluster- and/or peak-level. T = size of differences relative to the variation in the sample (*T*-statistics). T values indicate strengths of activation (color-coded). K = minimum number of voxels included in a cluster (extent threshold).

nucleus (CN; Fig. 6b, Table 6) and the right pre-SMA (Fig. 6b, Table 6).

4. Discussion

The purpose of this study was to investigate which brain structures are involved in the processing of discourse structure and its violation. To this end, we measured the brain's hemodynamic responses (fMRI) related to processing contextual information presupposed in a spoken discourse. Discourse coherence was manipulated by using the indefinite determiner and the definite determiner as PSP triggers in test sentences presupposing information that either matches or does not match the content of a preceding context sentence. Specifically, we expected that the definite determiner allows for a smooth understanding of the spoken discourse when referring to a unique and existing referent in the context (matching UD/ED). Furthermore, we expected that the indefinite determiner allows for a smooth understanding in case of multiple (non-unique) context items (matching UI) or novel items in the test sentence (matching EI). In contrast, we expected processing difficulties when the definite determiner is used in the context of multiple or non-existing items (non-matching UD/ED), or when an indefinite noun phrase occurs in a context where such an item has already been introduced (non-matching UI/EI).

The acceptability ratings clearly show that participants were processing the test sentences in a context-sensitive way, that is, they paid attention to discourse coherence. Violations of the discourse structure were reflected in lower acceptability ratings compared to a coherent discourse structure across all four conditions (UD, UI, ED, EI). However, not all mismatching conditions were judged to be incoherent to a similar degree. The analysis revealed differences in acceptability of non-matching stimuli, with the lowest acceptability ratings for the existence violation (ED), whereas the (non-) uniqueness (UI, UD) and novelty (EI) violations yielded higher acceptability ratings. As the grouping in high- and low-raters showed, participants had a divided opinion concerning the PSP violations. Whereas low-raters judged all mismatching conditions as being unacceptable, high-raters evaluated the mismatching indefinite and definite uniqueness (UI, UD) and the indefinite existence (EI) conditions as less unacceptable or almost somewhat good. We interpret this pattern of results as an indicator that some participants were able to "repair" PSP violations more easily than others. To gain more insight into a potential repair process, we compared the behavioral responses of participants with their hemodynamic responses.

4.1. Brain regions involved in the semantic interpretation of heard sentences

Brain regions related to the processing of the sentence pairs included primary auditory areas of both hemispheres as well as adjacent structures of the STG, MTG, and Tp. Additionally, activation within the left- and right-hemispheric IFG, the left pre-SMA, the cerebellum, and the IPL (AG) were observed. The involvement of these areas nicely fits our expectations as outlined in the Introduction, in which we hypothesized that these

Table 4

Hemodynamic responses for the post hoc SPM *T*-contrasts between the two levels of each factor (CONGRUENCE, SUBSET, DEFINITENESS). Regions were labeled using the AAL2 atlas. Montreal Neurological Institute (MNI) coordinates (x, y, z) in the left (L) and right (R) hemisphere.

			MNI Coordinates		
Region Label	Extent	*T*-value	x	y	z
a) Mismatch vs. match ($p < .05$, $k = 10$, FWE corr.)					
Supp_Motor_Area_L	260	9.451	−3	21	54
Frontal_Sup_Medial_L	260	8.772	3	36	42
Frontal_Inf_Tri_R	30	7.829	48	21	6
Insula_L	34	7.821	−27	21	−9
Cingulate_Ant_L	11	7.018	−9	33	24
Frontal_Inf_Oper_L	22	6.893	−51	12	3
Parietal_Inf_R	10	6.768	51	−48	42
b) Match vs. mismatch ($p < .05$, $k = 10$, FWE corr.)					
Rectus_L	46	8.359	−6	36	−18
c) Uniqueness vs. existence ($p < .001$, $k = 10$, FWE corr.)					
Insula_L	334	15.850	−45	6	3
Putamen_L	334	12.121	−24	−3	0
Frontal_Inf_Oper_R	155	15.368	45	9	6
Pallidum_R	155	9.977	24	−9	−3
Cingulate_Mid_L	299	14.103	−9	6	39
Supp_Motor_Area_L	299	13.305	−6	−9	60
Cerebelum_4_5_L	53	13.279	−18	−51	−24
Precentral_R	134	11.947	42	−15	57
SupraMarginal_R	35	11.635	60	−33	30
Temporal_Mid_R	17	10.405	42	−66	6
Cerebelum_4_5_L	18	10.138	−3	−63	−15
Occipital_Inf_R	14	9.884	36	−84	−6
d) Existence vs. uniqueness ($p < .001$, $k = 10$, FWE corr.)					
Temporal_Sup_R	245	18.515	60	−12	0
Temporal_Mid_R	245	12.457	60	0	−18
Temporal_Sup_L	189	14.533	−57	−9	0
Temporal_Sup_L	189	12.290	−42	−30	9
e) Indefinite vs. definite ($p < .0001$, $k = 35$, FWE cluster corr.)					
Frontal_Inf_Tri_L	335	7123	−48	36	−3
Frontal_Inf_Tri_L	335	6263	−54	18	15
Frontal_Mid_2_L	335	5841	−45	12	45
Parietal_Inf_R	142	6930	45	−57	48
Frontal_Sup_Medial_L	88	6731	−6	24	39
Temporal_Sup_R	195	6661	60	−30	3
Parietal_Inf_L	57	5940	−39	−54	42
Supp_Motor_Area_L	35	5886	−9	12	63
Temporal_Mid_L	92	5864	−54	−39	−3
Frontal_Inf_Tri_R	109	5763	51	24	0

Abbreviations: Inf = inferior, Sup = superior, Mid = middle, Tri = pars triangularis, Supp = supplementary, Oper = pars opercularis.

language-related brain areas are involved in the processing of discourse structure triggered by sentences containing a PSP. Accordingly, brain regions within the temporal lobe, specifically the primary auditory cortices, STG, MTG, and Tp, serve to manage phonetic perception, categorization of phonemes, lexico-phonological operations, and access to lexico-semantic representations, respectively. Further, the IFG and IPL (AG) as well as connections between those areas make sure that sentence materials are processed according to semantic and pragmatic rules recalled in working memory (Frey et al., 2008; Fuster, 2001; Gennari et al., 2007; Kelly et al., 2010; Mollo et al., 2018; Tromp et al., 2015; Tulving, 1995). Most interestingly, processing discourse structures evoked a language network extended towards the pre-SMA and related regions such as the cerebellum and the BG. This broad extension

Table 5

Hemodynamic responses for the conjunction analysis (a) of the three main effects of the factors CONGRUENCE (match/mismatch), SUBSET (uniqueness/existence), and DEFINITENESS (definite/indefinite), and ROI center-coordinates (b) within anatomical masks (seeking space) overlapping activated clusters from conjunction analysis. Regions were labeled using the AAL atlas. Montreal Neurological Institute (MNI) coordinates (x, y, z) in the left (L) and right (R) hemisphere.

Region Label	Extent	F-value	MNI Coordinates		
			x	y	z
a) Conjunction of main effects ($p < .005$, $k = 8$, uncorr.)					
Insula_L	108	25.032	−45	15	−3
Insula_R	110	19.006	42	18	−3
Parietal_Inf_R	119	18.334	45	−45	45
Frontal_Sup_Medial_L	122	17.326	−3	15	42
Supp_Motor_Area_R	122	11.379	12	12	57
Frontal_Mid_2_R	12	13.880	45	15	51
Angular_R	50	13.704	48	−57	33
Precuneus_R	13	12.594	15	−69	39
Parietal_Inf_L	46	12.528	−45	−45	42
Precentral_L	64	12.265	−42	9	48
Putamen_R	8	10.089	15	9	0
Pallidum_L	14	10.019	−15	6	3
Frontal_Inf_Tri_R	15	9.162	51	27	12

b) ROI center-coordinates

Anatomical mask (seeking space)	Region label	ROI center-coordinates		
Supp_Motor_Area _L	Supp_Motor_Area	−3	12	45
Supp_Motor_Area _R	Supp_Motor_Area	12	12	57
Frontal_Inf_Tri + Oper _L	Frontal_Inf_Tri	−48	15	−3
Frontal_Inf_Tri + Oper _R	Frontal_Inf_Oper	45	18	0
Parietal_Inf + Angular _L	Parietal_Inf	−45	−45	42
Parietal_Inf + Angular _R	Parietal_Inf	45	−45	45

Abbreviations: Inf = inferior, Oper = pars opercularis, Tri = pars triangularis, Sup = superior, Mid = middle, Supp = supplementary.

