frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 18 January 2012
doi: 10.3389/fpsyg.2011.00384

From reference to sense: how the brain encodes meaning for speaking

Laura Menenti [1,2] ***, Karl Magnus Petersson*** [1,3] ***and Peter Hagoort*** [1,3] ***

[1] *Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands*
[2] *Institute for Neuroscience and Psychology, University of Glasgow, Glasgow, UK*
[3] *Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands*

Edited by:
Andriy Myachykov, University of Glasgow, UK

Reviewed by:
Ken McRae, University of Western Ontario, Canada
Mante Nieuwland, Basque Center on Cognition, Brain and Language, Spain

***Correspondence:**
Laura Menenti, Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, Netherlands.
e-mail: laura.menenti@mpi.nl;
Peter Hagoort, Donders Institute for Brain, Cognition and Behaviour, Donders Centre for Cognitive Neuroimaging, Radboud University Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, Netherlands.
e-mail: peter.hagoort@donders.ru.nl

In speaking, semantic encoding is the conversion of a non-verbal mental representation (the reference) into a semantic structure suitable for expression (the sense). In this fMRI study on sentence production we investigate how the speaking brain accomplishes this transition from non-verbal to verbal representations. In an overt picture description task, we manipulated repetition of sense (the semantic structure of the sentence) and reference (the described situation) separately. By investigating brain areas showing response adaptation to repetition of each of these sentence properties, we disentangle the neuronal infrastructure for these two components of semantic encoding. We also performed a control experiment with the same stimuli and design but without any linguistic task to identify areas involved in perception of the stimuli *per se*. The bilateral inferior parietal lobes were selectively sensitive to repetition of reference, while left inferior frontal gyrus showed selective suppression to repetition of sense. Strikingly, a widespread network of areas associated with language processing (left middle frontal gyrus, bilateral superior parietal lobes and bilateral posterior temporal gyri) all showed repetition suppression to both sense and reference processing. These areas are probably involved in mapping reference onto sense, the crucial step in semantic encoding. These results enable us to track the transition from non-verbal to verbal representations in our brains.

Keywords: semantics, conceptual representation, language production, fMRI, fMRI adaptation

INTRODUCTION

Look at that guy hitting the other guy! After reading this sentence, you presumably have a mental representation of two adult male persons, of whom one is hitting the other. They are both male and adult but they are still two different persons. A linguistic distinction within the domain of semantics, is the difference between *reference* and *sense* of a linguistic expression (Frege, 1892). The sense of an expression is its linguistic meaning, the reference is the entity the expression refers to. In the representation of the sentence *Look at that guy hitting the other guy!* there are two *guys* (for instance a blond and a dark-haired guy, as indicated by "that" and "the other"), but they are both referred to by the same sense, the word g*uy* (an adult male person). This sense thus has two possible references. The reverse is also possible. If you knew more about the two *guys* you might be shouting: *Look at that man hitting his son!* in the same situation. *His son* and *the other guy* are then two possible senses which can have the same referent.

In the view of Jackendoff (2002), which we adopt in the current paper, referents are representations in our minds. For concrete objects they are representations in our minds, of objects in the real world, constructed by the perceptual system. These representations are considered concepts, which thus are non-linguistic in nature. Sense, then, is that part of meaning that is encoded in the form of the utterance. In other words, sense (linguistic meaning) is the interface between the conceptual system and linguistic form (spanning both phonology and syntax; Jackendoff, 2002). Speaking is the conversion of an intention to communicate a message into a linearized string of speech sounds. An essential step in this process is semantic encoding – the retrieval of the relevant concepts and the specification of semantic structure (Levelt, 1989). In this step, the intended reference needs to be mapped onto a sense, for it to be expressed. In this mapping process, certain semantic choices have to be made, such as referring to the entities in the referential domain by, for instance, "the guy" or "the man on the chair." From a processing point of view, then, reference forms the input to semantic encoding, while sense is the output. Semantic encoding itself is the computation necessary to map reference (the input) onto the sense (the output) in order to generate the appropriate output. In this paper, we consider sense to be equivalent to the *preverbal message* in sentence production (Levelt, 1989). The preverbal message is the semantic structure that forms the output of semantic encoding and the input to syntactic and phonological encoding.

