



Published in final edited form as:

Neuroimage. 2016 August 1; 136: 10–25. doi:10.1016/j.neuroimage.2016.05.020.

Sentence Understanding Depends on Contextual Use of Semantic and Real World Knowledge

Sarah Tune^{1,*}, Matthias Schlesewsky², Arne Nagels³, Steven L. Small^{1,¶}, and Ina Bornkessel-Schlesewsky^{3,¶}

¹Department of Neurology, University of California, Irvine, CA, USA

²School of Psychology, Social Work and Social Policy, University of South Australia, Adelaide, Australia

³Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany

Abstract

Human language allows us to express our thoughts and ideas by combining entities, concepts and actions into multi-event episodes. Yet, the functional neuroanatomy engaged in interpretation of such high-level linguistic input remains poorly understood. Here, we used easy to detect and more subtle “borderline” anomalies to investigate the brain regions and mechanistic principles involved in the use of real-world event knowledge in language comprehension. Overall, the results showed that the processing of sentences in context engages a complex set of bilateral brain regions in the frontal, temporal and inferior parietal lobes. Easy anomalies preferentially engaged lower-order cortical areas adjacent to the primary auditory cortex. In addition, the left supramarginal gyrus and anterior temporal sulcus as well as the right posterior middle temporal gyrus contributed to the processing of easy and borderline anomalies. The observed pattern of results is explained in terms of (i) hierarchical processing along a dorsal-ventral axis and (ii) the assumption of high-order association areas serving as cortical hubs in the convergence of information in a distributed network. Finally, the observed modulation of BOLD signal in prefrontal areas provides support for their role in the implementation of executive control processes.

Keywords

borderline anomalies; fMRI; language processing; prediction; semantic processing

1. Introduction

Human language, in conjunction with other sensory input, constitutes an important source of information that helps us to interact safely and successfully with our environment. To *make*

*Corresponding author: Sarah Tune, Ph.D., Department of Psychology I, University of Lübeck, Maria-Goeppert-Str. 9a, 23562 Lübeck, Germany, phone: +49 – 451 – 3101 – 3627, fax: +49 – 451 – 3101 – 3604, sarah.tune@uni-luebeck.de.

¶These authors contributed equally to this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

sense of linguistic input, i.e., to associate it with known objects, concepts, actions, and events, we can draw upon a myriad of relations that comprise our extensive semantic knowledge. An individual's semantic knowledge, or semantic memory, is grounded in and continuously shaped by perceptual experience. It not only enables the linking of word forms to objects, individuals, actions and abstract concepts, but also provides information on their respective attributes, purposes or intentions, and the associations among them. As such, semantic memory plays an important role in guiding action.

One of the prevailing questions addressed in recent neurobiological models of semantic memory is how modality-specific aspects of memory (e.g., reflecting different sensory, motoric or affective dimensions) are bound together and whether, in addition to the binding of these components, there is evidence for modality-independent “hubs” or “convergence zones” (Damasio 1989; Damasio et al. 2004; Martin 2007; Patterson et al. 2007; Binder and Desai 2011). The model proposed by Binder and Desai, for example, differentiates between low-level modal and high-level supramodal convergence zones: “*Modality-specific sensory, action, and emotion systems (...) provide experiential input to high-level temporal and inferior parietal convergence zones (...) that store increasingly abstract representations of entity and event knowledge*” (Binder and Desai 2011). From this perspective, high-level convergence zones, particularly those involving parts of the inferior parietal lobule (IPL), are deemed important for the emergence and use of event knowledge by serving to bind together a wealth of modality-specific information converging onto entities and events that interact in space and time to make up a particular episode. Numerous studies on language processing have used descriptions of very simple events (e.g., “The man on vacation lost a bag and a wallet”; Humphries et al. 2007) to investigate semantic processing at the single sentence level, and many of these study have implicated the angular gyrus and in some cases the adjacent supramarginal gyrus as critical regions (Ni et al. 2000; Friederici et al. 2003; Newman et al. 2003).

Beyond the level of semantic knowledge required for the interpretation of these simple events, however, semantic memory must also encompass information on broader episodes arising from the combination of multiple events. Here, we use the term scenario to denote such complex sets of events. Scenarios can describe a wide array of topics, covering a continuum that ranges from unique personal experiences or particular historic events (e.g., the Titanic's maiden voyage) to recurring patterns of events that can be generalized in so-called event schemata or cognitive scripts (e.g., hosting a dinner party). Independent of whether they describe unique or generalizable situations, scenarios are defined by the involvement of specific entities and their respective roles in the series and/or co-occurrence of individual events.

However, little research has been conducted to date examining the functional neuroanatomy underlying the processing of such complex stimuli: the majority of studies on semantic processing using language have generally focused on single words that vary along different dimensions or on the comparison of semantically congruent and incongruent single sentences (e.g., Ni et al. 2000; Newman et al. 2001; Vandenberghe et al. 2002; Friederici et al. 2003; Humphries et al. 2006; 2007). Moreover, at the sentence level, these types of manipulations typically involve a word or phrase that has a poor fit to the global context

(e.g., “*The thunderstorm was ironed*” vs. “*The shirt was ironed*”; example from Friederici et al., 2003). Given the lack of semantic association between the incongruent element and the surrounding context, it is unclear whether these types of experimental manipulations are adequate for tapping into the neural mechanisms of higher-level semantic processing. One might argue that changes in brain activity in response to such incongruent words or phrases embedded in single sentences reflect a lack of normally (ecologically) occurring contextual priming as opposed to an evaluation of message-level plausibility.

There are, however, other semantic anomalies, termed “anomalies at the borderline of awareness” or simply “borderline anomalies” (Sanford et al. 2011) that represent an interesting exception. Borderline anomalies - the “Moses Illusion” being the most famous example – are particularly hard to detect because the anomaly-inducing word has a strong semantic relationship to the meanings of the words and even to the propositional content of the context. In other words, the critical word is closely associated with the scenario as a whole but it does not constitute the correct filler for the “slot” in which it occurs. Borderline anomalies frequently include incorrect but closely related protagonists who perform or are affected by a described action while in other cases the action itself may be incorrect but again strongly associated with the scenario as a whole. The original Moses Illusion (“*How many animals of each kind did Moses take on the ark?*”) is an example of a distorted question that falsely presents the biblical figure Moses as the builder and sailor of the ark. Due to the high degree of semantic similarity (in most people's semantic memory) between Moses and Noah, the correct agent for the role, very few people notice the anomaly (Erickson and Mattson 1981). Other studies have embedded these types of semantic illusions in more elaborate contexts, thus creating the types of scenarios under discussion here:

“A pay dispute between lorry drivers and their employers reached a crisis in negotiation; even the professional mediators seemed dejected. After five days of discussion the government rejected outright the final conciliatory pay-offer and halted the talks” (Sanford et al. 2011).

In this case the incorrect term government is again closely related to the global scenario but its assigned role is unexpected given what we know about these kinds of situations. Scenarios of this type thus permit investigation of language comprehension in a way that requires the application of higher-level event knowledge, as it is not confounded with a lack of semantic association between the critical word or phrase and the overall scenario.

To date, several studies have investigated the processing of borderline (and easy to detect) anomalies using EEG, eye-tracking or fMRI. The fMRI study (Raposo and Marques, 2013) contrasted obvious semantic anomalies with more subtle incongruities in single sentences (e.g., “It was the terrible stepmother who tried to kill Cinderella with a poisoned apple”). The authors found higher levels of BOLD activity for subtle anomalies compared to true statements in the right IPL, which they link to processes of generating and integrating inferences. Comparing detected to undetected subtle anomalies increased activity in the dorsolateral prefrontal cortex, orbitofrontal cortex and insula, which they link to conflict monitoring and error detection. However, the choice of it-cleft structures as a means of directing the listeners' attention to a particular part of the sentence strongly discourages the integration of all involved components into a global meaning– the very process we aim to

investigate. Moreover, it is unclear whether neural responses to isolated it-cleft sentences are reflective of brain mechanisms underlying ecological language comprehension, as these structures cannot be felicitously uttered without supporting context.

The EEG and eye-tracking studies contrasted obvious and borderline anomalies embedded in sentence pair structures or short paragraphs (Daneman et al. 2007; Bohan and Sanford 2008; Sanford et al. 2011; Bohan et al. 2012; Tune et al. 2014). The results of these studies have provided important insights into the time course of semantic processing in context but they only allow limited conclusions about the underlying neural circuitry.

Thus, given the scarcity of functional neuroimaging studies focused on the question of how the brain accomplishes the interpretation of complex scenarios including the integration of multiple events and the evaluation of the derived global message against real-world knowledge, the neural mechanisms and circuitry involved in implementing these processes remain insufficiently characterized. Our study attempts to close this gap by taking advantage of the properties of contextually embedded borderline anomalies that cannot be categorized as anomalous based on a lack of semantic priming and therefore require a thorough semantic analysis and the application of world knowledge. As such, they provide a novel way of studying the brain mechanisms that support ecological language comprehension.

1.1 The present study

In the present study we set out to investigate the neuroanatomical substrates underlying the use of semantic knowledge to understanding sentences in context, with a particular focus on the interpretation of scenarios that call for extensive evaluation. More precisely, we were interested in identifying the set of brain regions and mechanistic principles implicated in the application of real-world event knowledge to sentence interpretation, i.e., the assessment of whether a described relation between events and their respective participants is consistent with expectations derived from previous experience. Further, we aimed to tease apart changes in brain activity that are the result of high-level comprehension processes from those that reflect processes at lower levels, e.g., effects of lexical-semantic associations. Finally, we aimed to address these issues in a way that would provide group-level results of high anatomical precision.

The experimental design and analysis of the present study address our specific goals in the following ways. We contrasted two different types of scenarios that were created by embedding borderline anomalies and easy (obvious) anomalies in a richer context. Each anomalous sentence was paired with a closely matched non-anomalous control sentence, thus yielding a 2×2 design with factors Plausibility (anomalous vs. non-anomalous) and Scenario Type (borderline vs. easy). Our study design builds on the logic that plausible, non-anomalous scenarios are more likely to confirm our expectations, whereas implausible, anomalous scenarios correspond less closely to our beliefs and assumptions about the world. The use of borderline anomalies is key to our goal of tapping into the mechanisms implementing high-level comprehension processes, since, in contrast to easy-to-detect anomalies, they cannot be classified as being anomalous simply based on the poor semantic fit of a single word or phrase and the broader semantic context. This specific design has been

adapted from that successfully employed in several EEG studies (Tune et al. 2014; Sanford et al., 2011).

Based on results from previous EEG and eye-tracking studies suggesting that detected and non-detected trials are processed differently (Bohan and Sanford 2008; Sanford et al. 2011; Bohan et al. 2012; Tune et al. 2014), we used a plausibility judgment task that would allow us to distinguish these trials from one another. Finally, to obtain group-level results of high anatomical precision, we used a surface-based approach in conjunction with statistical analyses at the levels of both whole brain and anatomically defined regions of interest (ROIs).

For the regional analyses, we focused on the left and right superior and middle temporal cortices, inferior parietal lobule, prefrontal, and insular cortices as these brain areas are frequently found to be sensitive to semantic manipulations at different levels of language comprehension (for reviews see Van Petten and Luka 2006; Vigneau et al. 2006; Lau et al. 2008; Binder et al. 2009; Price 2010; Binder and Desai 2011) and are thought to play key roles in the ventral and dorsal functional streams of auditory language processing (Hickok and Poeppel 2007; Saur et al. 2008; Friederici 2009; 2012; Bornkessel-Schlesewsky and Schlesewsky 2013; Bornkessel-Schlesewsky et al. 2015).

We had the following four specific hypotheses. First, since scenario interpretation requires both the convergence of partly modality-specific information about entities and events and its integration into more complex patterns in space and time (Damasio 1989; Damasio et al. 2004), we expected involvement of both modal and supramodal regions in the temporal and inferior parietal lobes.

Second, with regard to the comparison of borderline and easy anomalies, we assumed that the former would draw more strongly on multimodal regions in the posterior middle temporal gyrus, and more importantly, the inferior parietal lobule, as the interpretation and detection of borderline anomalies rely on an extensive analysis and application of event knowledge. The angular and supramarginal gyri of the IPL are thought to be involved in the combination and manipulation of high-order configuration of entities and actions set in time and space (Binder et al. 2009), which is why we expected these regions to be particularly sensitive to scenarios describing conceivable but nevertheless unexpected events.