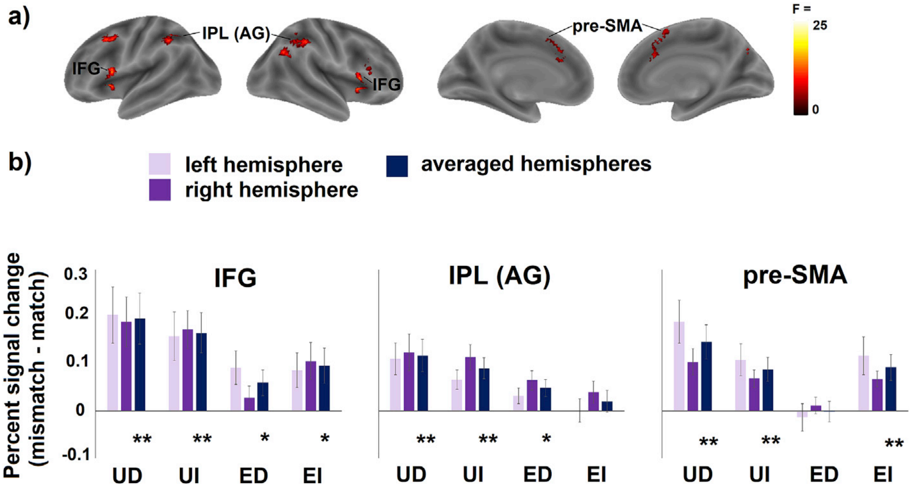

Fig. 5. ROI analyses of fMRI data. (a) Conjunction analyses revealed bilateral pre-SMA, IFG, and IPL (AG) as relevant ROIs. (b) Differences of signal change (mismatch minus match) were tested for each condition (UD, UI, ED, EI) for the values averaged across hemispheres (dark blue). Asterisks indicate significance level ($p_{\mathrm{Bonf}} < 0.01$ **, $p_{\mathrm{Bonf}} < 0.05$ *), error bars indicate standard error of the mean. F = ratio of variances (F-statistic) indicate strengths of activation (color-coded). UD = definite determiner in the uniqueness subset, UI = indefinite determiner in the uniqueness subset, ED = definite determiner in the existence subset, EI = indefinite determiner in the existence subset.

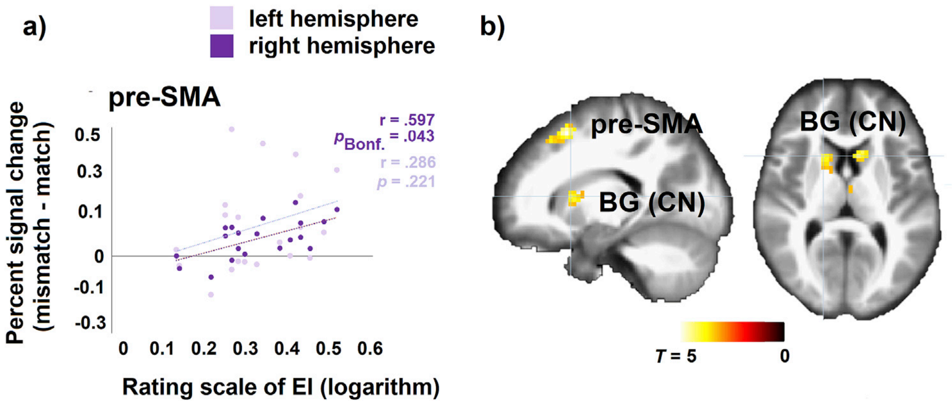

Fig. 6. The correlation of acceptability of discourse violation with hemodynamic responses. (a) Pearson's correlation between the hemodynamic mismatch effect and acceptability ratings shown for the existence indefinite (EI) condition within left and right pre-SMA. (b) Activation pattern when covarying acceptability ratings with the hemodynamic mismatch effect during EI within right pre-supplementary motor area (pre-SMA) and bilateral basal ganglia (BG) overlapping caudate nuclei (CN). T = size of differences relative to the variation in the sample (T-statistics) quantifying the magnitude of covariation (color-coded).

Table 6
Hemodynamic responses for the covariance analysis. Acceptability ratings (logarithm) covaried with hemodynamic responses of the existence indefinite (EI) condition contrasting mismatch versus match ($p < .005$. $k = 511$, corrected). Regions were labeled using the AAL2 atlas. Montreal Neurological Institute (MNI) coordinates (x, y, z) in the left (L) and right (R) hemisphere.

			MNI Coordinates		
Region Label	Extent	*T*-value	x	y	z
Frontal_Sup_2_R	511	5.799	18	18	57
Supp_Motor_Area_R		3.91	15	12	60
Caudate_L	644	4.990	−9	15	9
Caudate_R		4.703	18	15	9

Abbreviations: Sup = superior, Supp = supplementary.

contrasts with speech-related activation during passive listening to single words as this was found to be restricted to the temporal lobe (Dietrich et al., 2008). Our findings, however, are well in line with the results of Li et al. (2019), who mentioned the pre-SMA as belonging to the "core-language brain network" which is activated while performing distinct language tasks, for example, silent verb generation. We assume that the activation within the pre-SMA and the IFG serves to accomplish executive functions required by the task of this study. Specifically, the evaluation processes associated with the acceptability rating, as well as the task of keeping the sentence content available during discourse processing by means of inner speech (Friederici, 2012; Hertrich et al., 2016; Matchin et al., 2017; Rothermich and Kotz, 2013) most probably required executive functions (see e.g., Alderson-Day and Fernyhough, 2015, for a review). In addition, the activation of the cerebellum supports the function of the pre-SMA in discourse processing. Alongside this, the pre-SMA is a region receiving output from cerebellar-thalamic and basal ganglia-thalamic circuits (Schwartze et al., 2012a, 2012b), and Kotz and Schwartze (2010) assumed that the cerebellum strongly interacts with the pre-SMA in predictive sentence processing. Hence, the involvement of IFG, IPL (AG), and pre-SMA in discourse processing, requiring working memory, semantic/pragmatic integration, and executive functions fits nicely within the existing literature.

4.1.1. *Processing uniqueness vs. existence*

In addition to the network found to be relevant for sentence processing, distinct response patterns were observed with respect to each of the two stimulus subsets (uniqueness/existence). Contrasting the processing of uniqueness sentences to those of existence (UD, UI > ED, EI) yielded regions that have been assigned to premotor and motor cortex, right IFG pars opercularis, rolandic operculum, left SMA proper and postcentral regions, right precentral gyrus, and left insula extending to IFG. In contrast, the inverse contrast (ED, EI > UD, UI) revealed activations within bilateral STG and Tp. Regions with stronger activity during processing the uniqueness subset appear to belong to the dorsal stream of the language network (Hickok and Poeppel, 2007, 2004; Rauschecker and Scott, 2009), connecting sensory auditory (anterior/posterior temporal), parietal, and (pre-) motor-related regions to (pre-) frontal cortex through the superior longitudinal and arcuate fasciculus (Saur et al., 2010; see Bajada et al., 2015, for a review). Regions more active during processing the existence subset, however, appear to belong to the ventral stream linking the auditory cortex, the STG/MTG, and the Tp to inferior frontal regions via the extreme capsule and/or uncinate fasciculus (Friederici and Gierhan, 2013; Saur et al., 2010). The two streams overlap in the STG (Perrone-Bertolotti et al., 2017; Saur et al., 2010), a region that has been ascribed sub-lexical processing steps during speech perception (Binder et al., 2000; Scott and Johnsrude, 2003; see also Brauer et al., 2013; Dick et al., 2014; Dick and Tremblay, 2012; Parker et al., 2005, for evidence from tractography studies).

The dissociative pattern of brain activation resulting from processing the two types of sentence sets might be interpreted as mirroring the specific requirements for processing these two sets. Specifically, activation of sensory-motor speech regions might be due to specific verbal working memory processes maintaining speech sequences in order to analyze/reconstruct them. Here it is important to recall the experimental manipulations: Understanding (non-) uniqueness requires the identification of quantifiers in the context sentence (e.g., "several"/"one"). Thus, working memory (maintenance) mirrors morpho-syntactic and phonological operations. In contrast, comprehension of existence or novelty requires the recall of lexical information, which extends down to the word form. DeWitt and Rauschecker (2012, 2013) postulate for this function an auditory word form area in the left anterior superior temporal lobe. According to this interpretation, both morpho-syntactic and lexical memory functions are necessary to initiate deeper semantic operations, but these functions were recruited for use at a different level when processing the different PSP conditions. Because we assume that the uniqueness as well as the existence subsets were processed at a deep semantic level, additional brain areas are likely to be involved for sentence comprehension. The contrast method of analysis is likely to have resulted in many additional brain region activations canceling each other out, which is why these additional brain regions are not visible in the present analysis.

In general, the functional roles of the two linguistic processing streams as characterized so far with respect to their cognitive functions, that is, phonological/syntactic processing assigned to the dorsal part (Goucha and Friederici, 2015) and lexico-semantic processing assigned to the ventral part (Kümmerer et al., 2013; Saur et al., 2008), constitute a potential framework to understand the differential activation patterns induced by processing different types of PSPs.