In speaking establishing reference is the first step of semantic encoding, necessary to utter a sentence in the first place. As few neuroimaging studies investigating semantic encoding in sentence production have so far been conducted, in this fMRI study we aim to fill that gap. Picture naming paradigms have previously been used in fMRI albeit in single word studies. Retrieving a name for a picture has been shown to involve more activity in bilateral temporal areas, the left frontal lobe, bilateral occipital areas, bilateral parietal areas, and the anterior cingulate

(Kan and Thompson-Schill, 2004) than does making visual decisions about abstract pictures. A similar set of areas has been shown to increase activity in picture naming and reading aloud compared to counting (Parker Jones et al., 2011). These data suggest that a large network of areas may be involved in semantic encoding. Also, while part of this network are well-established language areas, some are not. Perhaps, then, these are areas encoding the reference for these materials.

Moving on to sentence production, in a previous fMRI adaptation study on sentence-level processing, we compared the neuronal structure underlying computation of semantic structure of an utterance in comprehension and production (Menenti et al., 2011). More specifically, we investigated the construction of *thematic role structure*, the relation between the different concepts and events, or "who does what to whom." This aspect of semantic structure forms a crucial interface between conceptual structure (the domain of reference) and syntactic structure (the grammatical roles). Schematically a thematic role structure can be stated as a predicate with arguments: ROB(THIEF, LADY(OLD)). There is a "ROB" event, performed by a THIEF (the agent of the action) to the expenses of a LADY (the patient of the action), who has the property of being OLD. In our study, photographs depicting transitive events (events requiring an agent and a patient, such as ROB, KISS, HIT) provided the context for the sentences, which were either produced or heard by the participants. We found bilateral posterior middle temporal gyri involved in this component of sentence processing.

While this study provided valuable insights on the neuronal infrastructure underlying different steps in sentence production and comprehension, the semantic encoding manipulation disregarded the distinction between reference and sense. The next question, and the one underlying the present study, then, is how the different areas involved in semantic encoding play a role in processing the input (reference), the output (sense), and the process of mapping the one onto the other. In this sentence production fMRI adaptation study we again focused on thematic role structure as an essential part of semantic structure. In a picture description paradigm, we manipulated repetition of semantic structure across subsequent sentences, crossing repetition of reference and sense.

Our paradigm involved pictures of transitive events being enacted by two actors. We operationalized sense as the literal sentence used to describe the picture. Reference we considered the sum of the action involved, the roles of actors as agents and patients, and the exact spatial configuration of agents and patients.

In our task, the actors in the picture were colored and these colors varied for the same depicted situation. Participants could therefore subsequently describe the same situation as "The yellow man hits the blue woman." and then as "The red man hits the green woman." Although the picture therefore looked slightly different in the two trials, the colors were an arbitrarily varying property of the individuals in the picture and the participants were made aware of this (see Materials and Methods). We do not consider such arbitrary variations to be part of reference. One might consider this parallel to the fact that we change clothes every day: they make us look different but do not thereby cause us to become different individuals. The reference of the expression was therefore kept constant but the sense changed. In a complementary fashion, the sentence "The red man hits the green woman." could be used in subsequent trials to describe a different hitting event involving different participants. Sense was kept constant, but the reference changed. This allowed us to distinguish the situation the participants spoke about from the utterance they used to speak about it.