Third, we expected easy anomalies to rely more on regions of the anterior portion of the superior temporal gyrus and sulcus, which have lower multimodal and integrative roles and appear to be involved in the gradual abstraction within one given modality. The latter hypothesis is based on the assumption that superior temporal regions forming part of the ventral stream of auditory language processing engage in the mapping of spectro-temporal patterns to increasingly complex auditory objects that encode concepts of varying complexity as well as associations between these concepts (Rauschecker and Scott 2009; Bornkessel-Schlesewsky and Schlesewsky 2013; Bornkessel-Schlesewsky et al. 2015).

Fourth, we expected certain brain responses to be engendered by the plausibility judgment task itself. Recall that the use of a behavioral task is essential to the design of our study since the classification of borderline anomalies into detected and non-detected trials would not be

possible otherwise. In accordance with accounts that propose a domain general role of the left IFG in the implementation of cognitive control mechanisms (Thompson-Schill et al. 1997; Miller and Cohen 2001; Thompson-Schill et al. 2005; Bornkessel-Schlesewsky and Schlewsky 2013), we suggest that this region should show differential degrees of activation as a function of classification difficulty across the critical sentence conditions in our study. Based on previous results, we expect a similar response profile – i.e. higher activation with higher decision uncertainty – for left frontal regions adjacent to the IFG, and particularly the anterior insula and deep frontal operculum (Grewe et al., 2005; Bornkessel-Schlesewsky, Schlewsky, & von Cramon, 2009). With respect to the decision-making processes leading up to the behavioral responses, we hypothesized that the classification of easy anomalies would be least demanding, while the degree of uncertainty should be higher for the plausible, easy-type control stimuli. In contrast to the easy anomalies, in which the non-fitting critical word provides a clear cue that the sentence is anomalous, the easy control stimuli do not include a clear cue for a judgment one way or the other. Finally, as the borderline stimuli are by definition difficult to classify as anomalous or non-anomalous, we expect that the decision uncertainty will be most pronounced for these sentences.

2. Materials and Methods

2.1 Participants

Twenty-two healthy, right-handed (indicated by an adapted German version of the Edinburgh Handedness Inventory; (Oldfield 1971), monolingually raised, native speakers of German participated in the fMRI study (11 women, mean age 24.5, range 20-30). All participants were recruited at the University of Marburg. They had normal or corrected-to-normal vision and no known neurological, psychiatric or auditory disorders. Prior to the scanning session, participants were screened to ensure that they met all safety and participation requirements and they gave written informed consent. Four participants had to be excluded: two due to excessive motion, one due to structural abnormalities and one because of an incomplete functional run. The study was approved by the ethics committee of the Medical Faculty at the University of Marburg.

2.2 Materials

The materials of the present study were identical to those used in the German ERP experiments reported in Tune et al. (2014). They were a translated and adapted version of the English stimuli employed in an ERP study by Sanford et al. (2011). The experimental conditions (see Table 1 for examples) belonged to two types of semantic violations: hard to detect (“borderline”) anomalies and more classic easy to detect anomalies (i.e., words with a poor fit to the context). Prior to the experiment, the materials were pre-tested in two separate questionnaire studies. The first study (n=70) ensured that borderline anomalies were reliably missed some of the time while easy anomalies were detected at least 95% of the time. The detection rates were 68.1 % (sd: 26.2%) for borderline anomalies and 98.4 % (sd: 8.6%) for easy anomalies. The second study (n=20) asked participants to rate the contextual fit of the anomaly-inducing word to the described scenario on a 7-point Likert scale. The ratings indicated a significantly better contextual fit for borderline anomalies compared to easy anomalies (mean ratings: 5.02 and 2.20; $t(210.201)=18.54$, $p<0.0001$).¹

In total, 180 stimulus sets consisting of an anomalous condition and a corresponding congruent control condition were constructed, 120 pairs for the borderline anomalies and 60 pairs for the easy to detect anomalies. All stimuli were composed of two semantically connected sentences, with the first sentence providing context, and a second, critical sentence, containing a target word that determined whether the sentence was semantically anomalous or not. All critical sentences consisted of 17 words. In the following, the different layouts will be described on the basis of German and English examples listed in Table 1.

Borderline anomalies—In the borderline anomaly condition, the second sentence was a thematic continuation of the context sentence and contained two alternative local context words (highlighted in italics in Table 1) and the critical target word (set in bold print in Table 1). The target word was always in the 13th word position and separated by five words from the contextual manipulation. It is the relation between local context word/phrase and target word that determines whether the latter is perceived to be anomalous. However, borderline anomalies are often missed because the target word is highly associated semantically to the overall context, despite its implausibility for the meaning of the particular sentence in which it appears.

Easy to detect anomalies—Sixty sets of easy to detect anomaly sentences were used. These sentence pairs were constructed in a similar manner to the borderline anomalies, but the internal structure of the critical sentences was less standardized because in these cases the anomaly is elicited by a single word with a very poor fit to both local and global context. The critical words (highlighted in bold print in Table 1) appeared in different positions across stimulus sentences. This ensured that participants had to pay close attention to the whole sentence and could not predict the critical region in the stimuli presented.

The duration of context sentences across conditions ranged from 3.5 to 10.5s (mean = 5.4s \pm 1.5s), while the critical sentence had a more standardized length with a mean of 6.4s (\pm 0.8s) and a range of 5-8.5s. Auditory stimuli were recorded by a female native speaker of German who read the sentences with clear and natural intonation. To ensure equal volume levels, the stimuli were normalized digitally. On average, the target word was presented 4.0s (\pm 0.64s) after sentence onset in borderline anomaly sentences and 3.73s (\pm 1.3s) in easy to detect anomaly sentences.

The stimuli were distributed across eight lists, each composed of 180 sentence pairs, consisting of 120 borderline and 60 poor fit stimuli. While poor fit stimuli were divided evenly, of the 120 borderline items, 90 contained an anomaly and the remaining 30 were plausible controls. We employed this asymmetrical design, adopted from Sanford et al. (2011), to obtain a similar number of trials for each of the three experimental conditions (detected anomalies, missed anomalies, plausible controls) for the analysis. Within a final list, an item was presented as either an anomalous or control condition, while across all lists, each condition of a stimulus pair was presented at least once. This was accomplished by rotating the borderline materials over four lists with 120 stimuli each. Poor fit materials were

¹For further information on the translation of materials and a description of the questionnaire studies, please refer to the details provided in Tune et al. (2014).

divided into two lists consisting of 60 stimuli. Merging each borderline list with each poor fit list yielded the eight final lists. The order of stimuli was then pseudo-randomized.

2.3 Procedure

Prior to the imaging experiment, participants performed a short training session. Stimuli used for training purposes were not part of the imaging experiment. In the experimental sessions, participants listened to the auditory stimuli via MRI compatible noise-cancelling headphones. Visual cues were displayed on a monitor placed at the rear of the scanner and viewed by the participants via a mirror attached to the head coil. Each trial started with the presentation of a fixation cross in the center of the screen, followed by the auditory presentation of the context and critical sentence. After the second sentence had ended, the fixation cross was replaced by an answer screen that served as a cue for the participants to indicate via button press using their right index and middle finger whether they had detected an anomaly. A schematic description of this paradigm is shown in Figure 1. In order to avoid anticipatory motor response preparation following the processing of the critical word, we did not employ a fixed assignment of response buttons to the “plausible” and “implausible” categorizations per participant. Rather, the assignment of the left and right buttons to “plausible” and “implausible” responses varied on a trial-by-trial basis and was signaled by two smiley faces (one laughing and one frowning). Across each session, the assignment of the “plausible” and “implausible” categories to the left and right buttons was counterbalanced. The maximal response time was set to 2000ms. Onset and timing of stimuli in the event-related design was optimized for efficiency by introducing short jittered rest periods between the presentation of the context and critical sentence, as well as between the offset of the critical sentence and the presentation of the response screen and after the 2000ms response interval and the beginning of the next trial. Jittered rest intervals had a mean duration of 1020ms (\pm 622ms) and a range of 500-3000ms. Each participant was presented with one of the eight lists that were split into six 10-minute runs of 30 trials each. Between each pair of experimental runs, we collected resting state data for one minute.

2.4 fMRI data acquisition

Functional imaging was performed on a 3T MR Magnetom, Tim Trio scanner (Siemens Medical Solutions, Erlangen Germany). Twenty-nine axial slices were acquired using a T2-weighted gradient-echo echo-planar image (EPI) sequence optimized for blood oxygen level dependent (BOLD) effects (TR = 1500ms, TE = 21ms, flip angle = 71°, slice thickness = 4mm, in-plane resolution = 3.8 × 3.8mm, 1mm gap, FOV = 240mm, matrix dimension = 64 × 64). Four dummy scans recorded at the beginning of the functional sequence were discarded to account for magnetization stabilization effects. A total of 2750 functional images were acquired. In addition, a high-resolution anatomical image was collected for each participant using a T1-weighted MPRAGE sequence (TR = 1900ms, TE = 2.52ms, flip angle = 9°, FOV = 256mm, 176 slices, voxel size = 1.0 × 1.0 × 1.0mm).

2.5 fMRI data analysis

Analysis of the neuroimaging data was conducted using Analysis of Functional Neuroimages / Surface Mapping (AFNI/SUMA; <http://afni.nimh.nih.gov>; Cox 1996; 2012; Saad and Reynolds 2012) and Freesurfer software packages (<http://>

surfer.nmr.mgh.harvard.edu; Dale et al. 1999; Fischl 1999; 2004). The ROI analysis was carried out using the statistical software package R (<http://www.R-project.org>; R Core Team 2014).

2.5.1 Pre-processing—All functional time series were subjected to the following pre-processing steps in the native volume domain: de-spiking, slice-timing correction, mean normalization, three-dimensional affine motion correction using a weighted linear least squares cost function for the alignment of three translational (x,y,z) and three rotational (pitch, roll, yaw) parameters, registration to the middle image of the functional run and to the structural image. Time points that showed excessive motion (> 1mm) were censored and subsequently excluded from regression analysis.

2.5.2 Single subject analysis—To assess the degree of event-related BOLD activity against an implicit resting baseline consisting of all jittered intervals plus the resting state data collected between runs, each voxel's time series was analyzed by multiple linear regression analysis. To model the hemodynamic response (HDR), we used fixed-shape impulse response functions (IRF) for context sentences and behavioral responses and data-driven deconvolution of the HDR for target sentences. AFNI's piecewise linear spline method allows the temporal shape of the response model to vary across voxels and thus provides a more accurate and flexible estimate of the brain response. Since we were primarily interested in the brain's response to target sentences, context sentences and behavioral response intervals were each collapsed across conditions and included in the analysis as single regressors. Context sentences showed substantial variation in length, this regressor was therefore convolved with a duration modulated square wave function by specifying onset and duration of each trial, while the regressor modeling the response was convolved with a fixed length (2s) square wave function. Depending on the behavioral response, target sentences were divided into five conditions of interest (*borderline detected*, *borderline non-detected*, *borderline control*, *easy detected*, *easy control*) and a condition of no interest that included false rejections and false alarm responses. The HDR to each target sentence condition was deconvolved by linear interpolation using AFNI's tent basis function. The HDR was modeled from the onset of the target sentence for a period of 21s to account for the initial delay and to cover the entire duration of the stimulus as well any ongoing processing after the end of the sentence. Each tent spanned an interval of 1.5s matching the length of a single volume acquisition (which equals the TR in this case). Mean, linear and quadratic trend components along with the six parameters obtained from the spatial alignment procedure were included as nuisance regressors. Finally, to further minimize the influence of nuisance variance unlikely to reflect signal of interest, we regressed out signal from both lateral ventricles and bilateral white matter (Fox 2005; Dick et al. 2010).

The analysis resulted in regression weights (reflecting percent signal change relative to the implicit resting baseline) along with an associated t statistic assessing their reliability. For regressors modeled with a fixed shape impulse response function (IRF), this first-level single-subject analysis resulted in individual regression coefficients per voxel and condition, while the data-driven, deconvolution analysis used for target sentences yielded an estimated HDR shape described by 15 regression coefficients per voxel and condition. All subsequent

analyses used the average of the regression coefficients modeling the brain response in the time-window of 3-12s post stimulus onset. The choice of this time window respects the initial delay in the hemodynamic response and accounts for the entire duration of the target sentence.