4.1.2. *Processing definite vs. indefinite determiners*

The whole-brain data revealed a differential activation pattern regarding definite and indefinite determiners serving as PSP triggers. Stronger hemodynamic responses to the indefinite compared with the definite determiner (UI, EI > UD, ED) were found within bilateral STG, IFG, and IPL while the inverse comparison (UD, ED > UI, EI) did not yield significant differences in any brain region. This pattern of results suggests higher processing demands for the indefinite compared to the definite determiner. It might be interpreted in terms of theoretical and empirical literature (e.g., Mangold-Allwinn et al., 1995), theorizing that the usage of determiners and the different PSPs they trigger depend on distinct aspects during conversation, such as the communicative aim, prior knowledge about items, and discourse structure. In line with this assumption, Mangold-Allwinn et al. (1995) reported that the probability of using definite determiners increases when aiming to strengthen the salience of an item. Related to our results, processing the definite determiner as a salient indicator within a discourse might indicate somewhat automated processing, requiring fewer cognitive resources compared to those necessary for processing indefinite determiner phrases.

4.2. Processing of discourse incoherence

In line with our expectation that listening to violations of discourse structure would elicit higher cognitive effort than listening to a coherent discourse, the obtained fMRI data revealed stronger hemodynamic responses in response to incoherence relative to coherence within brain regions such as the IFG, pre-SMA and IPL (AG). In the following sections, we will discuss the results of the ROI analyses against the background of discourse violations (mismatching effect). These results point to specific functional roles of the brain areas for discourse understanding and clarify the functions of brain areas related to cognitive processes such as reference processes, semantic/pragmatic integration, and handling of mismatch. One might well ask, however, whether the mismatching effect indicating higher cognitive load when processing PSP violations is specific to language processing. The effect might instead be a response to processing of complexity in general. Considering the behavioral results of the current study as well as data from a recent reading time study (Rolke et al., 2019) that used the same stimulus subsets, it is clear that the response pattern was modulated by the kind of discourse violation. These differential modulations due to different discourse violations suggest that subjects were processing the test sentences in a speech-related, context-sensitive way and therefore that the mismatching effect – most probably – results from speech-related comprehension processes. Further, the activation pattern found in this study differs from that found in studies investigating cognitive load, for example, by using non-speech stimuli in a spatial working memory task (tracking the location of items) (Huang et al., 2016). Even though this study also observed activation within the inferior frontal cortex, inferior parietal regions, and the pre-SMA, the activation peaks of these data differed consistently from those of the current study: Inferior frontal and parietal activations are more dorsal relative to those of the current data, and pre-SMA activation is more anterior relative to this of the current data. Thus, although a general effect of cognitive load cannot be completely excluded, the arguments mentioned above justify deeper analyses of the mismatching effect as reflecting higher cognitive demands in case of speech-specific violations of discourse coherence.

4.2.1. IFG – reference process

Context-mismatching PSPs elicited a strong activation of IFG in all four conditions (UD, UI, ED, EI). This common role of IFG might be related to the functions of the left IFG for encoding of incoming information, “online” working memory maintaining contextual information (Gerton et al., 2004), detecting errors, and encoding of new representations (Kessler and Oberauer, 2015; Morris and Jones, 1990). The increased activation in case of a mismatching PSP trigger might thus mirror the detection of an erroneous reference process, which necessitates a re-check of the context in order to interpret the PSP.

Further support for the assumption that the hemodynamic response in IFG mirrors item categorization and the initiation of reference processes can be found in the enhanced mismatching effect of the IFG for the (non-) uniqueness violations compared to the existence violations. This enhanced effect might be related to the type of query which has to be made for the different PSP violations: While processing (non-) uniqueness might require a complex syntactical operation (singular/plural), processing (non-) existence of an item might elicit a simple logical binary query (available or not). These processing differences might modulate the strength of the mismatching effect when the listeners integrate their syntactic knowledge about the determiners and their specific usage. Taken together, we assume that activation of the IFG mirrors trigger categorization, error detection and the beginning of reference process, which is accompanied by different processing demands depending on the type of PSP trigger at hand.

4.2.2. IPL (AG) – semantic/pragmatic integration

IPL (AG) showed mismatching effects in (non-) uniqueness violations (UD, UI) and existence violations (ED), while novelty violations (EI) failed to activate this brain region. This pattern of results suggests that the IPL (AG) is involved in cognitive mismatching processing relevant for all conditions but not for the indefinite determiner of the existence subset (EI). As outlined in the Introduction, IPL (AG) has been found to be involved in a number of language-related processes, such as memory retrieval, semantic processing, logical reasoning, or attention (see Seghier, 2013 for a review). These functions include those relevant for PSP processing, that is, evaluating references in discourse including meaning identification of noun phrases and recognition of dependence (similarity) between critical noun phrases. Moreover, the identification of the meaning of critical noun phrases in test and context sentences requires access to lexical/semantic representations stored within long-term memory and the integration of noun phrases into the context by means of pragmatic knowledge. Further, Mars et al. (2011) points at two subdivisions of AG: a posterior-ventral part and an anterior-dorsal part. The IPL (AG) region of the current study was located in the latter being found functionally connected with the anterior prefrontal cortex and basal ganglia (BG; Mars et al., 2011; Uddin et al., 2010). The activation of this region as one involved in language processes converges with the ideas of Uddin et al. (2010), who emphasized that the anterior-dorsal part of IPL (AG) is linked more closely with anterior language regions in the inferior frontal gyrus than the posterior-ventral part of IPL (AG). In line with Seghier (2013), we argue that the IPL (AG) is involved in the retrieval of content (i.e., semantic) and episodic memories (including pragmatic experiences). Thus, it may be involved in a feedback strategy to ascertain semantic and pragmatic matching.

As there was no mismatching effect in the existence indefinite condition (EI), it seems that this condition does not require the above described processes. Theoretically, the existence indefinite condition (EI) presupposes a non-existing or independent (novel) item, while the uniqueness definite and indefinite conditions (UD, UI) and the existence definite (ED), even under negation, presuppose a referent mentioned in the context. It thus seems that a semantic/pragmatic reference process evoking enhanced IPL (AG) activation in case of mismatch is required for all conditions but the existence indefinite (EI) condition, because the test sentence in this condition could be interpreted without establishing a context reference. Support for this assumption comes from event-related brain potential (ERP) studies. Functional-anatomical mapping studies have related the posterior N400 ERP-component, which has been associated with difficulties in semantic integration (Kutas and Federmeier, 2011; Kutas and Hillyard, 1980), to IPL (AG) activity (Brouwer and Hoeks, 2013). Furthermore, Burkhardt and Roehm (2007) have reported that processing definite determiner phrases of an unintroduced new protagonist leads to a larger N400 than processing an introduced protagonist, indicating higher processing costs for difficult or impossible reference processes. Thus, the mismatching effects for conditions requiring a reference process in this study are in line with findings reported from Burkhardt and Roehm (2007) in the way that independency or dependency of noun phrases affect the N400 component that mirrors

IPL (AG) involvement.

Furthermore, (non-) uniqueness violations showed stronger hemodynamic responses than existence violations within IPL (AG). This difference might be due to facilitated semantic/pragmatic operations based on a simple logical binary query in case of the existence violation (ED) and complex syntactic singular/plural operations (UD, UI) in case of uniqueness PSPs. This idea receives support from the observation that negation seems to contribute to fast and effortless processing. For instance, in line with our results, Tettamani et al. (2008) have reported reduced hemodynamic responses in brain circuits for negated information compared to un-negated material. The AG activation peaks found in the current study are quite superior and anterior relative to those typically implicated in linguistic complexity (Pallier et al., 2011), but the general topography of the reported activations is compatible with conceptual-semantic processing (Price et al., 2015; Binder et al., 2009). For example, Bonhage et al., 2014 investigated sentence comprehension and manipulated linguistic material with respect to the number of words per sentences, testing working memory load as well as response to grammaticality of sentences. They found inferior parietal activation peaks in response to grammatically correct related word strings (vs. ungrammatical word strings) close to that of the current study. Taken together with this prior evidence, IPL (AG) activation in the current study might reflect processing of coherence between items within sentences rather than pure lexico-semantic access.

To summarize, the present IPL (AG) activation pattern appears to reflect requirements of semantic retrieval and pragmatic integration in order to process dependency between critical noun phrases in the test sentence and context sentence. We hypothesize that an IPL-based process was not necessary in case of novelty violations (EI), because the indefinite determiner in the existence condition does not need a semantic reference and pragmatic integration process.