As can be seen in **Figure 1**, this means that our relevant "novel reference" condition still has considerable overlap with the prime. We chose this approach to eliminate any potential confounds. For instance, repeating the actors between prime and "repeated reference" target but not between prime and "novel reference" target would leave effects open to, for instance, the alternative interpretation that we are looking at face repetition effects. By choosing the most narrow comparison possible, we can be more sure of the interpretation of the results, while admittedly running the risk of missing some other potentially relevant effects.

To further investigate the distinction between non-verbal and verbal processing of meaning, we performed a control experiment. In this experiment we showed participants the exact same stimulus sequences, but this time paired with a non-linguistic task. Any brain areas involved in processing only the non-linguistic, conceptual representations involved in interpreting the pictures (i.e., the reference), should also show an adaptation effect without a linguistic task. On the other hand, brain areas involved in converting meaning into language (the sense), should not show adaptation effects in such a setting.

Our hypothesis was that areas involved in processing the conceptual input to semantic encoding should show adaptation effects for repetition of reference in both the speaking and control experiments, while not showing sensitivity to repetition of sense. Areas involved in semantic encoding itself, that is, in mapping reference onto sense, should show adaptation to repetition of both reference and sense. Finally, areas involved in processing the output of semantic encoding, the sense, should show adaptation to repetition of sense in the speaking experiment, and should not show sensitivity to repetition of reference.

MATERIALS AND METHODS

PARTICIPANTS

Twenty-four right-handed subjects took part in each experiment (speaking: 12 female; mean age 25.2 years, SD 7.5; control: 13 female; mean age 22.8 years, SD 3.17). All subjects were healthy native Dutch speakers with normal or corrected-to-normal vision and had attended or were attending university education in the Netherlands. All subjects gave written informed consent prior to the experiment and received a fee or course credit for their participation. No participants took part in both experiments.

STIMULI

Our target stimulus set contained 1152 photographs that depicted 36 transitive events such as *kiss, help, strangle* with the agent and patient of this action. Four couples performed each action (2 × men/women; 2 × boy/girl), in two configurations (one with the man/boy as the agent and one with the woman/girl). These