In preparation for second-level group analysis, we used Freesurfer to create two-dimensional surface models from each participant's high-resolution anatomical scan. White and grey matter of the anatomical volumes were automatically segmented, the results inspected and manually improved when needed, separate cortical surfaces for each hemisphere were inflated and brought into registry with a common template of average gyral and sulcal folding (Dale et al. 1999; Fischl 1999; Fischl et al. 1999). Using SUMA, the individual surface models were then imported into the AFNI three-dimensional space and regression coefficients obtained from first-level analysis were interpolated from the three-dimensional volume domain to the two-dimensional surface model. To decrease spatial noise due to inter-individual variation, functional data were smoothed on the surface using a 6mm full width at half maximum (FWHM) heat kernel (Chung et al. 2005). Finally, for display purposes an average of all 18 individual cortical surface representations was created using Freesurfer.

2.5.3 Region of interest group analysis—For this analysis, we chose a set of 14 anatomically defined cortical regions per hemisphere using the automated parcellation scheme implemented in Freesurfer (Fischl 2004; Destrieux et al. 2010). The parcellation procedure that yields 74 cortical regions per hemisphere uses an algorithm based on macroanatomical landmarks that follow the anatomical conventions and nomenclature described by Duvernoy (1999). As we were mostly interested in examining the neural responses of regions that are i) situated along the ventral and dorsal streams of auditory language processing and ii) known to be sensitive to semantic manipulations in language stimuli, we focused on a set of bilateral frontal, temporal and inferior parietal regions, and in line with established practices in our lab, manually subdivided some of them into smaller sub-regions known to show distinct structural connectivity profiles (Dick et al. 2009; 2012; Mashal et al. 2012). A detailed definition of all ROIs is given in Table 2. Frontal regions encompassed the anterior insula (INSa), pars opercularis (IFG_Op), triangularis (IFG_Tr) and orbitalis (IFG_Orb) of the inferior frontal gyrus, while inferior parietal regions included the supramarginal (SMG) and angular gyri (AG). For the temporal regions, we subdivided the lateral superior temporal gyrus into an anterior (STGa) and a posterior portion (STGp) using a vertical plane drawn from the anterior tip of the transverse temporal gyrus as the boundary. The transverse temporal gyrus and sulcus were combined to form one region describing the primary auditory cortex (AUD). Using the same delimiting landmark as the STG, the superior temporal sulcus was partitioned into an anterior and a posterior portion that were subsequently further subdivided into their respective upper and lower banks with the fundus of the sulcus serving as boundary. This yielded four superior temporal regions per hemisphere (STSa_lower, STSa_upper, STSp_lower, STSp_upper). Lastly, we included the posterior middle temporal gyrus (MTGp) that was delimited from the anterior segment of the gyrus by the same vertical plane described above. All sub-regions were defined in reference to known differences in structural connectivity gleaned from non-human primate studies (Jones and Powell 1970; Seltzer and Pandya 1978; Petrides and Pandya 1984;

Matelli et al. 1986; Seltzer and Pandya 1989; 1994; Gloor 1997; Romanski et al. 1999; Rizzolatti and Matelli 2003; Schmahmann et al. 2007).

At the level of the single subject, we extracted mean beta values per region and experimental condition focusing only on correctly classified trials (*borderline detected*, *borderline control*, *easy detected*, *easy control*). At the group-level, we then ran repeated-measures ANOVAs including the factors Scenario Type (borderline vs. easy) and Plausibility (anomalous vs. non-anomalous) across all ROIs and proceeded by resolving significant interactions by Scenario Type. To correct for multiple comparisons across ROIs we applied an FDR procedure (Benjamini and Hochberg, 1995; Genovese et al. 2002).

2.5.4 Whole-brain group analysis—A mixed effects (repeated-measures) analysis of variance (ANOVA) with CONDITION (5) as a fixed factor and PARTICIPANT (18) as a random factor was performed on a vertex-by-vertex basis using the normalized beta weights from each individual's regression model as the dependent variable. Simple contrasts against the resting baseline were obtained for each condition. Following the results of the ROI analysis that revealed robust and widespread differences between the two scenario types (see Results section), we decided to restrict the exploratory whole-brain contrasts to comparisons within each type (*borderline detected* versus *borderline control* and *easy detected* versus *easy control*)². A Monte Carlo simulation approach (cf Forman et al. 1995) was used to identify significant clusters of activated vertices, with an individual vertex threshold of $p < 0.001$ for simple contrasts against resting baseline and $p < 0.01$ for comparisons among conditions, and a family-wise error (FWE) correction at the whole-brain level at $p < 0.05$. The difference in individual vertex thresholds is motivated by results obtained from signal-to-noise analyses that indicate a power advantage for contrasts against baseline compared to contrasts between conditions (Dick et al. 2009). Given the individual vertex thresholds, the Monte Carlo simulation determined a required minimum cluster size of 83 nodes for simple contrasts and 234 nodes for contrasts between conditions.

3 Results

3.1 Behavioral results

Analysis of the behavioral data showed that the detection rate for borderline anomalies was 78.1% ($\pm 8.5\%$), while easy to detect anomalies were correctly judged as implausible at a rate of 96.8% ($\pm 4.6\%$). Sentences without anomalies were correctly judged as plausible at a rate of 76.9% ($\pm 11.5\%$) for borderline and 91.1% (± 4.4) for easy scenarios.

3.2 Neuroimaging results

3.2.1 Whole brain comparisons against baseline—We first computed both increases and decreases in BOLD signal for each of the five conditions with respect to the resting baseline at the level of the whole brain. The results of these analyses are presented in Figure

²In a third contrast, which was not directly related to our specific research questions, we also compared the brain responses to non-detected borderline anomalies and borderline control sentences. The contrast, presented in Figure S3 of the supplementary materials, revealed only a single bilateral cluster in the ventral part of the central sulcus reflecting higher levels of activity for control sentences compared to non-detected anomalies. Therefore, the results do not provide any evidence for a subconscious detection of the anomaly.

2. All five conditions show signal changes across a large set of activation clusters in the bilateral frontal, temporal and parietal lobes. Even though most of these activation clusters are bilateral, the activation intensity and volume tends to be greater in the left hemisphere. Areas showing an increase in brain activity relative to the baseline include regions of the bilateral lateral and medial frontal cortices, the superior and middle temporal cortices, parts of the left inferior parietal lobe, and the posterior thalamus bilaterally. Although this thalamic activation appears to be in the pulvinar, the chosen surface-based analysis approach does not provide complete or reliable results in areas of subcortical grey matter. To make a more substantiated statement about the role of subcortical structures like the thalamus, amygdala or basal ganglia, a separate analysis that encompasses these regions would be necessary.

Frontal activation is most pronounced and widespread in the three borderline conditions with bilateral activation covering large parts of the anterior insula, lateral orbital sulcus, pars opercularis and triangularis of the inferior frontal gyrus, the inferior frontal sulcus extending into the middle frontal gyrus, and the medial portion of the superior frontal gyrus. In the temporal lobe, all borderline conditions show a cluster of positive signal change in the left – and in the *borderline detected* condition also the right – superior temporal sulcus that extends into the middle temporal gyrus. In the parietal lobe, all conditions elicited an increase in activity in the left intraparietal sulcus, and for the two conditions involving successful anomaly detection, also in the neighboring left supramarginal gyrus.

The examination of negative signal deflections from baseline revealed large clusters centered bilaterally around the middle and posterior portions of the superior temporal lobe that also include portions of the subcentral gyrus and sulcus, the central sulcus and in some cases the inferior portion of the supramarginal gyrus. Additional clusters with higher levels of activity during baseline than during language processing are found in the medial portions of the parietal and occipital lobes bilaterally. In principle, there are at least two possible explanations for the observed negative signal deflections (“deactivation”) from baseline. One explanation relates to the fact that our implicit baseline included longer periods of rest collected between experimental runs. During this time, participants may have engaged in daydreaming or mind-wandering which has been shown to draw on similar cognitive mechanisms and anatomical areas as many semantic tasks (Binder et al., 1999; Humphreys et al., 2015). However, to explain the observed pattern along these lines, one would need to assume that conceptual processing during rest led to higher levels of neural activity than semantic processes underlying scenario comprehension. In the present study, a more likely explanation is that the decrease found in regions of the primary auditory cortex and neighboring areas reflects effects of habituation and concomitant repetition suppression in the BOLD signal. We suggest that these effects are due to the persistent auditory stimulation throughout the entire duration of the trials as context and target sentences were separated only by a short jittered inter-stimulus interval.³ These findings will not be further addressed in the discussion.

³This interpretation is corroborated by the following observations: The deactivation is most pronounced in the primary auditory cortex and decreases as a function of distance from this area. Moreover, the activation found for preceding context sentences merged across

3.2.2 ROI Analysis—In the ROI analysis, we examined changes in brain responses in 14 regions per hemisphere with the main goal of identifying brain regions that are preferentially responsive to either borderline or easy anomalous sentences compared to their plausible counterparts, i.e., regions showing an interaction of Plausibility and Scenario Type. In line with our hypotheses, significant interactions were resolved by Scenario Type. In addition, we also wanted to investigate if any of the included ROIs would show higher (or lower) levels of hemodynamic activity for anomalous compared to non-anomalous sentences irrespective of the particular scenario type (main effect Plausibility). The results of the ROI analysis are shown in Figure 3 below and the effects per region are described in detail in Table 3.

The most stable pattern revealed by the analysis was a main effect of Scenario Type, reflecting a difference in activation between the two scenario types irrespective of whether or not they contained an anomaly. This effect was found in 22 of the 28 examined regions. Among these regions sensitive to the Scenario Type, three showed an additional and independent main effect of Plausibility: the upper and lower bank of the left anterior STS and the right posterior MTG. In all but one region, the right SMG, the effect reflects higher levels of brain activity for borderline compared to easy target sentences.

We found a significant two-way interaction in one left hemisphere regions (AUD) and two right hemisphere regions (INSA and STGp). When resolved by Scenario Type, all regions showed a significant main effect of Plausibility for easy but not for borderline sentences. Finally, the left SMG was the only region with a significant main effect of Plausibility but not Scenario Type. This main effect reflects a relative increase in signal for anomalous versus non-anomalous target sentences of both scenario types.

To rule out that the pronounced differences in activation across the two stimulus types were induced by differences in the preceding context sentences, we ran a follow-up analysis in which we split the context sentences based on the categorization of subsequent target sentences and included them as separate regressors in the analysis. Next, we ran comparisons between the two anomalous (borderline detected vs. easy detected) and the two control conditions (borderline control vs. easy control) separately for context and target sentences. The results, depicted in Figure S1 and S2 of the supplementary materials, clearly show that the extensive differences in brain responses found for the comparisons of target sentences are not mirrored by comparable effects for the context sentences. In fact, little to no differences were revealed by contrasting brain response to context sentences across stimulus types. The contrast of context sentences leading up to easy and borderline anomalies produced a single cluster in the left precentral sulcus. The comparison of contexts preceding plausible control sentences resulted in three clusters, two of which were located in the anterior cingulate gyrus of each hemisphere and one in the marginal branch of the right cingulate sulcus.

all conditions (data not shown) indicates an increase in signal relative to baseline throughout the primary auditory cortex and neighboring regions, thereby providing evidence for a neural habituation effect specific to auditory processing.

3.2.3 Whole brain: Easy detected versus easy control—In light of the widespread and pronounced differences between target sentences of the two scenario types, we decided to focus the whole brain analysis on differences between anomalous and non-anomalous target sentences for each scenario type. The results of the contrast of detected easy anomalous and (correctly classified) easy control sentences are shown in Figure 4 and Table 4. At the chosen threshold, we found eight significant clusters in the left hemisphere and nine in the right. Several of the reported clusters fall into regions examined in the ROI analysis and the observed patterns of brain responses are in line with the regional signal averages. The analysis revealed additional clusters along the medial wall of the left and right temporal and parietal lobes as well as the right prefrontal cortex. Notably, all lateral frontal clusters show higher level of BOLD signal for easy control sentences than for easy detected anomalies.

3.2.4 Whole brain: Borderline detected versus borderline control—The comparison of detected borderline anomalies with their plausible controls produced a number of significant clusters all of which reflect relatively stronger brain responses for the anomalous sentences. The results are detailed in Figure 5 and Table 4. Again, the analysis of direct contrasts at the whole brain level extends the results of the ROI analysis by revealing differential involvements of several medial wall structures. Here, detected borderline anomalies elicited stronger responses than correctly classified control sentences in the left anterior cingulate cortex, the right amygdala and both thalami. Furthermore, we found two distinct clusters in the anterior and posterior part of the superior supramarginal gyrus, which provides additional detail to the pattern observed in the ROI analysis.