4.2.3. Right hemisphere – monitoring and error detection

The present study was not specifically designed to obtain hemispheric differences, and most of the observed fMRI effects were bilateral rather than lateralized to one hemisphere. However, suggesting (in line with recent models of the language network in the brain) that language is predominantly processed in the left hemisphere, some considerations about the present right-hemispheric activations might be of interest here. Two right hemisphere regions, the IFG and the IPL (AG), were found to be activated during discourse violations. Many studies suggest a dominant role of the right hemisphere in error perception and processing (Gauvin et al., 2016; Raposo and Marques, 2013) and, thus, it might play a role in monitoring to minimize the number of misunderstandings. Researchers investigating contextual sentence processing have argued that the right hemisphere IFG, relative to the left hemisphere IFG, is more sensitive to discourse modulations, such as textual coherence and discourse anomalies (Kuperberg et al., 2006; Menenti et al., 2008; Nieuwland, 2012). Furthermore, previous studies have shown activation of right hemispheric IFG during presentation of incongruent, unfamiliar, irregular, or violated speech and non-speech stimuli such as music (Cheung et al., 2018; Hassabis and Maguire, 2007), indicating that monitoring, that is, falsification/verification of incoming items with existing contextual information, might play a role. Finally, while the phonologically organized verbal working memory is located in the left hemisphere including BA 44 (Rong et al., 2018), the right hemisphere seems to house a supra-modal pragmatic episodic memory (Brand et al., 2009; Ptak and Schnider, 2004) that might be used as an interface for the integration of language and non-language information (Parola et al., 2016) or, as in the present study, to match a test sentence with a context sentence that has already been encoded into supra-modal memory structures. We conclude that the right hemisphere might support the function of monitoring of speech sequences, especially when stimulus materials comprise erroneous information.

4.2.4. Pre-SMA – inhibitory control

The pre-SMA was sensitive to mismatching contexts while participants listened to (non-) uniqueness violations (UD, UI) and novelty violations (EI). In contrast, existence violations (ED) were not reflected in mismatching effects in the pre-SMA. We assume that the pre-SMA is involved in PSP processing because of its role in providing executive functions. As outlined in the Introduction, the pre-SMA has been reported to be involved in various superordinate control functions and monitoring processes (Geranmayeh et al., 2017; Hertrich et al., 2016; Dietrich et al., 2015, 2018). In language processing, meta-analyses of fMRI studies (Adank, 2012a, 2012b) have shown activation of the pre-SMA in case of sentence processing under adverse listening conditions. This activation was hypothesized to be due to its function of disambiguating linguistic information when lexical/semantic access is difficult. Although little is known about the linguistic function of the pre-SMA in detail, regarding the perception of accelerated speech, Vagharchakian et al., 2012 found pre-SMA activation near the limit of intelligibility when predictive top-down strategies must be engaged. Interestingly, they report a sudden collapse of the activation in the pre-SMA when speech becomes unintelligible (Vagharchakian et al., 2012). They interpreted these findings as showing a buffer function of the pre-SMA in the sense that as long as incoming items could be perceived, the pre-SMA coordinates speech comprehension. Transferred to the present study, the temporal coordination, re-analyses, and speech execution (comprehension) is not relevant when the existence PSP is violated– the trigger presupposes a negated item – and the condition is irreparable. However, if discourse, despite errors or atypical features, is evaluated as potentially interpretable, the function of pre-SMA should be necessary. In a recent behavioral study, Rolke et al. (2019) obtained some evidence which points in the direction of this assumption. These authors presented the sentences of the existence subset and showed that reading times for the definite existence condition (ED) was shorter compared to the indefinite existence condition (EI). The results of the present study suggest that in case of an inconsistent discourse structure, an increased pre-SMA involvement mirrors increased disambiguation and interpretation effort. Obviously, such an error management was not present in case of existence violations using the definite determiner (ED).

We can understand this pattern when considering the acceptability ratings, which revealed a pattern analogous to pre-SMA activity: The lowest acceptability ratings were observed for the existence violation (ED), whereas the (non-) uniqueness (UI, UD) or novelty (EI) violations yielded higher acceptance. This higher acceptance of discourse situations (e.g., in case of UD, UI, EI) might indicate that comprehension is not seriously threatened by an error since the recipient is able to handle the violation, possibly by guessing, tolerating slight errors, or modifying the (assumed) context in order to obtain an acceptable discourse meaning (see e.g. Kirsten et al., 2014). By contrast, in case of clear “unacceptable” errors during existence violations (ED), discourse meaning can easily be rejected by interrupting the language comprehension system. Various studies document pre-SMA involvement in repair mechanisms occurring under high demand conditions (Adank and Devlin, 2010; Scott et al., 2004; see Lima et al., 2016, for a review). Thus, pre-SMA appears to contribute to a top-down monitoring and error management process leading to a repair of mismatching PSPs in order to facilitate comprehension.

In order to guarantee fluency of speech comprehension, inner speech processes have to be slowed down when errors need to be managed. Kim et al. (2010) have reported that the major input to SMA comprises subcortical signals from the BG and cerebellum via the thalamus (see also Schwartze et al., 2012a; Schwartze et al., 2012b). Functionally, these signals appear to serve mostly inhibitory mechanisms delaying the procedural stream of speech comprehension in order to adjust the ongoing process to various demands (Kotz and Schwartze, 2010). In line with these findings, inhibitory control of executive functions (Wiecki and Frank, 2013), for example, in a stop movement task, has been discussed with respect to the pre-SMA (Kwon and Kwon, 2013). Further, the fact that anatomical connectivity between the SMA and the IFG (aslant tract) exists (see Introduction) might support these interactions in terms of controlling for speech comprehension. Overall, we assume that the pre-SMA manages inhibitory signals and seems to transiently slow down speech execution in order to gain time for error management.

4.2.5. *Pre-SMA and BG – successful re-interpretation*

Ultimately, the question remains in which region the decision for (i) rejection of erroneous speech materials, (ii) toleration as “passable” error (e.g., a slip of tongue), or (iii) repair (e.g., re-interpretation) is made. The acceptability ratings indicate that (non-) uniqueness (UD, UI) and novelty (EI) violations were evaluated as “less unacceptable”, while existence violations (ED) were rejected as “unacceptable”. Theoretically, the novelty violation (EI) is the only condition that could be repaired by means of re-interpretation of the context: Since the indefinite determiner presupposes non-existence, the introduction of a new item in the test sentence in addition to the already existing item mentioned in the context might be a rather easy way to re-interpret discourse meaning. Such a strategy might lead to successful comprehension and, in turn, higher acceptance or weaker rejection. This pattern of results was observed at least in a sub-group of participants whom we termed as high-raters.

Hemodynamic responses of novelty violations (EI) covaried with behavioral acceptability ratings within pre-SMA and bilateral BG at the level of the caudate nucleus (CN). This relationship suggests that the CN is involved in repair processes coming along with PSP violations. The CN is linked to the pre-SMA, but not SMA proper (Kim et al., 2010), and Grahn et al. (2008) have suggested that the CN selects appropriate sub-goals based on evaluation (see Jahanshahi et al., 2015, for a review). Related to our study, that would mean that the goal in case of a PSP violation, that is, the successful comprehension, would be initiated by the CN, potentially by a re-interpretation process. However, (non-) uniqueness violations (UD, UI) did not show any correlation/covariance between hemodynamic responses and acceptability ratings. This result occurred in spite of high acceptability ratings in some participants (high-raters), which might reflect successful repair processes. This differential pattern of behavioral results and hemodynamic responses might be viewed as a hint that high-raters did not really re-interpret discourse meaning in order to gain comprehension, but instead just tolerated the error (e.g., as a slip of tongue) without an attempt to create a new scenario. This in turn might explain the absence of acceptability-related BG activity in this group. Geiger et al. (2018) have documented strong interactions between the BG and the prefrontal cortex during encoding of novel (versus previously trained) working memory items and discussed facilitation of the “input-gating” of novel and relevant items as a function of the BG. Thus, both violation of novelty caused by the use of the indefinite determiner (EI) and re-interpreting the item in the test sentence as a novel (further) item in order to make the context plausible might justify activation of the BG for updating of working memory. In sum, tolerating implausible meaning during (non-) uniqueness violations (UD, UI) or re-interpreting implausible into plausible meaning evoked by novelty violations (EI) seem to be different processes as shown in different BG involvement. However, from the linguistic field’s view, it is still debatable whether accommodation means only re-interpretation including successful comprehension or whether accommodation also comprises tolerance of slight errors. In this regard, Giles et al. (1991) have mentioned the possibility that accommodation processes and their pragmatic implications may show considerable variability. Considering the acceptability rating data of the present study, accommodation can be considered as an adaptive behavior with the goal to understand the target sentence, including both re-interpretation as well as tolerance. However, the brain activations obtained in the present study suggest that behavioral acceptance of mismatching items may be due to different mechanisms such as tolerating slight errors in case of (non-) uniqueness violations (UI, UD) or content-related re-interpretation in case of novelty violations (EI).