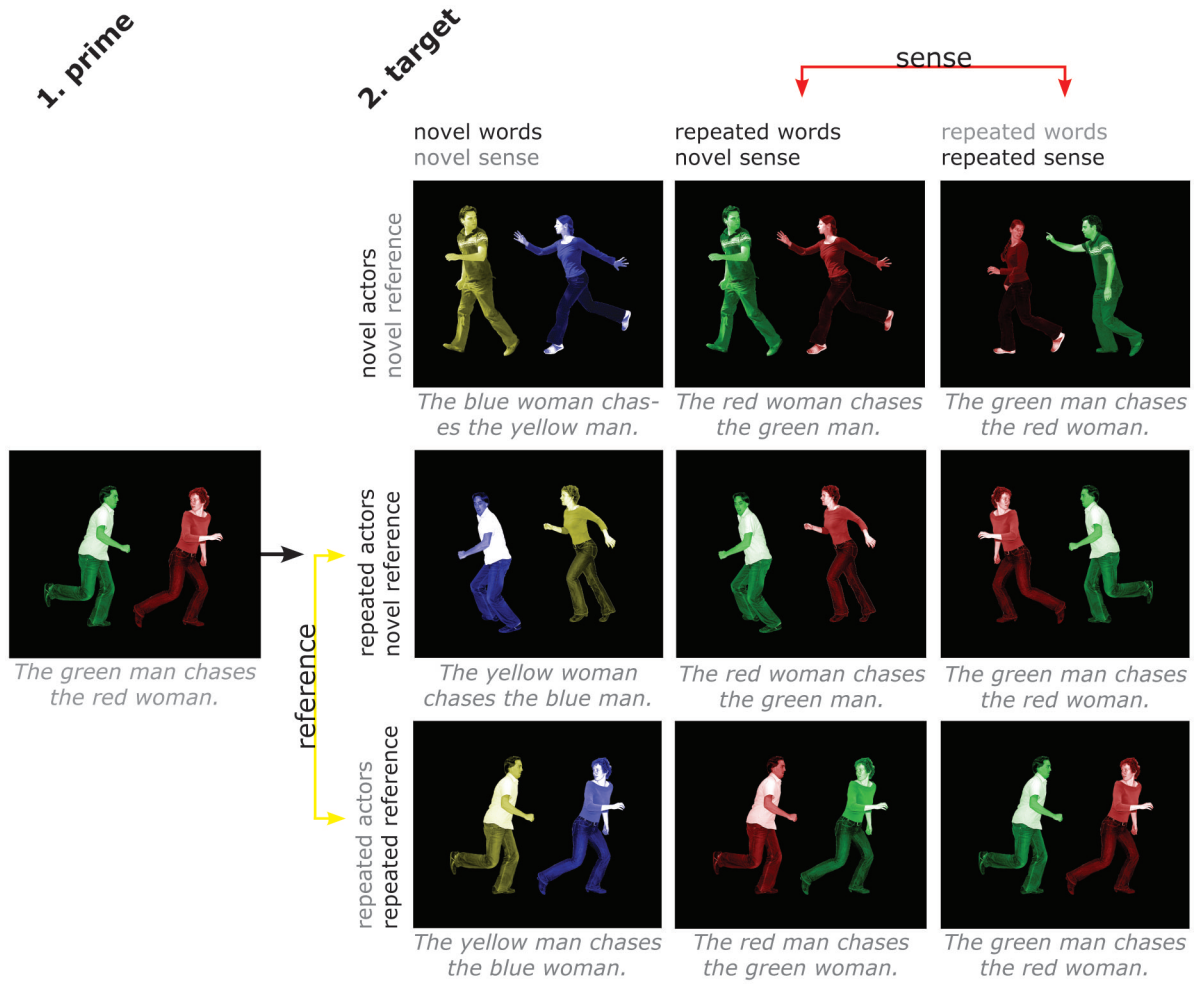

FIGURE 1 | Design and stimuli. Subjects described scenes with short sentences containing action, agent, and patient and their respective colors. There are four factors (words, sense, actors, reference), with two levels (repeated/novel) each. Separate repetition of sense and individual words allows distinguishing areas that are sensitive to the overall meaning of a sentence from those that are sensitive to repetition of words but not to repetition of sentence meaning. Separate repetition of actors and reference allows distinguishing areas that are sensitive to the overall situation from those that are sensitive to the people involved irrespective of the specific situation. The four factors cannot be fully crossed, since it is not possible to repeat sense without, at least partly, repeating words, as it is not possible to repeat a situation while changing the people involved. Therefore, each of the factors varies only at one level of the factor it could potentially be confounded with. The contrasts of interest are between the middle and right column for sense (hence keeping repetition of words controlled) and between the middle and bottom row for reference (hence keeping repetition of actors controlled). Likewise, the effect of word repetition was computed by comparing the leftmost and middle column, and the effect of actor repetition by comparing the top and middle row.

$36 \times 4 \times 2$ pictures were further edited so that the agent and patient each had a different color (red–green, green–red, blue–yellow, yellow–blue), and these $36 \times 4 \times 2 \times 4$ pictures were also flipped so that the position of the agent could be either left or right on the picture. The filler stimuli contained pictures depicting 868 intransitive (e.g., *The boy runs.*) and 160 locative (e.g., *The balls lie on the table.*) events. The actors and objects in these pictures were also colored in red, green, yellow or blue. The control experiment further included catch stimuli, which constituted 10% of the trials. These were pictures similar to the target pictures, but containing a range of visual defects that the subjects had to detect. The stimuli are available for use from the authors.