4. Discussion

The aim of the present study was to shed light on how the brain accomplishes the comprehension of multiple event scenarios by using easy and borderline anomaly stimuli and their plausible counterparts. The most striking finding was unexpected, i.e., widespread and robust differences between the two scenario types that were independent of their plausibility. Since this main effect is independent of the effects of interests, i.e., plausibility effects and their modulation across the two scenario types, the discussion will only briefly touch upon this finding. Given the considerable overlap in the pattern of results revealed by whole brain direct contrasts and the regional analysis, the discussion will mostly focus on the findings of the more detailed two-factor ROI analysis.

Scenario comprehension engages a bilateral frontal-temporal-parietal network

For the comparison against baseline, we found that analysis and interpretation of all stimulus types involved a large set of brain regions in the frontal, temporal and inferior parietal lobes. In contrast to the results of single sentence studies (e.g., Ni et al. 2000; Newman et al. 2001; Vandenberghe et al. 2002; Friederici et al. 2003; Humphries et al. 2006; 2007), most of the observed clusters were bilateral, though the activation intensity and volume was found to be slightly greater in the left hemisphere. The quite extensive right hemispheric contribution observed here is in line with a growing body of evidence that supports an important role of the right hemisphere in ecological language comprehension. Virtually all studies of sentence

comprehension and discourse comprehension that do not rely on contrasts between highly similar conditions show bilateral activation (Mazoyer et al. 1993; Indefrey and Cutler 2004; Vigneau et al. 2006; Wilson et al. 2008; Dick et al. 2009; Straube et al. 2009; Dick et al. 2010). Various theories have been proposed to account for functional differentiation of the two hemispheres, but suffice it to say that an enormous network of brain activation for ecological language, encompassing many regions in both hemispheres, reflects a very large number of concurrent processes, including those at the level of sounds, words, meaning, grammar, prosody, emotion, inference, memory, and others.

Familiarity as a potential influence on scenario comprehension

The ROI analysis revealed an increase in BOLD signal for borderline compared to easy stimuli in virtually all ROIs (except for the left SMG and right AG and IFG_Op). However, it seems highly unlikely that it was caused by differences in specific low-level stimulus or design properties. Target sentences were matched for overall length as well as length and frequency of the critical word and our follow-up analysis showed that the observed differences in easy and borderline target sentences could not be attributed to the preceding context sentences. Furthermore, based on the results of the same analysis, it seems unlikely that the effect is driven by a difference in the number of trials per stimulus type. As described in the results section and shown in Figures S1 and S2, both contrasts yield the same robust effect despite the fact that trials numbers were much more similar for the contrast of control sentences across stimulus types.⁴

We reasoned *post hoc* that the widespread differences were most likely triggered by a factor that would impact the processes underlying scenario processing within a large distributed network. In interpreting scenarios, the particular components (e.g., people, objects, events and their associated modality-specific and/or modality-independent information) need to be bound together at different levels of abstraction in order to arrive at the global meaning of the received input. We hypothesized that participants were more familiar with the everyday events described in easy scenarios than with the predominantly non-recurring events described in borderline scenarios, and that the degree of familiarity has implications for how the brain accomplishes scenario comprehension.⁵

Scenario processing is implemented by a hierarchically organized neural architecture

One of our specific aims was to tease apart the processing of low-level lexical-semantic mismatches in easy anomalies and message-level scenario comprehension necessary for successful detection of borderline anomalies. Following the idea of a hierarchically organized neural architecture implementing anterior ventral and posterior dorsal streams of auditory language processing (Rauschecker and Scott 2009; Bornkessel-Schlesewsky and Schlewsky 2013; Bornkessel-Schlesewsky et al. 2015), we hypothesized that easy

⁴In fact, in the comparison of borderline control vs. easy control target sentences, more trials were included in the easy condition (due to lower accuracy for borderline sentences), whereas the opposite was true for the contrast of borderline detected vs. easy detected target sentences.

⁵To further test this hypothesis, we conducted an additional post hoc questionnaire study to assess potential differences in familiarity across the two scenario types using a group of participants comparable to that of the fMRI study. The results show that participants were indeed more familiar with the events described in easy scenarios than with those of borderline scenarios. The details of this study and an extended discussion of the results can be found in the supplementary materials.

anomalies would rely primarily on processing in early association areas of the ventral stream in the anterior superior temporal lobe, whereas borderline anomalies would engage higher-order association cortices in the dorsal pathways of the posterior temporal and inferior parietal lobes. We found an interaction of Scenario Type and Plausibility in a number of regions, including the left primary auditory cortex and right posterior STG. In these regions, the plausibility effect was observable for the easy anomalies but not for the borderline anomalies relative to the control conditions. The presence of plausibility effects for easy but not borderline anomalies has previously been found using EEG. While the processing of easy stimuli compared to controls elicits strong N400 effects, depending on the tested language and task settings, the N400 effect for detected borderline anomalies compared to controls is greatly reduced and in some cases even completely absent (Sanford et al. 2011; Bohan et al. 2012; Tune et al. 2014).

Contrary to our hypothesis, we did not find that the processing of easy anomalies preferentially engaged portions of the anterior superior temporal lobe. Nevertheless, the effects of plausibility for easy anomalies were mostly found in primary auditory cortex and adjacent association cortices in the superior temporal lobe, thus lending some support to the idea of hierarchical processing. The fact that we see these effects in early auditory regions suggest that they may reflect expectation mismatches at a sensory level due to the high degree of predictability for the correct target word. This assumption is supported by a number of observations: The majority of sentences with easy to detect anomalies described very general, everyday occurrences such as the following example (Original in German; English translation given):

Fred was feeling really tired when he got home. All he wanted was to sleep after he closed his bedroom curtains and curled up in bed/crayons.

In this example, the context renders the correct sentence final word *bed* highly predictable. In fact, as revealed by a sentence completion task, target words of easy scenarios show higher cloze probability rates than those of borderline scenarios (Heinen, 2013; unpublished B.A. Thesis). Therefore, it could be argued that the relatively high expectation for a specific continuation led to a predictability down to the level of sensory representations. This is in line with the notion of predictive coding via hierarchically organized internal models, serving to link more abstract predictions with concrete predictions at a sensory level (Friston and Kiebel 2009; Friston 2012; for an application to higher-order language see Bornkessel-Schlesewsky et al. 2015). The stronger activity observed for anomalous items would thus reflect the mismatch between a context-based expectation pertaining the physical form of a stimulus and the effectively encountered input. Evidence for this assumption stems from a series of visual MEG studies that showed that violation of form-related predictions led to an increased amplitude of an early event-related component localized in the visual cortex (Dikker et al. 2009; 2010; Dikker and Pylkkänen 2011). These results demonstrate that context-based predictions that pertain to the physical form of future input can have an impact even on the earliest sensory processing of words.

Contributions of temporal and parietal higher-order association areas to the processing of message-level plausibility

The processing of meaning requires the orchestration of a multitude of neural computations even for the simplest tasks such as object recognition and naming. Therefore, there can be no doubt that semantic analysis in more ecologically valid contexts such as scenario comprehension builds on a distributed network of regions acting in concert. In such a large-scale network, higher-order multimodal association cortices are assumed play an important role in the convergence of information across different scales, modalities and systems. The following three brain regions are frequently discussed as critical components in the neural circuitry that implements conceptual processing in language: the anterior temporal lobe, the middle temporal gyrus, and the inferior parietal lobule (Stowe et al. 2005; Binder et al. 2009; Visser et al. 2010; Binder and Desai 2011; Hoffman et al. 2012; Noonan et al. 2013; Seghier 2013).

The results of our ROI analysis found evidence for involvement of all three regions in the analysis of scenario plausibility. More specifically, we observed plausibility effects for the upper and lower banks of the left anterior STS, the left supramarginal gyrus, and the right posterior MTG. The left supramarginal gyrus was the only ROI that showed a main effect of Plausibility without an effect of Scenario Type, i.e., stronger responses to detected anomalies (easy and borderline) than to plausible control sentences. The direct contrasts between conditions performed at the whole brain level extend these findings by revealing increased BOLD signal in the right supramarginal gyrus and angular gyrus for easy anomalies compared to control sentences.

A detailed review of the literature on the role of all three regions in the processing of semantic information in linguistic contexts is beyond the scope of this discussion. We will thus focus on the most robust and novel result - the involvement of the supramarginal gyri in both hemispheres (and the angular gyrus of the right hemisphere) – and only briefly address the contributions of anterior and middle temporal cortices as part of a distributed semantic network.

Inferior parietal lobule—The inferior parietal lobule, encompassing the supramarginal and angular gyri, is a functionally heterogeneous brain region consistently implicated in neuropsychological and neuroimaging studies of semantic processing (Ni et al. 2000; Friederici et al. 2003; Newman et al. 2003; Humphries et al. 2006; 2007; Lau et al. 2008; Binder et al. 2009; Menenti et al. 2009; Binder and Desai 2011). The observed responses of the two inferior parietal sub-regions to the differing degree of message-level plausibility fits well with the following assumption that Binder et al. (2009, p. 2776) made on the basis of their extensive meta-analysis on semantic processing:

“Considering these various lines of evidence, we propose that the AG occupies a position at the top of a processing hierarchy underlying concept retrieval and conceptual integration. Though it is involved in all aspects of semantic processing, it may play a particular role in behaviors requiring fluent conceptual combination, such as sentence comprehension, discourse, problem solving, and planning.”

Notably, several of the clusters reported in this meta-analysis extended into the adjacent posterior portion of the SMG (which corresponds most closely to the areas PFG and PF in the monkey; (Pandya and Seltzer 1982), similar to the location the significant clusters in our whole-brain analysis. Due to its strategic location in the heteromodal parietal association cortex and its rich structural and functional connectivity, the inferior parietal lobule is considered a multimodal association area, with functional characterizations as “convergence zone” (Damasio, 1989) or “hub” for cross-modal integration (Joseph 1982; Hagmann et al. 2008; Tomasi and Volkow 2011) that plays a role in a number of core processes, such as categorization of events, access to semantic representations, fact retrieval and attention shifting (Seghier 2013). This perspective is in line with the processing demands imposed by anomalous sentences embedded in context and more specifically with hard to detect anomalies that call for a more extensive analysis and application of world knowledge and for which we observed two distinct clusters in the SMG at the whole-brain level.

Another explanation for the increased engagement of the supramarginal gyri in the processing of detected anomalies relative to control sentences lies in accounts on task- and attention-related processes (Shulman et al. 2003; Corbetta et al. 2008; Kristensen et al. 2013). Shulman and colleagues (2003) found sustained “deactivation” (negative deflection from baseline during task compared to rest) in the temporal-parietal region bilaterally during periods in which attention was strongly focused on the search for a behaviorally relevant target stimulus until a significant target had been detected. The authors argue that suppression of activity in the ventral attention network acts as a filter to protect the attention system from distractions caused by irrelevant stimuli and thus improves task performance. In the present task environment that emphasized the detection of anomalies, the attenuation of this negative deflection observed for semantic anomalies of both types could therefore indicate that detection of an anomaly has taken place, while the usual pattern for control sentences could point to a prolonged search for a relevant target.

Note that the suggested contributions of both inferior parietal lobules to attention-related or conceptual processes are not necessarily mutually exclusive as the anatomical heterogeneity and close proximity to several neural pathways of this region highlight the possibility of several functionally distinct sub-regions and their participation in a variety of different neural computations. However, more work will be required to disentangle and determine the precise contributions of the left and right SMG and AG to the implementation of these (and other) processes.