4.3. *A neuroanatomical model of discourse processing*

To account for the processing differences across experimental conditions (UD, UI, ED, EI), we propose a neuroanatomical PSP processing model, which is based on previous cognitive models for sentence comprehension. For instance, Bornkessel-Schlesewsky and Schlesewsky (2008) have proposed a three-step model comprising (1) word category identification, (2) combination of arguments within a sentence as well as control of plausibility of these combinations, and (3) generalized mapping responsible for detection of overall sentence plausibility. In order to extend the model to PSP processing, Kirsten et al. (2014) have assumed that the PSP trigger might be (1) identified as a referential signal in the test sentence, (2) referred in a syntactic, semantic, and pragmatic plausible way to an item in the context sentence, (3) evaluated, in case of detected errors, for accommodation or rejection, and (4) eventually accommodated within the discourse structure in order to provide an understanding.

The present study suggests that these presumed processes increase cognitive load in case of violations of the discourse structure and that this is associated with stronger hemodynamic responses within the IFG, IPL (AG), pre-SMA, and BG. Our results, moreover, allow us to disentangle the contributions of different brain areas to different discourse understanding processes by closely examining the example of PSP processing. We suggest the following processing chain assigned to functional regions (see Fig. 7): The IFG seems to be assigned to working memory functions, which might be involved in scanning the ongoing speech signal. This scanning process comprises the consideration of reference triggers, maintenance of information, and initiation of a reference process. The IPL (AG) seems to be responsible for semantic/pragmatic retrieval and integration. It might recognize dependencies between noun phrases in the context and test sentence. The right hemisphere seems to be involved in monitoring processes in order to detect errors and, speculatively, might contribute to the decision of whether violations should be tolerated, repaired, or rejected. Finally, the pre-SMA seems to be involved in control functions, and it might become active when speech comprehension is difficult. For example, if discourse violations occur and errors are evaluated as manageable, the pre-SMA might slow down the speech comprehension system (IFG) in order to facilitate a restart. In case of reparable errors, the BG might be involved in re-interpretation as kind of accommodation process for successful comprehension.

The current study provides a starting point for investigating processing of context-dependent speech by means of PSPs. In comparing

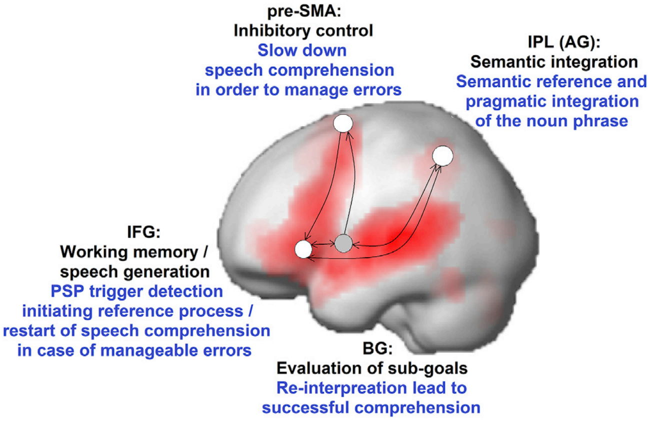

Fig. 7. Suggested functional neuroanatomical model for discourse management based on the processing network outlined in the present results. Black text: functions of IFG, IPL (AG), pre-SMA, and BG (dots) reported in literature; blue text: hypothesized functions for the processing of presuppositions (PSPs); arrows: hypothesized functional connections between IFG, IPL (AG), pre-SMA, and BG.

processing of discourse violations to processing of a coherent discourse, we aimed to isolate brain structures which were involved in processing references within discourse structures. To estimate whether the brain areas found to be active in the present study take a specific role in processing PSPs and PSP violations, or whether the described network generally mirrors discourse understanding, further studies are necessary. One interesting further approach would be to compare sentences with PSPs to a matched control condition, that is, to sentences containing a similar degree of phonological, lexical, syntactic, and semantic complexity, but lacking a PSP trigger. When considering the maxims of Grice (1975), one might propose different hypotheses for this contrast. On the one hand, in keeping a message informative, brief, and relevant by using PSPs (p. 3), the hemodynamic activation would presumably be higher in a control condition without PSPs than in PSP sentences because, in the former, more linguistic entities need to be processed, and redundant information would increase cognitive load. On the other hand, because such a control condition does not need reference processes, initial comprehension might perhaps be simpler in this situation, leading to possible reductions in brain activation while processing the sentences. So far, it seems that experimental studies support the latter alternative as PSP processing enhanced reading times in behavioural studies (e.g., Rolke et al., 2019; Tiemann et al., 2011). All in all, PSPs seem to be a promising tool for investigating discourse comprehension and discovering which brain structures underlie the human ability to understand a discourse and its pragmatic implications.

5. Conclusion

This fMRI study addressed the neuronal structures involved in spoken discourse comprehension. We manipulated discourse coherence by using PSP triggers (i.e., the indefinite determiner or definite determiner) that either correspond or fail to correspond to items in the discourse context. Bilateral IFG, IPL (AG), and pre-SMA were active during violations of discourse coherence. We discussed these areas in light of their respective, previously-established contexts: of working memory functions being relevant for initiation of the reference process (IFG), of semantic/pragmatic integration required for identification of plausible referents (IPL/AG), and of an inhibitory control network being important to slow down speech comprehension in order to manage errors and restart comprehension (pre-SMA). Furthermore, we assume that right hemispheric processes might contribute to the evaluation of how to integrate erroneous materials: rejection, tolerance as “passable” error (e.g., a slip of tongue), or repair by means of re-interpretation. In the last two cases, these are manageable errors requiring inhibitory control functions (pre-SMA). Additionally, re-interpretation considered as a kind of sub-goal leading to successful comprehension might be associated with a strong activation of BG – an input region of pre-SMA – and might reflect the initiation of a potential accommodation process.

Acknowledgement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors thank Fotini Scherer for excellent technical assistance as well as Jürgen Dax and Michael Erb for installing technical devices for stimulus presentation and the recording of behavioral data in the fMRI scanner.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.116047.

References

Adank, P., 2012a. Design choices in imaging speech comprehension: an activation likelihood estimation (ALE) meta-analysis. Neuroimage 63, 1601–1613.

Adank, P., 2012b. The neural bases of difficult speech comprehension and speech production: two activation likelihood estimation (ALE) meta-analyses. Brain Lang. 122, 42–54.

Adank, P., Devlin, J., 2010. On-line plasticity in spoken sentence comprehension: adapting to time-compressed speech. Neuroimage 49, 1124–1132. https://doi.org/10.1016/j.neuroimage.2009.07.032.

Alderson-Day, B., Fernyhough, C., 2015. Inner speech: development, cognitive functionsnd neurobiology. Psychol. Bull. 141, 931–965. https://doi.org/10.1037/bul0000021.

Alm, P.A., 2011. Cluttering: a neurological perspective. In: Ward, D., Scott, K.S. (Eds.), Cluttering: A Handbook of Research, Intervention, and Education. Psychology Press, Hove (UK), pp. 3–28.

Alonso-Ovalle, L., Menendez-Benito, P., Schwarz, F., 2011. Maximize presupposition and two types of definite competitors. In: Lima, S., Mullin, K., Sith, B. (Eds.), *Proceedings Of the 39th Meeting of the North East Linguistic Society*. GLSA, Amherst, MA, pp. 29–40.

Amunts, K., Lenzen, M., Friederici, A.D., Schleicher, A., Morosan, P., Palomero-Gallagher, N., Zilles, K., 2010. Broca’s region: novel organizational principles and multiple receptor mapping. PLoS Biol. 8 (9), e1000489 doi: 101371/journal.pbio.1000489.

Anderson, J.E., Holcomb, P.J., 2005. An electrophysiological investigation of the effects of coreference and synonymy. Brain Lang. 94, 200–216. https://doi.org/10.1016/j.bandl.200501001.

Baddeley, A., 2003. Working memory: looking back and looking forward. Nat. Rev. Neurosci. 4, 829–839.

Bajada, C.J., Lambon-Ralph, M.A., Cloutman, L.L., 2015. Transport for language south of the Sylvian fissure: the routes and history of the main tracts and stations in the ventral language network. Cortex 69, 141–151.

Beaver, D., Zeevat, H., 2007. Accommodation. In: Ramchand, G., Reiss, C. (Eds.), Oxford Handbook of Linguistic Interfaces. University Press, Oxford, pp. 503–538.

Binder, J., 2017. Current controversies on Wernicke’s area and its role in language. Curr. Neurol. Neurosci. Rep. 17, 58. https://doi.org/10.1007/s11910-017-0764-8.