Anterior temporal lobe and posterior middle temporal gyrus—There is a growing body of evidence encompassing both functional neuroimaging and neuropsychological findings that consistently implicate the middle temporal gyrus – and more specifically its left posterior portion – and various parts of the bilateral anterior temporal lobe in processes of multimodal convergence and conceptual combinatorics (Martin 2007; Binder et al. 2009; Visser et al. 2010; 2012; Binder and Desai 2011; Noonan et al. 2013). Both regions are defined as higher-order association cortices based on their rich structural connectivity: The lateral anterior temporal lobe receives inputs from three different sensory systems (auditory, visual, olfactory) and is tightly connected to limbic and paralimbic regions (Olson et al. 2007). The posterior MTG lies at the confluence of several major white matter pathways

including direct and indirect segments of the arcuate fasciculus (AF), the inferior occipito-frontal fasciculus (IOFF), and the inferior and medial longitudinal fasciculus (ILF; MLF). The posterior MTG is connected to the angular gyrus via short fibers of the indirect segment of the AF and to anterior parts of the superior temporal lobes via the MLF (Buckner et al. 2009; Turken and Dronkers 2011). The functional relevance of these connections as part of a neural circuit underlying conceptual processing is supported by the results of recent studies examining functional connectivity during task and rest, as well as by detailed meta-analyses (Simmons et al. 2010; de Zubicaray et al. 2011; Visser et al. 2012; Noonan et al. 2013; Hurley et al. 2015). The fact that the discussed regions in the temporal and inferior parietal lobule appear to be functionally connected and show similar response profiles to semantic manipulations, however, does not entail that their specific contributions to the processing of semantic information are the same. Examining the differences in their structural and functional connections with other brain regions and systems can provide valuable information for conclusions about their respective roles within a common network. Anterior temporal regions, for example, have been linked to encoding emotional and social aspects of semantic memory based to their connections to the limbic system (Olson et al. 2007; Simmons et al. 2010). The posterior MTG, on the other hand, has been suggested to participate in mechanisms of controlled semantic retrieval in light of its frequent coactivation with the left IFG in tasks requiring semantic selection and control (Whitney, Jefferies, et al. 2011; Whitney, Kirk, et al. 2011; Dick et al. 2012; Noonan et al. 2013).

Most of these proposals are based on results from strictly controlled experimental designs of simple verbal and non-verbal stimuli such as single words, phrases or pictures. The present study provides additional and converging evidence to the emerging picture of a distributed network supporting conceptual processing by including results from high-level language comprehension. We used semantically complex stimuli that require transfer and convergence of information along different modalities and scales and found support for the involvement of cortical hubs in the anterior and middle temporal lobe. The challenge for future research lies not only in more precisely identifying the respective components of the network based on structural and functional connectivity during task and rest, but also in uncovering the basic mechanistic principles of information transfer and convergence, as well as their dynamic, contextual adaptations.

Task-related effects in the prefrontal cortex and the anterior insula

Our analyses revealed distinct levels of BOLD activation across stimulus types as well as for detected anomalies compared to control sentences in several prefrontal regions and in the anterior insula. We argue that the observed modulation of neural responses can be linked to the use of an overt plausibility judgment task and to distinct contributions of these brain regions in the implementation of goal-directed behavior.

The effects found in the ROI and whole-brain analysis for the anterior insula are particularly interesting as they show stronger responses to control than to anomalous scenarios of the easy type. We suggest that these findings reflect the varying degree of uncertainty associated with arriving at a final decision for the plausibility judgment (Bornkessel-Schlesewsky et al. 2009; Chang et al. 2013). Why might easy control sentences be more difficult to judge than

sentences with easy to detect anomalies? Recall that the different types of sentences were presented in pseudorandom order and required an overt decision. Sentences with easy to detect anomalies can be quickly and easily categorized as anomalous upon encountering the critical word, whereas sentences with a hard to detect anomaly or no anomaly at all both require more extensive and complete processing. This assumption is corroborated by the behavioral results: participants were most accurate in judging easy anomalies, while the accuracy rate was considerably lower for easy control stimuli and lowest for borderline stimuli. Note that only correctly judged control sentences were included in the analysis.

The role of (left) prefrontal areas, particularly the inferior frontal gyrus, in language processing is a topic of intense debate. At the core of this debate lies the question whether these regions are involved in language-specific operations or whether their contributions to language processing reflect domain general executive functions (Thompson-Schill et al. 1997; Stowe et al. 1998; Miller and Cohen 2001; Stowe et al. 2005; Thompson-Schill et al. 2005; Bornkessel-Schlesewsky and Schlewsky 2013). The results of the present study lend support to the latter perspective: The observed graded levels of hemodynamic activity across the four critical conditions that mirror the relative ease or difficulty of decision making in these cases (as evidenced by the behavioral results). Moreover, the results cannot be explained by accounts assuming a role of the prefrontal cortex in operations such as semantic unification (Hagoort 2005; 2014) since they would predict generally higher levels of activation for anomalous compared to non-anomalous sentences as unification would be more effortful for the former. As such, both prefrontal cortex and anterior insula appear to be engaged in goal-directed behavior as required by the experimental task. However, we argue that prefrontal regions are engaged in very general mechanisms of cognitive control, whereas the (anterior) insula serves as an interface between affective processing, higher order cognition and action.

5. Conclusion

We have shown here that scenario processing draws on a large set of brain areas distributed across the bilateral frontal, temporal and inferior parietal lobes. The present study provides novel evidence for the involvement of the supramarginal gyri of both hemispheres in the comprehension of linguistic stimuli at the level of more complex scenarios. Overall, our pattern of results supports the notion of a processing hierarchy, as the processing of easy anomalies preferentially engaged brain regions close to primary sensory regions involved in lower-level processes. The observed contributions of the anterior temporal and posterior middle temporal lobes and IPL in the processing of plausibility points to a distributed network engaged in ensuring the convergence of information along different dimensions and degrees of abstraction. This contrasts with the view that contextual integration occurs in a single specialized brain region. Lastly, we observed effects in the anterior insular cortex and inferior frontal gyrus bilaterally that seem to correlate with task-related demands of goal-directed behavior.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

When this research was performed, ST and IBS were at the University of Marburg. We would like to thank Sabine Frenzel for her help in the preparation of the stimulus materials, Fiona Weiß for assistance in the data acquisition, Matthew Schiel and Forrest Hippensteel for their support in data analysis, as well as Michael Andric, Anthony Dick, and Salomi Asaridou for helpful discussions. We also thank Jessica Heinen for conducting the sentence completion test.

Funding: This work was funded by the National Institutes of Health (grant DC-R01-3378 to SLS) and by a German Academic Exchange Service scholarship awarded to ST.

References

- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*. 1995; 57:289–300.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Rao SM, Cox RW. Conceptual Processing during the Conscious Resting State: A Functional MRI Study. *Journal of Cognitive Neuroscience*. 1999; 11:80–93. [PubMed: 9950716]
- Binder JR, Desai RH. The neurobiology of semantic memory. *Trends in Cognitive Sciences*. 2011; 15:527–536. [PubMed: 22001867]
- Binder JR, Desai RH, Graves WW, Conant LL. Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*. 2009; 19:2767–2796. [PubMed: 19329570]
- Bohan J, Leuthold H, Hijikata Y, Sanford AJ. The processing of good-fit semantic anomalies: An ERP investigation. *Neuropsychologia*. 2012; 50:3174–3184. [PubMed: 22975191]
- Bohan J, Sanford A. Semantic anomalies at the borderline of consciousness: An eye-tracking investigation. *The Quarterly Journal of Experimental Psychology*. 2008; 61:232–239. [PubMed: 17886160]
- Bornkessel-Schlesewsky I, Schlesewsky M. Reconciling time, space and function: A new dorsal-ventral stream model of sentence comprehension. *Brain and Language*. 2013; 125:60–76. [PubMed: 23454075]
- Bornkessel-Schlesewsky I, Schlesewsky M, Cramon von DY. Word order and Broca's region: Evidence for a supra-syntactic perspective. *Brain and Language*. 2009; 111:125–139. [PubMed: 19853290]
- Bornkessel-Schlesewsky I, Schlesewsky M, Small SL, Rauschecker JP. Neurobiological roots of language in primate audition: common computational properties. *Trends in Cognitive Sciences*. 2015; 19:142–150. [PubMed: 25600585]
- Buckner RL, Sepulcre J, Talukdar T, Krienen FM, Liu H, Hedden T, Andrews-Hanna JR, Sperling RA, Johnson KA. Cortical Hubs Revealed by Intrinsic Functional Connectivity: Mapping, Assessment of Stability, and Relation to Alzheimer's Disease. *Journal of Neuroscience*. 2009; 29:1860–1873. [PubMed: 19211893]
- Chang LJ, Yarkoni T, Khaw MW, Sanfey AG. Decoding the Role of the Insula in Human Cognition: Functional Parcellation and Large-Scale Reverse Inference. *Cerebral Cortex*. 2013; 23:739–749. [PubMed: 22437053]
- Chung MK, Robbins SM, Dalton KM, Davidson RJ, Alexander AL, Evans AC. Cortical thickness analysis in autism with heat kernel smoothing. *NeuroImage*. 2005; 25:1256–1265. [PubMed: 15850743]
- Corbetta M, Patel G, Shulman GL. The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron*. 2008; 58:306–324. [PubMed: 18466742]
- Cox RW. AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages. *Computers and Biomedical Research*. 1996; 29:162–173. [PubMed: 8812068]
- Cox RW. AFNI: What a long strange trip it's been. *NeuroImage*. 2012; 62:743–747. [PubMed: 21889996]
- Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage*. 1999; 9:179–194. [PubMed: 9931268]

- Damasio AR. Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*. 1989; 33:25–62. [PubMed: 2691184]
- Damasio H, Tranel D, Grabowski T, Adolphs R, Damasio A. Neural systems behind word and concept retrieval. *Cognition*. 2004; 92:179–229. [PubMed: 15037130]
- Daneman M, Lennertz T, Hannon B. Shallow semantic processing of text: Evidence from eye movements. *Language and Cognitive Processes*. 2007; 22:83–105.
- de Zubicaray GI, Rose SE, McMahon KL. The structure and connectivity of semantic memory in the healthy older adult brain. *NeuroImage*. 2011; 54:1488–1494. [PubMed: 20807579]
- Destrieux C, Fischl B, Dale A, Halgren E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*. 2010; 53:1–15. [PubMed: 20547229]
- Dick AS, Goldin-Meadow S, Hasson U, Skipper JI, Small SL. Co-speech gestures influence neural activity in brain regions associated with processing semantic information. *Hum Brain Mapp*. 2009; 30:3509–3526. [PubMed: 19384890]
- Dick AS, Mok EH, Beharelle AR, Goldin-Meadow S, Small SL. Frontal and temporal contributions to understanding the iconic co-speech gestures that accompany speech. *Hum Brain Mapp*. 2012; 35:900–917. [PubMed: 23238964]
- Dick AS, Solodkin A, Small SL. Neural development of networks for audiovisual speech comprehension. *Brain and Language*. 2010; 114:101–114. [PubMed: 19781755]
- Dikker S, Pyllkänen L. Before the N400: Effects of lexical–semantic violations in visual cortex. *Brain and Language*. 2011; 118:23–28. [PubMed: 21458057]
- Dikker S, Rabagliati H, Farmer TA, Pyllkänen L. Early Occipital Sensitivity to Syntactic Category Is Based on Form Typicality. *Psychological Science*. 2010; 21:629–634. [PubMed: 20483838]
- Dikker S, Rabagliati H, Pyllkänen L. Sensitivity to syntax in visual cortex. *Cognition*. 2009; 110:293–321. [PubMed: 19121826]
- Duvernoy, HM. *The Human Brain*. Springer; 1999.
- Erickson TD, Mattson ME. From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*. 1981; 20:540–551.
- Fischl B. Cortical Surface-Based Analysis II: Inflation, Flattening, and a Surface-Based Coordinate System. *NeuroImage*. 1999; 9:195–207. [PubMed: 9931269]
- Fischl B. Automatically Parcellating the Human Cerebral Cortex. *Cerebral Cortex*. 2004; 14:11–22. [PubMed: 14654453]
- Fischl B, Sereno MI, Tootell RB, Dale AM. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp*. 1999; 8:272–284. [PubMed: 10619420]
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC. Improved Assessment of Significant Activation in Functional Magnetic Resonance Imaging (fMRI): Use of a Cluster-Size Threshold. *Magn Reson Med*. 1995; 33:636–647. [PubMed: 7596267]
- Fox MD. From The Cover: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *PNAS*. 2005; 102:9673–9678. [PubMed: 15976020]
- Friederici AD. Pathways to language: fiber tracts in the human brain. *Trends in Cognitive Sciences*. 2009; 13:175–181. [PubMed: 19223226]
- Friederici AD. The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences*. 2012; 16:262–268. [PubMed: 22516238]
- Friederici AD, Rüschemeyer SA, Hahne A, Fiebach CJ. The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cerebral Cortex*. 2003; 13:170–177. [PubMed: 12507948]
- Friston K. Predictive coding, precision and synchrony. *Cognitive Neuroscience*. 2012; 3:238–239. [PubMed: 24171746]
- Friston K, Kiebel S. Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2009; 364:1211–1221.
- Genovese CR, Lazar NA, Nichols T. Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate. *NeuroImage*. 2002; 15:870–878. [PubMed: 11906227]
- Gloor, P. *The Temporal Lobe and Limbic System*. Oxford University Press; USA: 1997.

- Grewe T, Bornkessel I, Zysset S, Wiese R, von Cramon DY, Schlesewsky M. The emergence of the unmarked: A new perspective on the language-specific function of Broca's area. *Human Brain Mapping*. 2005; 26:178–190. [PubMed: 15929098]
- Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen Van J, Sporns O. Mapping the Structural Core of Human Cerebral Cortex. *PLoS Biol*. 2008; 6:1479–1493.
- Hagoort P. On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences*. 2005; 9:416–423. [PubMed: 16054419]
- Hagoort P. ScienceDirectNodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*. 2014; 28:136–141. [PubMed: 25062474]
- Heinen, J. Unpublished BA Thesis. University of Marburg; Germany: 2013. Kann eine Satzvervollständigungsaufgabe Aufschluss über die Generierung von kontextbasierten Prädiktionen geben? Eine empirische Studie zur Moses-Illusion. [= Can a sentence completion task provide insights into the generation of context-based predictions? An empirical study on the Moses Illusion.].
- Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci*. 2007; 8:393–402. [PubMed: 17431404]
- Hoffman P, Pobric G, Drakesmith M, Lambon Ralph MA. Posterior middle temporal gyrus is involved in verbal and non-verbal semantic cognition: Evidence from rTMS. *Aphasiology*. 2012; 26:1119–1130.
- Humphreys GF, Hoffman P, Visser M, Binney RJ, Lambon Ralph MA. Establishing task- and modality-dependent dissociations between the semantic and default mode networks. *PNAS*. 2015; 112:7857–7862. [PubMed: 26056304]
- Humphries C, Binder JR, Medler DA, Liebenthal E. Syntactic and Semantic Modulation of Neural Activity during Auditory Sentence Comprehension. *Journal of Cognitive Neuroscience*. 2006; 18:665–679. [PubMed: 16768368]
- Humphries C, Binder JR, Medler DA, Liebenthal E. Time course of semantic processes during sentence comprehension: An fMRI study. *NeuroImage*. 2007; 36:924–932. [PubMed: 17500009]
- Hurley RS, Bonakdarpour B, Wang X, Mesulam MM. Asymmetric Connectivity between the Anterior Temporal Lobe and the Language Network. *Journal of Cognitive Neuroscience*. 2015; 27:464–473. [PubMed: 25244113]
- Huynh H, Feldt LS. Conditions under which the mean-square ratios in repeated measurement designs have exact F-distributions. *Journal of the American Statistical Association*. 1970; 65:1582–1589.
- Indefrey P, Cutler A. Prelexical and lexical processing in listening. *The cognitive neurosciences*. 2004; 3:759–774.
- Jones EG, Powell T. An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain*. 1970
- Joseph R. The neuropsychology of development: Hemispheric laterality, limbic language, and the origin of thought. *J Clin Psychol*. 1982; 38:4–33. [PubMed: 7056873]
- Kristensen LB, Wang L, Petersson KM, Hagoort P. The Interface Between Language and Attention: Prosodic Focus Marking Recruits a General Attention Network in Spoken Language Comprehension. *Cerebral Cortex*. 2013; 23:1836–1848. [PubMed: 22763170]
- Lau EF, Phillips C, Poeppel D. A cortical network for semantics: (de)constructing the N400. *Nat Rev Neurosci*. 2008; 9:920–933. [PubMed: 19020511]
- Martin A. The representation of object concepts in the brain. *Annu Rev Psychol*. 2007; 58:25–45. [PubMed: 16968210]
- Mashal N, Solodkin A, Dick AS, Chen EE, Small SL. A Network Model of Observation and Imitation of Speech. *Front Psychology*. 2012; 3
- Matelli M, Camarda R, Glickstein M. Afferent and efferent projections of the inferior area 6 in the macaque monkey. *Journal of Comparative Neurology*. 1986; 251:281–298. [PubMed: 3021823]
- Mazoyer BM, Tzourio N, Frak V, Syrota A, Murayama N, Levrier O, Salamon G, Dehaene S, Cohen L, Mehler J. The cortical representation of speech. *Journal of Cognitive Neuroscience*. 1993; 5:467–479. [PubMed: 23964919]

- Menenti L, Petersson KM, Scheeringa R, Hagoort P. When elephants fly: differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. *Journal of Cognitive Neuroscience*. 2009; 21:2358–2368. [PubMed: 19016600]
- Miller EK, Cohen JD. An Integrative Theory of Prefrontal Function. *Annu Rev Neurosci*. 2001; 24:167–202. [PubMed: 11283309]
- Newman AJ, Pancheva R, Ozawa K, Neville HJ, Ullman MT. An Event-Related fMRI Study of Syntactic and Semantic Violations. *J Psycholinguist Res*. 2001; 30:339–364. [PubMed: 11523278]
- Newman SD, Just MA, Keller TA, Roth J, Carpenter PA. Differential effects of syntactic and semantic processing on the subregions of Broca's area. *Cognitive Brain Research*. 2003; 16:297–307. [PubMed: 12668239]
- Ni W, Constable RT, Mencl WE, Pugh KR, Fulbright RK, Shaywitz SE, Shaywitz BA, Gore JC, Shankweiler D. An Event-related Neuroimaging Study Distinguishing Form and Content in Sentence Processing. *Journal of Cognitive Neuroscience*. 2000; 12:120–133. [PubMed: 10769310]
- Noonan KA, Jefferies E, Visser M, Lambon Ralph MA. Going beyond Inferior Prefrontal Involvement in Semantic Control: Evidence for the Additional Contribution of Dorsal Angular Gyrus and Posterior Middle Temporal Cortex. *Journal of Cognitive Neuroscience*. 2013; 25:1824–1850. [PubMed: 23859646]
- Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*. 1971; 9:97–113. [PubMed: 5146491]
- Olson IR, Plotzker A, Ezzyat Y. The Enigmatic temporal pole: a review of findings on social and emotional processing. *Brain*. 2007; 130:1718–1731. [PubMed: 17392317]
- Pandya DN, Seltzer B. Association areas of the cerebral cortex. *Trends in Neurosciences*. 1982; 5:386–390.
- Patterson K, Nestor PJ, Rogers TT. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci*. 2007; 8:976–987. [PubMed: 18026167]
- Petrides M, Pandya DN. Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. *Journal of Comparative Neurology*. 1984; 228:105–116. [PubMed: 6480903]
- Price CJ. The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*. 2010; 1191:62–88. [PubMed: 20392276]
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2014. URL <http://www.R-project.org/>
- Raposo A, Marques JF. The contribution of fronto-parietal regions to sentence comprehension: Insights from the Moses illusion. *NeuroImage*. 2013; 83:431–437. [PubMed: 23796543]
- Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci*. 2009; 12:718–724. [PubMed: 19471271]
- Rizzolatti G, Matelli M. Two different streams form the dorsal visual system: anatomy and functions. *Exp Brain Res*. 2003; 153:146–157. [PubMed: 14610633]
- Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat Neurosci*. 1999; 2:1131–1136. [PubMed: 10570492]
- Saad ZS, Reynolds RC. SUMA. *NeuroImage*. 2012; 62:768–773. [PubMed: 21945692]
- Sanford AJ, Leuthold H, Bohan J, Sanford AJ. Anomalies at the borderline of awareness: An ERP study. *Journal of Cognitive Neuroscience*. 2011; 23:514–523. [PubMed: 19925201]
- Saur D, Kreher BW, Schnell S, Kummerer D, Kellmeyer P, Vry MS, Umarova R, Musso M, Glauche V, Abel S, Huber W, Rijntjes M, Hennig J, Weiller C. Ventral and dorsal pathways for language. *PNAS*. 2008; 105:18035–18040. [PubMed: 19004769]
- Schmahmann JD, Pandya DN, Wang R, Dai G, D'Arceuil HE, de Crespigny AJ, Wedeen VJ. Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography. *Brain*. 2007; 130:630–653. [PubMed: 17293361]
- Seghier ML. The Angular Gyrus Multiple Functions and Multiple Subdivisions. *The Neuroscientist*. 2013; 19:43–61. [PubMed: 22547530]

- Seltzer B, Pandya DN. Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Research*. 1978; 149:1–24. [PubMed: 418850]
- Seltzer B, Pandya DN. Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *J Comp Neurol*. 1989; 281:97–113. [PubMed: 2925903]
- Seltzer B, Pandya DN. Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: A retrograde tracer study. *J Comp Neurol*. 1994; 343:445–463. [PubMed: 8027452]
- Shulman GL, McAvoy MP, Cowan MC, Astafiev SV, Tansy AP, d'Avossa G, Corbetta M. Quantitative analysis of attention and detection signals during visual search. *Journal of Neurophysiology*. 2003; 90:3384–3397. [PubMed: 12917383]
- Simmons WK, Reddish M, Bellgowan PSF, Martin A. The Selectivity and Functional Connectivity of the Anterior Temporal Lobes. *Cerebral Cortex*. 2010; 20:813–825. [PubMed: 19620621]
- Stowe LA, Broere CA, Paans AM, Wijers AA, Mulder G, Vaalburg W, Zwarts F. Localizing components of a complex task: sentence processing and working memory. *NeuroReport*. 1998; 9:2995–2999. [PubMed: 9804304]
- Stowe LA, Haverkort M, Zwarts F. Rethinking the neurological basis of language. *Lingua*. 2005; 115:997–1042.
- Straube B, Green A, Weis S, Chatterjee A, Kircher T. Memory Effects of Speech and Gesture Binding: Cortical and Hippocampal Activation in Relation to Subsequent Memory Performance. *Journal of Cognitive Neuroscience*. 2009; 21:821–836. [PubMed: 18578601]
- Thompson-Schill SL, Bedny M, Goldberg RF. The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology*. 2005; 15:219–224. [PubMed: 15831406]
- Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences*. 1997; 94:14792–14797.
- Tomasi D, Volkow ND. Association between Functional Connectivity Hubs and Brain Networks. *Cerebral Cortex*. 2011; 21:2003–2013. [PubMed: 21282318]
- Tune S, Schlesewsky M, Small SL, Sanford AJ, Bohan J, Sassenhagen J, Bornkessel-Schlesewsky I. Cross-linguistic variation in the neurophysiological response to semantic processing: Evidence from anomalies at the borderline of awareness. *Neuropsychologia*. 2014; 56:147–166. [PubMed: 24447768]
- Turken U, Dronkers NF. The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Frontiers in systems neuroscience*. 2011
- Van Petten C, Luka BJ. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language*. 2006; 97:279–293. [PubMed: 16343606]
- Vandenberghe R, Nobre AC, Price CJ. The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience*. 2002; 14:550–560. [PubMed: 12126497]
- Vigneau M, Beaucousin V, Herve PY, Duffau H, Crivello F, Houde O, Mazoyer B, Tzourio-Mazoyer N. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage*. 2006; 30:1414–1432. [PubMed: 16413796]
- Visser M, Jefferies E, Embleton KV, Lambon Ralph MA. Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. *Journal of Cognitive Neuroscience*. 2012; 24:1766–1778. [PubMed: 22621260]
- Visser M, Jefferies E, Lambon Ralph MA. Semantic Processing in the Anterior Temporal Lobes: A Meta-analysis of the Functional Neuroimaging Literature. *Journal of Cognitive Neuroscience*. 2010; 22:1083–1094. [PubMed: 19583477]
- Whitney C, Jefferies E, Kircher T. Heterogeneity of the Left Temporal Lobe in Semantic Representation and Control: Priming Multiple versus Single Meanings of Ambiguous Words. *Cerebral Cortex*. 2011; 21:831–844. [PubMed: 20732899]
- Whitney C, Kirk M, O'Sullivan J, Ralph MAL, Jefferies E. The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cerebral Cortex*. 2011; 21:1066–1075. [PubMed: 20851853]

Wilson SM, Molnar-Szakacs I, Iacoboni M. Beyond Superior Temporal Cortex: Intersubject Correlations in Narrative Speech Comprehension. *Cerebral Cortex*. 2008

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

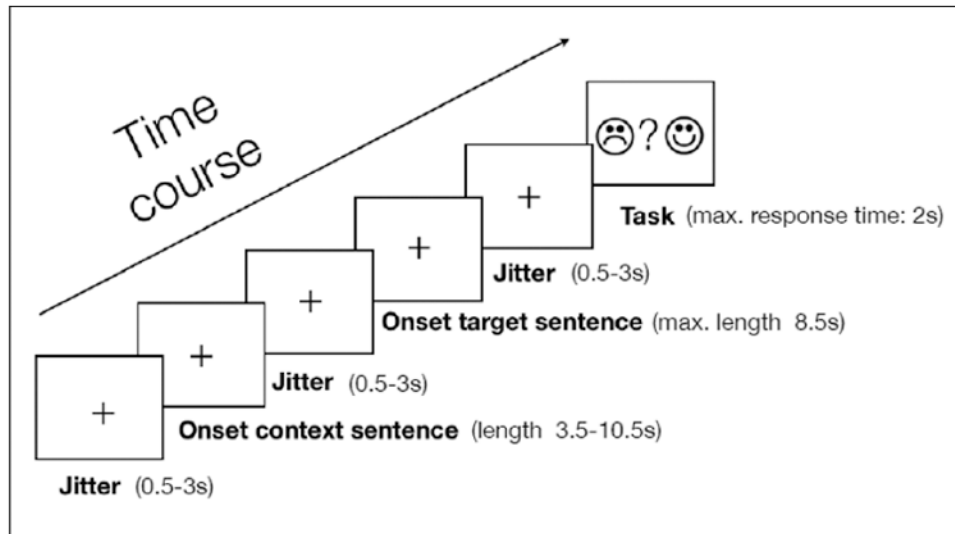


Figure 1.
Schematic depiction of the fMRI paradigm for a single trial.