Binder, J., Desai, R., Graves, W., Conant, L., 2009. Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. Cerebral Cortex 19, 2767–2796. https://doi.org/10.1093/cercor/bhp055.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellqowan, P.S., Springer, J.A., Kaufman, J.N., Possing, E.T., 2000. Human temporal lobe activation by speech and nonspeech sounds. Cerebr. Cortex 10 (5), 512–528.

Bonhage, C., Fiebach, C., Bahlmann, J., Müller, J., 2014. Brain Signature of Working Memory for Sentence Structure: Enriched Encoding and Facilitated Maintenance. Journal of Cognitive Neuroscience 26 (8), 1654–1671. https://doi.org/10.1162/jocn_a_00566.

Bornkessel-Schlesewsky, I., Schlesewsky, M., 2008. An alternative perspective on “semantic P600“ effects in language comprehension. Brain Res. Rev. 59, 55–73. https://doi.org/10.1016/j.brainresrev.2008.05.003.

Brand, M., Eggers, C., Reinhold, N., Fujiwara, E., Kessler, J., Heiss, W.D., Markowitsch, H.J., 2009. Functional brain imaging in 14 patients with dissociative amnesia reveals right inferolateral prefrontal hypometabolism. Psychiatry Res. Neuroimaging 174, 32–39. https://doi.org/10.1016/j.pscychresns.2009.03.008.

Brauer, J., Anwander, A., Perani, D., Friederici, A.D., 2013. Dorsal and ventral pathways in language development. Brain Lang. 127, 289–295. https://doi.org/10.1016/j.bandl.2013.03.001.

Brouwer, H., Hoeks, J.C.J., 2013. A time and place for language comprehension: mapping the N400 and the P600 to a minimal cortical network. Front. Hum. Neurosci. 7, 758. https://doi.org/10.3389/fnhum.2013.00758.

Burkhardt, P., Roehm, D., 2007. Differential effects of saliency: an event-related brain potential study. Neurosci. Lett. 413, 115–120. https://doi.org/10.1016/j.neulet.2006.11.038.

Camos, V., Barrouillet, P., 2014. Attentional and non-attentional systems in the maintenance of verbal information in working memory: the executive and phonological loops. Frontiers in Human Neurosciences 8, 900.

Chemla, E., 2008. An epistemic step for anti-presuppositions. J. Semant. 25, 141–173. https://doi.org/10.1093/jos/ffm017.

Cheung, V.K.M., Meyer, L., Friederici, A.D., Kölsch, S., 2018. The right inferior frontal gyrus processes nested non-local dependencies in music. Nat. Sci. Rev. 8, 3822. https://doi.org/10.1038/s41598-018-22144-9.

DeWitt, I., Rauschecker, J.P., 2012. Phoneme and word recognition in the auditory ventral stream. Proc. Natl. Acad. Sci. U.S.A. 109 (8), E505–E514. https://doi.org/10.1073/pnas.1113427109.

DeWitt, I., Rauschecker, J.P., 2013. Wernicke’s area revisited: parallel streams and word processing. Brain Lang. 127, 181–191. https://doi.org/10.1016/j.bandl.2013.09014.

Dick, A.S., Bernal, B., Tremblay, P., 2014. The language connectome: new pathways, new concepts. The Neuroscientist 20, 453–467.

Dick, A.S., Tremblay, P., 2012. Beyond the arcuate fasciculus: consensus and controversy in the connectional anatomy of language. Brain 135, 3529–3550. https://doi.org/10.1093/brain/aws222.

Dietrich, S., Hertrich, I., Alter, K., Ischebeck, A., Ackermann, H., 2008. Understanding the emotional expression of verbal interjections: a functional MRI study. Neuroreport 19 (18), 1751–1755.

Dietrich, S., Hertrich, I., Ackermann, H., 2015. Network modeling for functional magnetic resonance imaging (fMRI) signals during ultra-fast speech comprehension in late-blind listeners. PLoS One 10, e132196. https://doi.org/10.1371/journal.pone.0132196.

Dietrich, S., Hertrich, I., Müller-Dahlhaus, F., Ackermann, H., Belardinelli, P., Desideri, D., et al., 2018. Reduced performance during a sentence repetition task by continuous theta-burst magnetic stimulation of the pre-supplementary motor area. Front. Neurosci. 12, 1–13. https://doi.org/10.3389/fnins.2018.00361.

Domaneschi, F., Di Paola, S., 2018. The processing costs of presupposition accommodation. J. Psycholinguist. Res. 47, 483–503. https://doi.org/10.1007/s10936-017-9534-7.

Frey, S., Campbell, J.S.W., Pike, G.B., Petrides, M., 2008. Dissociating the human language pathways with high angular resolution diffusion fiber tractography. J. Neurosci. 28 (45), 11435–11444. https://doi.org/10.1523/JNEUROSCI.2388-08.2008.

Friederici, A.D., 2012. The cortical language circuit: from auditory perception to sentence comprehension. Trends Cogn. Sci. 16, 262–268. https://doi.org/10.1016/j.tics.2012.04.001.

Friederici, A.D., Gierhan, S.M.E., 2013. The language network. Curr. Opin. Neurobiol. 23, 250–254. https://doi.org/10.1016/j.conb.2012.10.002.

Friston, K.J., Rotshtein, P., Geng, J.J., Sterzer, P., Henson, R.N., 2006. A crtical of functional localisers. Neuroimage 30, 1077–1087. https://doi.org/10.1016/j.neuroimage.2005.08.012.

Fuster, J.M., 2001. The prefrontal cortex – an update: time is of the essence. Neuron 30, 319–333. https://doi.org/10.1016/S0896-6273(01)00285-9.

Garrod, S.C., Sanford, A.J., 1982. The mental representation of discourse in a focused memory system: implications for the interpretation of anaphoric noun phrases. J. Semant. 1, 21–41. https://doi.org/10.1093/jos/1.1.21.

Gauvin, H.S., De Baene, W., Brass, M., Hartsuiker, R.J., 2016. Conflict monitoring in speech processing: an fMRI study of error detection in speech production and perception. Neuroimage 126, 96–105. https://doi.org/10.1016/j.neuroimage.2015.11.037.

Geiger, L., Moessnang, C., Schäfer, A., Zang, Z., Zangl, M., Cao, H., et al., 2018. Novelty modulates human striatal activation and prefrontal-striatal effective connectivity during working memory encoding. Brain Struct. Funct. https://doi.org/10.1007/s00429-018-1679-0.

Gennari, S.P., MacDonald, M.C., Postle, B.R., Seidenberg, M.S., 2007. Context-dependent interpretation of words: evidence for interactive neural processes. Neuroimage 35, 1278–1286. https://doi.org/10.1016/j.neuroimage.2007.01.015.

Geranmayeh, F., Chau, T.W., Wise, R.J.S., Leech, R., Hampshire, A., 2017. Domain-general subregions of the medial prefrontal cortex contribute to recovery of language after stroke. Brain 140, 1947–1958.

Gerton, B.K., Brown, T.T., Meyer-Lindenberg, A., Kohn, P., Holt, J.L., Olsen, R.K., Berman, K.F., 2004. Shared and distinct neurophysiological components of the digits forward and backward tasks as revealed by functional neuroimaging. Neuropsychologia 42, 1781–1787. https://doi.org/10.1016/j.neuropsychologia.2004.04.023.

Giles, H., Coupland, N., Coupland, J., 1991. Accommodation theory: communication, context, and consequence. In: Giles, H., Coupland, J., Coupland, N. (Eds.), *Contexts Of Accommodation: Developments In Applied Sociolinguistics* (Studies in Emotion and Social Interaction. Cambridge University Press, Cambridge, pp. 1–68. https://doi.org/10.1017/CBO9780511663673.001.

Goucha, T., Friederici, A.D., 2015. The language skeleton after dissecting meaning: a functional segregation within Broca's Area. Neuroimage 114, 294–302.

Grahn, J.A., Parkinson, J.A., Owen, A.M., 2008. The cognitive functions of the caudate nucleus. Prog. Neurobiol. 86, 141–155. https://doi.org/10.1016/j.pneurobio.2008.09.004.

Grice, H.P., 1975. Logic and conversation. In: Cole, P., Morgan, J.L. (Eds.), Syntax and Semantics Vol. 3 Speech Acts. Academic Press, New York, pp. 41–58.

Hassabis, D., Maguire, E.A., 2007. Deconstruction episodic memory with construction. Trends Cogn. Sci. 11, 299–306. https://doi.org/10.1016/j.tics.2007.05.001.

Heim, I., 1982. The Semantics of Definite and Indefinite Noun Phrases. Doctoral dissertation. University of Massachusetts, Amherst. http://eecopock.info/DynamicSemantics/Readings/Heim1982-Dissertation.pdf.

Heim, I., Kratzer, A., 1998. Semantics and Generative Grammar. Blackwell, Malden.

Hertrich, I., Kirsten, M., Tiemann, S., Beck, S., Wühle, A., Ackermann, H., Rolke, B., 2015. Context-dependent impact of presuppositions on early magnetic brain responses during speech perception. Brain Lang. 149, 1–23.

Hertrich, I., Dietrich, S., Ackermann, H., 2016. The role of the supplementary motor area for speech and language processing. Neurosci. Biobehav. Rev. 68, 602–610. https://doi.org/10.1016/j.neubiorev.2016.06.030.

Hickok, G., 2012. Computational neuroanatomy of speech production. Nat. Rev. Neurosci. 13, 135–145. https://doi.org/10.1038/nrn3158.

Hickok, G., Poeppel, D., 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92, 67–99. https://doi.org/10.1016/j.cognition.2003.10.011.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402. https://doi.org/10.1038/nrn2113.

Huang, A.S., Klein, D.N., Leung, H.-C., 2016. Load-related brain activation predicts spatial working memory performance in youth aged 9-12 and is associated with executive function at earlier ages. Dev. Cogn. Neurosci. 17, 1–9. https://doi.org/10.1016/j.dcn.2015.10.007.

Jahanshahi, M., Obeso, I., Rothwell, J.C., Obeso, J.A., 2015. A fronto-striatosubthalamic-pallidal network for goal-directed and habitual inhibition. Nat. Rev. Neurosci. 16, 719–732. https://doi.org/10.1038/nrn4038.

Kelly, C., Uddin, L.Q., Shehzad, Z., Margulies, D.S., Castellanos, F.X., Milham, M.P., Petrides, M., 2010. Broca's region: linking human brain functional connectivity data and non-human primate tracing anatomy studies. Eur. J. Neurosci. 32, 383–398. https://doi.org/10.1111/j.1460-9568.2010.07279.x.

Kessler, Y., Oberauer, K., 2015. Forward scanning in verbal working memory updating. Psychon. Bull. Rev. 22, 1770–1776. https://doi.org/10.3758/s13423-015-0853-0.

Kim, J.-H., Lee, J.-M., Kim, S.H., Lee, J.H., Kim, S.T., Saad, Z.S., 2010. Definig functional SMA and pre-SMA subregions in human MPC using resting state fMRI: functional connectivity-based parcellation method. Neuroimage 49, 2375. https://doi.org/10.1016/j.neuroimage.2009.10.016.

Kirsten, M., Tiemann, S., Seibold, V.C., Hertrich, I., Beck, S., Rolke, B., 2014. When the polar bear encounters many polar bears: event-related potential context effects evoked by uniqueness failure. Lang. Cognit. Neurosci. 29, 1147–1162. https://doi.org/10.1080/23273798.2014.899378.

Kotz, S.A., Schwartze, M., 2010. Cortical speech processing unplugged: a timely subcortico-cortical framework. Trends Cogn. Sci. 14, 392–399. https://doi.org/10.1016/j.tics.2010.06.005.

Krahmer, E.J., 1998. Presupposition and determinedness. In: Krahmer, E. (Ed.), Presupposition and Anaphora. CSLI Publications, Stanford, pp. 193–224.

Kümmerer, D., Hartwigsen, G., Kellmeyer, P., Glauche, V., Mader, I., Klöppel, S., et al., 2013. Damage to ventral and dorsal language pathways in acute aphasia. Brain 136, 619–629. https://doi.org/10.1093/brainaws354.

Kuperberg, G.R., Lakshmanan, B.M., Caplan, D.N., Holcomb, P.J., 2006. Making sense of discourse: an fMRI study of causal inferencing across sentences. Neuroimage 33, 343–361. https://doi.org/10.1016/j.neuroimage.2006.06.001.

Kutas, M., Federmeier, K., 2011. Thirty years and counting: finding meaning in the N400 component of the event related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.

Kutas, M., Hillyard, S., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205. https://doi.org/10.1126/science.7350657.

Kwon, Y.H., Kwon, J.W., 2013. Response inhibition induced in the stop-signal task by transcranial direct current stimulation of the pre-supplementary motor area and primary sensoriomotor cortex. J. Phys. Ther. Sci. 25, 1083–1086.

Li, Q., Del Ferraro, G., Pasquini, L., Peck, K.K., Makse, H.A., Holodny, A.I., 2019. Core language brain network for fMRI-language task used in clinical applications. Quant. Biol. Neurons Cognit. 14 arXiv: 1906.07546.

Lima, C.F., Krishnan, S., Scott, S.K., 2016. Roles of supplementary motor areas in auditory processing and auditory imagery. Trends Neurosci. 39, 527–542. https://doi.org/10.1016/j.tins.2016.06.003.

Mangold-Allwinn, R., Baratelli, S., Kiefer, M., Kölbing, H.-G., 1995. Empirische befunde: determinanten der Artikelwahl. In: Mangold-Allwinn, R., Baratelli, S., Kiefer, M., Kölbing, H.-G. (Eds.), Wörter für Dinge. Von flexiblen Konzepten zu variablen Benennungen. Westdeutscher Verlag GmbH, Opladen, pp. 19–62.

Mars, R.B., Jbabdi, S., Sallet, J., O'Reilly, J.X., Croxson, P.L., Olivier, E., et al., 2011. Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. J. Neurosci. 31 (11), 4087–4100.

Martin, A., Chao, L.L., 2001. Semantic memory and the brain: structure and processes. Curr. Opin. Neurobiol. 11, 194–201. https://doi.org/10.1016/S0959-4388(00)00196-3.

Masia, V., Canal, P., Ricci, I., Vallauri, E.L., Bambini, V., 2017. Presupposition of new information as a pragmatic garden path: evidence from event-related brain potentials. J. Neurolinguistics 42, 31–48. https://doi.org/10.1016/j.jneuroling.2016.11.005.

Matchin, W., Hammerly, C., Lau, E., 2017. The role of the IFG and pSTS in syntactic prediction: evidence from a parametric study of hierarchical structure in fMRI. Cortex 88, 106–123. https://doi.org/10.1016/j.cortex.2016.12010.

Menenti, L., Petersson, K.M., Scheeringa, R., Hagoort, P., 2008. When elephants fly: differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. J. Cogn. Neurosci. 21 (12), 2358–2368. https://doi.org/10.1162/jocn.2008.21163.

Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. 24 https://doi.org/10.1146/annurev.neuro.24.1.167, 167-20.

Mollo, G., Jefferies, E., Cornelissen, P., Gennari, S.P., 2018. Context-dependent lexical ambiguity resolution: MEG evidence for the time-course of activity in inferior frontal gyrus and posterior middle temporal gyrus. Brain Lang. 177, 23–36. https://doi.org/10.1016/j.bandl.2018.01.001.

Morris, N., Jones, D.M., 1990. Memory updating in working memory: the role of the central executive. Br. J. Psychol. 81, 111–121. https://doi.org/10.1111/j.2044-8295.1990.tb02349.x.

Napoli, M., 2013. When the indefinite article implies uniqueness: a case study from Old Italian. Folia Linguist. 47, 183–235. https://doi.org/10.1515/flin.2013.008.

Nieuwland, M.S., 2012. Establishing propositional truth-value in counterfactual and real-world contexts during sentence comprehension: differential sensitivity of the left and right inferior frontal gyri. Neuroimage 59 (4), 3433–3440. https://doi.org/10.1016/j.neuroimage.2011.11.018.

Pallier, C., Devauchelle, A.-D., Dahaene, S., 2011. Cortical representation of the constituent structure of sentences. PNAS 108 (6), 2522–2527. https://doi.org/10.1073/pnas.1018711108.

Parker, G.J., Luzzi, S., Alexander, D.C., Wheeler-Kingshott, C.A., Ciccalrelli, O., Lambon-Ralph, M.A., 2005. Lateralization of ventral and dorsal auditory-language pathways in the human brain. Neuroimage 24, 656–666. https://doi.org/10.1016/j.neuroimage.2004.08.047.

Parola, A., Gabbatore, I., Bosco, F.M., Bara, B.G., Cossa, F.M., Gindri, P., Sacco, K., 2016. Assessment of pragmatic impairment in right hemisphere damage. J. Neurolinguistics 39, 10–25. https://doi.org/10.1016/j.jneuroling.2015.12.003.

Perrone-Bertolotti, M., Kauffmann, L., Pichat, C., Vidal, J.R., Baciu, M., 2017. Effective connectivity between ventral occipito-temporal and ventral inferior frontal cortex during lexico-semantic processing. A dynamic causal modeling study. Front. Hum. Neurosci. 11, 325. https://doi.org/10.3389/fnhum.2017.00325.

Poldrack, R.A., 2007. Region of interest analysis for fMRI. Soc. Cogn. Affect. Neurosci. 2, 67–70. https://doi.org/10.1093/scan/nsm006.