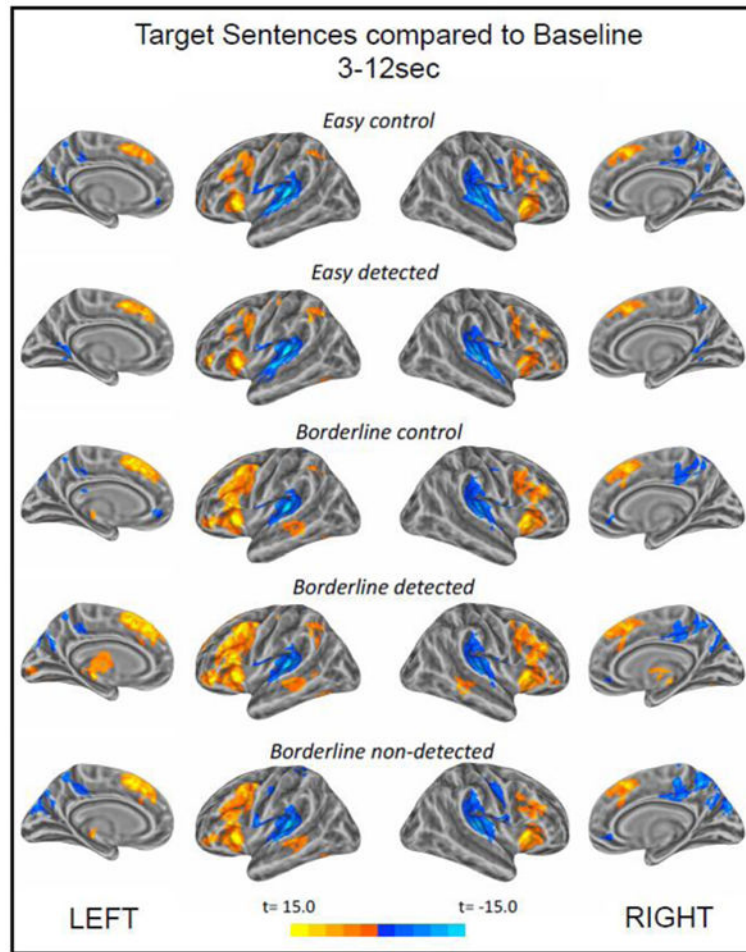


Figure 2. Whole-brain analysis results for each condition compared to baseline in the time window of 3-12 sec post target sentence onset. The individual per-vertex threshold was $p < .001$ (corrected FWE $p < .05$).

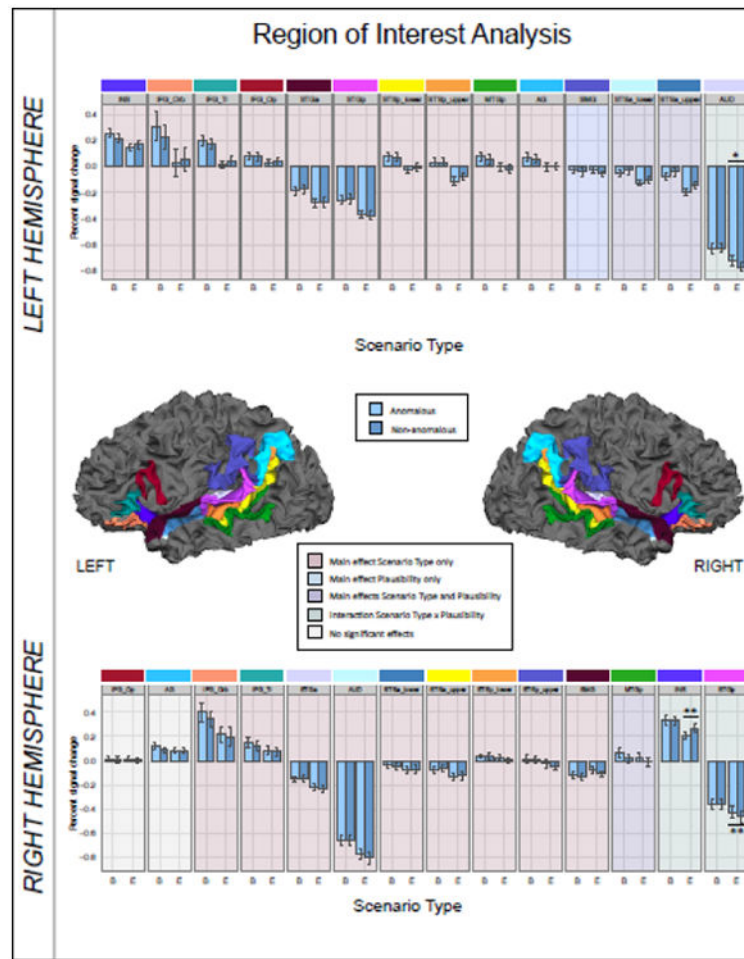


Figure 3. Results of the region of interest analysis including 14 regions per hemisphere. Regions are outlined on the white matter surface model of one representative subject. Error bars depict standard error of the mean. B = borderline scenarios. E = easy scenarios. * $p < 0.05$ (corrected); ** $p < 0.01$ (corrected).

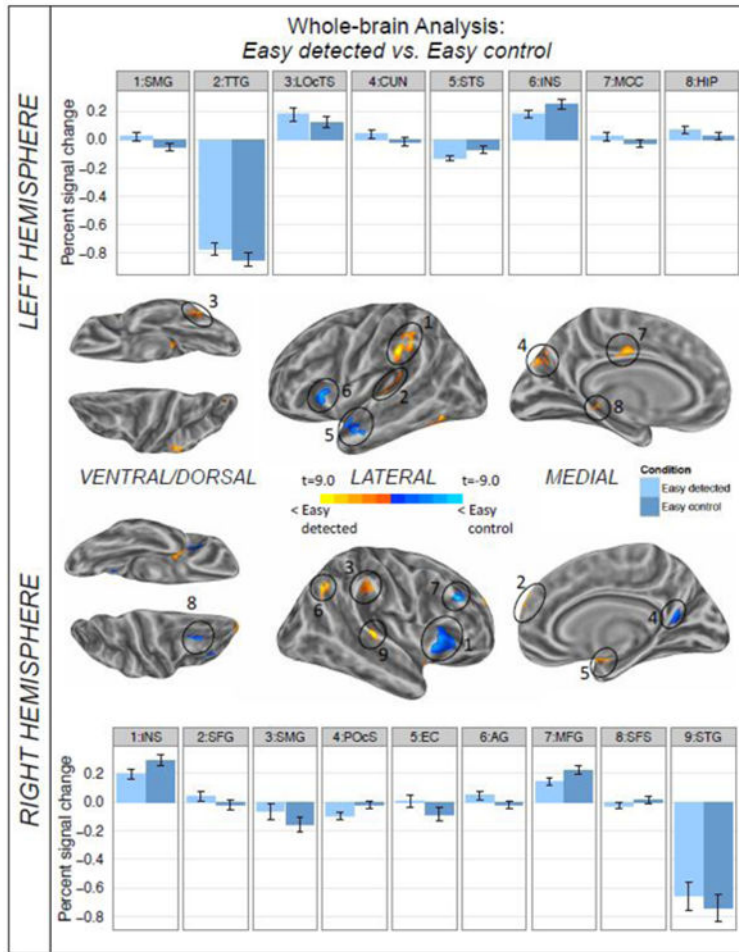


Figure 4. Results of the whole-brain comparison between the *easy detected* and *easy control* condition for the time window of 3-12 sec post sentence onset shown on inflated average brain surfaces of the left and right hemisphere. Labels refer to the local activation maximum of each cluster. The individual per-vertex threshold was $p < .01$ (corrected FWE $p < .05$). Warm colors indicate higher levels of activation for *easy detected*; cold colors reflect greater activation for the *easy control* sentences. Bar graphs present the difference between conditions for the local activation maximum of each cluster. Error bars depict standard error of the mean. AG=angular gyrus, CUN=cuneus, EC=entorhinal cortex, HIP= hippocampus, INS= anterior insula, LOcTS=lateral occipito-temporal sulcus, MCC=middle cingulate cortex, MFG=middle frontal gyrus, POcS=posterior occipital sulcus, SFS=superior frontal sulcus, SFG=superior frontal gyrus, SMG=supramarginal gyrus, STS=superior temporal sulcus, STG=superior temporal gyrus, TTG=transverse temporal gyrus.

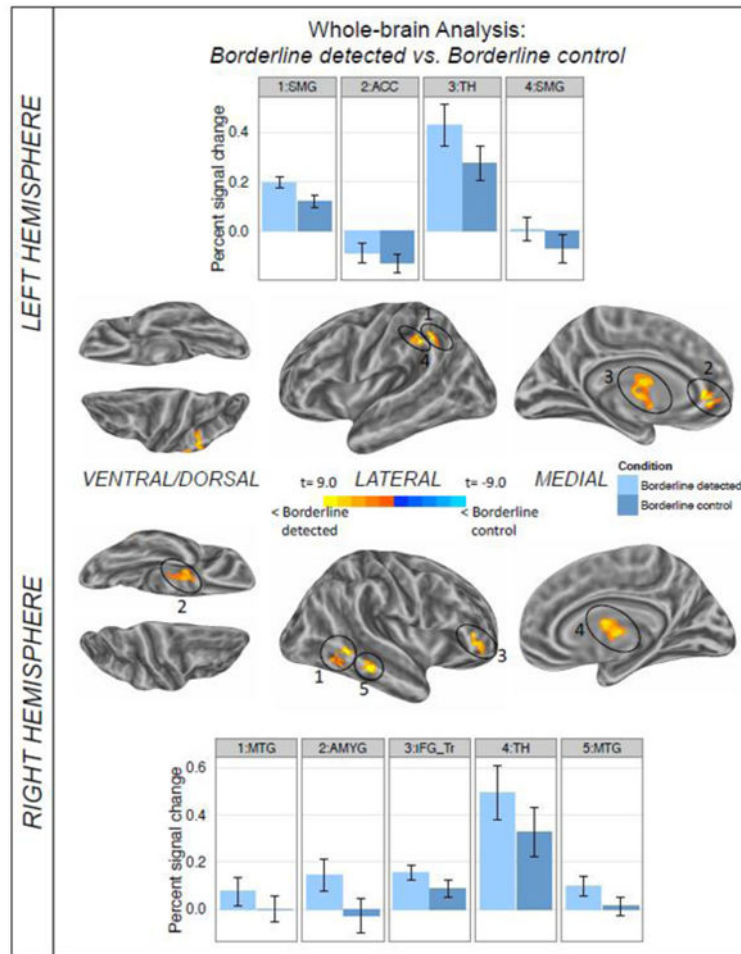


Figure 5.