Poldrack, R.A., Mumford, J.A., 2009. Independence in ROI analysis: where is the voodoo? Soc. Cogn. Affect. Neurosci. 4, 208–213. https://doi.org/10.1093/scan/nsp011.

Price, A., Bonner, M., Peele, J., Grossman, M., 2015. Converging Evidence for the Neuroanatomic Basis of Combinatorial Semantics in the AngularGyrus. The Journal of Neuroscience 35 (7), 3276–3284. https://doi.org/10.1523/JNEUROSCI.3446-14.2015.

Ptak, R., Schnider, A., 2004. Disorganised memory after right dorsolateral prefrontal damage. Neurocase 10, 52–59. https://doi.org/10.1080/13554790490960495.

Raposo, A., Marques, J.F., 2013. The contribution of fronto-parietal regions to sentence comprehension: insights from the Moses illusion. Neuroimage 83, 431–437. https://doi.org/10.1016/j.neuroimage.2013.06.052.

Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat. Neurosci. 12, 718–724. https://doi.org/10.1038/nn.2331.

Robertson, D.A., Gernsbacher, M.A., Guidotti, S.J., Robertson, R.R.W., Irwin, W., Mock, B.J., Campana, M.E., 2000. Functional neuroanatomy of the cognitive process of mapping during discourse comprehension. Psychol. Sci. 11 (3), 255–260. https://doi.org/10.1111/1467-9280.00251.

Rodriquez-Fornells, A., Cunillera, T., Mestres-Misse, A., Diego-Balaguer, R., 2009. Neurophysiological Mechanisms Involved in Language Learning in Adults, vol. 364. Philosophical Transaction of the Royal Society B, pp. 3711–3735. https://doi.org/10.1098/rstb.2009.0130.

Rolke, B., Kirsten, M., Seibold, V.C., Dietrich, S., Hertrich, I., 2019. Processing References in Context: when the Polar Bear Does Not Meet a Polar Bear (Manuscript in preparation).

Rong, F., Isenberg, A.L., Sun, E., Hickok, G., 2018. The neuroanatomy of speech sequencing at the syllable level. PLoS One 13, e0196381. https://doi.org/10.1371/journal.pone.0196381.

Rothermich, K., Kotz, S.A., 2013. Predictions in speech comprehension: fMRI evidence on the meter-semantic interface. Neuroimage 70, 89–100. https://doi.org/10.1016/j.neuroimage.2012.12.013.

Saur, D., Kreher, B.W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.S., et al., 2008. Ventral and dorsal pathways for language. Proc. Natl. Acad. Sci. U. S. A 105, 18035–18040. https://doi.org/10.1073/pnas.0805234105.

Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Küpper, H., Kellmeyer, P., et al., 2010. Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. Neuroimage 49, 3187–3197. https://doi.org/10.1016/j.neuroimage.2009.11.009.

Schlenker, P., 2012. Maximize presupposition and gricean reasoning. Nat. Lang. Semant. 20, 391–429. https://doi.org/10.1007/s11050-012-9085-2.

Schwartze, M., Rothermich, K., Kotz, S.A., 2012a. Functional dissociation of pre-SMA and SMA-proper in temporal processing. Neuroimage 60, 290–298. https://doi.org/10.1016/j.neuroimage.2011.11.089.

Schwartze, M., Tavano, A., Schröger, E., Kotz, S.A., 2012b. Temporal aspects of prediction in audition: cortical and subcortical neural mechanisms. Int. J. Psychophysiol. 83, 200–207. https://doi.org/10.1016/j.ijpsycho.2011.11.003.

Schwarz, F., 2007. Processing presupposed content. J. Semant. 24, 373–416. https://doi.org/10.1093/jos/ffm011.

Schwarz, F., 2016. Experimental work in presupposition and presupposition projection. Annu. Rev. Ling. 2, 273–292. https://doi.org/10.1146/annurev-linguistics-011415-040809.

Schwarz, F., Tiemann, S., 2017. Presupposition projection in online processing. J. Semant. 34, 61–106. https://doi.org/10.1093/jos/ffw005.

Scott, S.K., Johnsrude, I.S., 2003. The neuroanatomical and functional organization of speech perception. Trends Neurosci. 26, 100–107. https://doi.org/10.1016/S0166-2236(02)00037-1.

Scott, S.K., Rosen, S., Wickham, L., Wise, R.J., 2004. A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. J. Acoust. Soc. Am. 115, 813–821. https://doi.org/10.1121/1.1639336.

Seghier, M.L., 2013. The angular gyrus: multiple functions and multiple subdivisions. The Neuroscientist 19, 43–61. https://doi.org/10.1177/10738S8412440596.

Simons, M., 2003. Presupposition and accommodation: understanding the stalnakerian picture. Philos. Stud. 112, 251–278. https://doi.org/10.1023/A:1023004203043.

Singh, R., Fedorenko, E., Mahowald, K., Gibson, E., 2016. Accommodating presuppositions is inappropriate in implausible contexts. Cogn. Sci. 40, 607–634. https://doi.org/10.1111/cogs.12260.

Stalnaker, R., 2002. Common ground. Linguist. Philos. 25, 701–721. https://doi.org/10.1023/A:1020867916902.

Tettamani, M., Manenti, R., Della Rosa, P.A., Falini, A., Perani, D., Cappa, S.F., Moro, A., 2008. Negation in the brain: modulating action representations. Neuroimage 43, 358–367. https://doi.org/10.1016/j.neuroimage.2008.08.004.

Thiebaut de Schotten, M., Dell'Acqua, F., Valabreque, R., Catani, M., 2012. Monkey to human comparative anatomy of the frontal lobe association tracts. Cortex 48, 82–96. https://doi.org/10.1016/j.cortex.2011.10.001.

Tiemann, S., Schmid, M., Bade, N., Rolke, B., Hertrich, I., Ackermann, H., et al., 2011. Psycholinguistic evidence for presuppositions: on-line and off-line data. In: Reich, I., Horch, E., Pauly, D. (Eds.), Proceedings of the 2010 Annual Conference of the Gesellschaft für Semantik, Sinn und Bedeutung 15. Universaar – Saarland University Press, Saarbrücken, Germany, pp. 581–595.

Tromp, D., Dufour, A., Lithfous, S., Pebayle, T., Despres, O., 2015. Episodic memory in normal aging and Alzheimer disease: insights from imaging and behavioral studies. Ageing Res. Rev. 24, 232–262. https://doi.org/10.1016/j.arr.2015.08.006.

Tulving, E., 1995. Organization of memory: quo vadis? In: Gazzaniga, M.S. (Ed.), The Cognitive Neurosciences. MIT Press, Canada, pp. 839–847.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289. https://doi.org/10.1006/nimg.2001.0978.

Uddin, L.Q., Supekar, K., Amin, H., Rykhlevskaia, E., Nguyen, D.A., Greicius, M.D., et al., 2010. Dissociable connectivity within human angular gyrus and intraparietal sulcus: evidence from functional and structural connectivity. Cerebr. Cortex 20 (11), 2636–2646.

Vagharchakian, L., Dehaene-Lambertz, G., Pallier, C., Dehaene, S., 2012. A Temporal Bottleneck in the Language Comprehension Network. The Journal of Neuroscience 32 (26), 9089–9102.

Van Berkum, J.J.A., Brown, C.M., Hagoort, P., 1999a. Early referential context effects in sentence processing: evidence from event-related brain potentials. J. Mem. Lang. 41, 147–182. https://doi.org/10.1006/jmla.1999.2641.

Van Berkum, J.J.A., Hagoort, P., Brown, C.M., 1999b. Semantic integration in sentence and discourse: evidence from the N400. J. Cogn. Neurosci. 11, 657–671. https://doi.org/10.1162/089892999563724.

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. Cogn. Brain Res. 17, 701–718. https://doi.org/10.1016/S0926-6410(03)00196-4.

Vassal, F., Boutet, C., Lemaire, J.-J., Nuti, C., 2014. New insights into the functional significance of the frontal aslant tract: an anatomo-functional study using intraoperative electrical stimulations combined with diffusion tensor imaging-based fiber tracking. Br. J. Neurosurg. 28, 685–687. https://doi.org/10.3109/02688697.2014.889810.

von Fintel, K., 2008. What is presupposition accommodation, again? Philos. Perspect. 22, 137–170. https://doi.org/10.1111/j.1520-8583.2008.00144.x.

Wiecki, T.V., Frank, M.J., 2013. A computational model of inhibitory control in frontal cortex and basal ganglia. Psychol. Rev. 120, 329–355. https://doi.org/10.1037/a0031542.

Zhang, J.X., Leung, H.-C., Johnson, M.K., 2003. Frontal activations associated with accessing and evaluating information in working memory: an fMRI study. Neuroimage 20, 1531–1539. https://doi.org/10.1016/j.neuroimage.2003.07.016.