Results of the whole-brain comparison between the *borderline detected* and *borderline control* condition for the time window of 3-12 sec post sentence onset shown on inflated average brain surfaces of the left and right hemisphere. Labels refer to the local activation maximum of each cluster. The individual per-vertex threshold was $p < .01$ (corrected FWE $p < .05$). Warm colors indicate higher levels of activation for *borderline detected*; cold colors reflect greater activation for the *borderline control* sentences. Bar graphs present the difference between conditions for the local activation maximum of each cluster. Error bars depict standard error of the mean. ACC=anterior cingulate cortex, AMYG=amygdala, IFG_Tr=pars triangularis of the inferior frontal gyrus, MTG=middle temporal gyrus, SMG=supramarginal gyrus, TH=thalamus.

Table 1

Example items for each of the four experimental conditions. For each sentence, a word-by-word translation is given along with the corresponding English example from Sanford et al. (2011). Each trial consists of a context sentence followed by an anomalous or control target sentence. Target words are highlighted in bold print, contextual manipulations are given in italics.

Context Sentence	Target Sentence
<p>Ein amerikanischer Jumbo Jet wurde von bewaffneten Extremisten gezwungen, An American jumbo jet was by armed extremists forced in Kanada zu landen und Experten waren schnell vor Ort, um zu vermitteln. in Canada to land and experts were quickly on site for to mediate. A North American jumbo jet was forced at gunpoint to land in Canada, and experts were quickly on hand to help.</p>	<p>Borderline anomalous</p> <p>Die Lage beruhigte sich durch die <i>Verhandlungen</i> der Behörden mit den verängstigten Geiseln. The situation calmed itself by the <i>negotiations</i> of-the authorities with the scared hostages die im Flugzeug saßen. who in-the plane sat. First of all the authorities' initial <i>negotiations</i> with the scared and desperate hostages helped calm the situation.</p> <p>Borderline control</p> <p>Die Lage beruhigte sich durch die <i>Kommunikation</i> der Behörden mit den verängstigten Geiseln. The situation calmed itself by the <i>communications</i> of-the authorities with the scared hostages die im Flugzeug saßen. who in-the plane sat. First of all the authorities' initial <i>communications</i> with the scared and desperate hostages helped calm the situation.</p> <p>Easy anomalous</p> <p>Zuerst servierte ihnen der Maier die falschen Gerichte und dann wurde ihnen auch noch zu viel berechnet. First served them the painter the wrong meals and then was them also still too much charged. They were given the wrong meals by the painter and then they were overcharged for their meals.</p> <p>Easy control</p> <p>Zuerst servierte ihnen der Kellner die falschen Gerichte und dann wurde ihnen auch noch zu viel berechnet. First served them the waiter the wrong meals and then was them also still too much charged. They were given the wrong meals by the waiter and then they were overcharged for their meals.</p>
<p>Christian und Julias gemeinsamer Abend in dem neuen Restaurant war ein Christian and Julia's joint evening in the new restaurant was a richtiger Reinfall. Denise and Fred's date to the new restaurant was a complete disaster.</p>	<p>Easy anomalous</p> <p>Zuerst servierte ihnen der Maier die falschen Gerichte und dann wurde ihnen auch noch zu viel berechnet. First served them the painter the wrong meals and then was them also still too much charged. They were given the wrong meals by the painter and then they were overcharged for their meals.</p> <p>Easy control</p> <p>Zuerst servierte ihnen der Kellner die falschen Gerichte und dann wurde ihnen auch noch zu viel berechnet. First served them the waiter the wrong meals and then was them also still too much charged. They were given the wrong meals by the waiter and then they were overcharged for their meals.</p>

Table 2

Anatomical description of the cortical regions of interest.

ROI	Anatomical structure	Delimiting Landmarks
IFG_Op	Pars opercularis of the inferior frontal gyrus	A = Anterior vertical ramus of the lateral fissure P = Inferior precentral sulcus S = Inferior frontal sulcus I = Anterior horizontal ramus of the lateral fissure to the border with insular cortex
IFG_Tr	Pars triangularis of the inferior frontal gyrus	A = Coronal plane defined as the rostral end of the anterior horizontal ramus of the sylvian fissure P = Vertical ramus of the lateral fissure S = Inferior frontal sulcus I = Anterior horizontal ramus of the lateral fissure
IFG_Orb	Pars orbitalis of the inferior frontal gyrus	A = Posterior lateral orbital sulcus P = Vertical ramus of the lateral fissure S = Anterior horizontal ramus of the lateral fissure I = Orbital gyrus
INSA	Anterior segment of the circular sulcus of the insula	A = Lateral orbital gyrus P = Short insular gyri S = Superior segment of the circular sulcus of the insula I = Inferior segment of the circular sulcus of the insula
STGa	Anterior portion of the superior temporal gyrus and planum polare	A = Inferior circular sulcus of the insula P = Vertical plane drawn from the anterior extent of the transverse temporal gyrus S = Anterior horizontal ramus of the lateral fissure I = Dorsal aspect of the upper bank of the superior temporal sulcus
AUD	Transverse temporal gyrus and sulcus (Heschl's gyrus and sulcus)	A = Planum polare and inferior circular sulcus of the insula P = Planum temporale M = Lateral fissure L = Superior temporal gyrus
STGp	Posterior portion of the superior temporal gyrus and planum temporale	A = Vertical plane drawn from the anterior extent of the transverse temporal gyrus P = Angular gyrus S = Supramarginal gyrus I = Dorsal aspect of the upper bank of the superior temporal sulcus
STSa_upper	Anterior portion of the upper bank of the superior temporal sulcus	A = Superior and middle temporal gyrus P = Vertical plane drawn from the anterior extent of the transverse temporal gyrus S = Superior temporal gyrus I = Fundus of the sulcus
STSa_lower	Anterior portion of the lower bank of the superior temporal sulcus	A = Superior and middle temporal gyrus P = Vertical plane drawn from the anterior extent of the transverse temporal gyrus S = Fundus of the sulcus I = Middle temporal gyrus
STSp_upper	Posterior portion of the upper bank of the superior temporal sulcus	A = Vertical plane drawn from the anterior extent of the transverse temporal gyrus P = Angular gyrus and middle occipital gyrus and sulcus S = Angular gyrus I = Fundus of the sulcus
STSp_lower	Posterior portion of the lower bank of the superior temporal sulcus	A = Vertical plane drawn from anterior extent of the transverse temporal gyrus P = Angular gyrus and middle occipital gyrus and sulcus S = Fundus of the sulcus I = Middle temporal gyrus
MTGp	Posterior portion of the middle temporal gyrus	A = Vertical plane drawn from anterior extent of the transverse temporal gyrus P = Anterior occipital sulcus and temporo-occipital incisure S = Superior temporal sulcus I = Inferior temporal sulcus
SMG	Supramarginal gyrus	A = Postcentral sulcus P = Sulcus intermedius primus of Jensen S = Intraparietal sulcus I = Lateral fissure
AG	Angular gyrus	A = Sulcus intermedius primus of Jensen P = Dorsal part of the anterior occipital sulcus S = Intraparietal sulcus I = Superior temporal sulcus and lateral occipital sulcus

A = anterior limit. P = posterior limit. S = superior limit. I = inferior limit. M = medial limit. L = lateral limit.

Table 3

Results of the repeated-measures ANOVA per region of interest.

Region of Interest	Left Hemisphere	Right Hemisphere
INSa	Type: $F(1,17) = 84.8, p < 0.001$	Plausibility \times Type, $F(1,17)=13.7, p = 0.007$ Easy: Type, $F(1,17)=9.1, p = 0.026$
IFG_Orb	Type: $F(1,17) = 88.9, p < 0.001$	Type: $F(1,17) = 33.2, p < 0.001$
IFG_Tr	Type: $F(1,17) = 111.3, p < 0.001$	Type: $F(1,17) = 10.8, p = 0.016$
IFG_Op	Type: $F(1,17) = 35.3, p < 0.001$	----
STGa	Type: $F(1,17) = 97.1, p < 0.001$	Type: $F(1,17) = 77.5, p < 0.001$
STGp	Type: $F(1,17) = 73.7, p < 0.001$	Plausibility \times Type, $F(1,17)=7.2, p = 0.04$ Easy: Type, $F(1,17)=12.0, p = 0.009$
AUD	Plausibility \times Type, $F(1,17) = 7.6, p = 0.037$ Easy: Type, $F(1,17) = 9.8, p = 0.02$	Type: $F(1,17) = 139.2, p < 0.001$
STSa_lower	Type: $F(1,17) = 51.0, p < 0.001$ Plaus: $F(1,17) = 7.9, p = 0.034$	Type: $F(1,17) = 12.3, p = 0.01$
STSa_upper	Type: $F(1,17) = 77.4, p < 0.001$ Plaus: $F(1,17) = 8.8, p = 0.028$	Type: $F(1,17) = 48.6, p < 0.001$
STSp_lower	Type: $F(1,17) = 68.5, p < 0.001$	Type: $F(1,17) = 8.0, p = 0.034$
STSp_upper	Type: $F(1,17)=150.5, p < 0.001$	Type: $F(1,17) = 18.9, p = 0.002$
MTGp	Type: $F(1,17) = 27.6, p < 0.001$	Type: $F(1,17) = 8.4, p = 0.03$ Plaus: $F(1,17) = 12.8, p = 0.009$
SMG	Plaus: $F(1,17) = 7.9, p = 0.034$	Type: $F(1,17) = 18.0, p = 0.002$
AG	Type: $F(1,17) = 29.8, p < 0.001$	----

Only statistically significant effects ($p < 0.05$) are reported. Listed p-values are FDR-corrected ($q = 0.05$). All two-way interactions were resolved by Scenario Type.

Table 4

Regions showing reliable differences for the whole-brain contrasts of *easy detected* versus *easy control* and *borderline detected* versus *borderline control*, respectively.

Hemi	Local Maximum/Cluster Extent	MNI Coordinates			Cluster Size		
		X	Y	Z	Nodes	Area	MI
Easy detected > Easy control							
LEFT	Supramarginal gyrus	-61	-32	26	956	257.2	0.12
	Transverse temporal gyrus and sulcus	-50	-15	2	700	212.3	0.10
	Lateral occipito-temporal sulcus extending to the inferior temporal gyrus	-47	-43	-18	387	196.8	0.11
	Cuneus extending into the parieto-occipital sulcus	-7	-82	31	401	166.3	0.10
	Middle portion of the cingulate cortex	-4	-16	33	415	117.4	0.08
	Hippocampus	-22	-37	1	356	67.9	0.09
RIGHT	Superior frontal gyrus	17	54	23	523	196.7	0.08
	Supramarginal gyrus	58	-29	35	318	145.8	0.12
	Entorhinal cortex	33	5	-16	799	127.3	0.17
	Angular gyrus	47	-58	38	384	101.4	0.11
	Superior frontal sulcus	27	15	43	329	82.7	0.06
	Superior temporal gyrus	65	-22	1	304	72.4	0.16
Easy control > Easy detected							
LEFT	Superior temporal sulcus extending dorsally into the superior temporal sulcus and ventrally into the middle temporal gyrus	-53	1	-17	303	164.5	0.12
	Insula	-28	23	0	335	120.2	0.08
RIGHT	Insula	31	28	-5	1006	373.1	0.16
	Parieto-occipital sulcus extending into the posterior-dorsal part of the cingulate gyrus	11	-54	13	349	130.3	0.08
	Middle frontal gyrus	47	26	30	272	87.6	0.12
	Superior frontal sulcus	27	15	43	329	82.7	0.06
Borderline detected > Borderline control							
LEFT	Supramarginal gyrus extending dorsally into the intraparietal sulcus	-49	-45	45	645	300.0	0.14
	Anterior portion of the cingulate cortex	-7	40	4	356	161.4	0.08

Hemi	Local Maximum/Cluster Extent	MNI Coordinates			Cluster Size		
		X	Y	Z	Nodes	Area	MI
	Thalamus	-2	-9	12	679	157.8	0.28
	Supramarginal gyrus	-59	-34	39	359	82.4	0.10
RIGHT	Middle temporal gyrus extending ventrally into the inferior temporal sulcus and gyrus	63	-47	-2	395	173.8	0.14
	Amygdala	18	-4	-14	592	148.2	0.30
	Pars triangularis of the IFG extending rostrally into the lateral orbital sulcus	51	33	-1	436	144.1	0.13
	Thalamus	2	-7	9	497	119.1	0.28
	Middle temporal gyrus	63	-32	-14	266	103.8	0.12