DESIGN

The design is illustrated in **Figure 1**, and was identical for both experiments. There were four factors (words, sense, actors, reference), with two levels (repeated/novel) each. Contrasting repetition of sense and individual words allowed us to distinguish areas that are sensitive to the overall meaning of a sentence, and those that are sensitive to repetition of words but not to repetition of sentence meaning. The verb and nouns were always repeated for target trials, and only the adjectives could vary. This was necessary due to the constraints on repetition of elements in the pictures for the different conditions. For instance, since "repeated reference" entailed repeating both the action and the people involved, this meant also repeating the words used to refer to these elements.

Contrasting repetition of actors and reference allowed us to distinguish areas that are sensitive to the overall situation from those that are sensitive to the people involved irrespective of the specific situation. The four factors could not be fully crossed, since it is not possible to repeat sense without, at least partly, repeating words, like it is not possible to repeat an event (the reference) while changing the people involved. Therefore, we performed the relevant comparison for each factor at only one level of the factor it could potentially be confounded with (see **Figure 1**).

The target items were presented in 78 mini-blocks with an average length of 5.4 items (range 3–7 items). The target blocks were alternated with filler blocks, with an average length of 3.5 items. Filler blocks served the purpose of increasing variability in syntactic structures, words, and visual properties of the sentences and pictures. Subjects were unaware of the division in blocks, as the items were presented at a constant rate. We used a running priming paradigm where each target item also served as prime for the subsequent target item. No condition was repeated twice in a row. Since there were 78 target blocks, 78 transitive items (the first of each mini-block) served as primes only. The remaining 315 transitive items (2–6 per block depending on block length) constituted the target trials so that there were 35 items per condition. Each subject saw a different randomized list, which consisted of 393 transitive (78 prime-only and 315 target items) stimuli and 262 filler stimuli. For the speaking experiment, these were randomly sampled from the 868 intransitive and 160 locative pictures in the filler stimulus set. In the control experiment, the 262 pictures were always 65 catch (10% of total number of trials), 67 locative and 130 intransitive pictures.

TASK AND PROCEDURE

Speaking experiment: participants first read the instructions and were given the opportunity to ask questions. The instructions not only explained the task, but also introduced all the different frequently occurring actors as separate individuals, along with the same photo of them in every color. This way, we made sure that the participants were aware that the colors were arbitrarily varying properties of the different actors.

Each target picture was preceded by its corresponding verb. Participants described the picture with a short sentence, using the presented verb. In this sentence they had to mention both persons and their colors. The experiment consisted of two runs of 39 min. This served the purpose of not keeping participants in the MRI-scanner for too long; the runs were otherwise completely equivalent. The participants underwent a 5-min anatomical scan after the first run, and were then taken out of the MR-scanner for a break before they underwent the second run. The responses were recorded in order to extract reaction times (RTs). The experimenter coded the participant's responses online for correctness and prevoicing. Prevoicing was coded to ensure correct measurement of RTs, which were extracted through thresholding of the speech recording (see below for details). Each trial lasted 7000 ms and consisted of the following events: the verb was presented with a jittered start time of 0–1000 ms after the start of the trial, and a duration of 500 ms. After an ISI of 500–2500 ms the picture was presented for 2000 ms.

Control experiment: participants first read the instructions and were given the opportunity to ask questions. The participant's task was to act as a "proof viewer" scanning a set of pictures for misprints. They were given examples of both correct pictures and possible misprints. They were instructed to press a button whenever they detected a misprint, and to do nothing if the pictures were ok. The experiment consisted of two runs of 22 min. The participants underwent a 5-min anatomical scan between runs. Each trial lasted 4000 ms, in which the picture was displayed with a jittered start time of 0–1500 ms from trial onset, and stayed on screen for 1000 ms. We chose different timing parameters for this experiment, to avoid it becoming incredibly boring.

DATA ACQUISITION AND ANALYSIS

Data acquisition took place in a 3-T Siemens Magnetom Tim-Trio MRI-scanner. Participants were scanned using a 12-channel surface coil. To acquire our functional data we used parallel-acquired inhomogeneity-desensitized fMRI (Poser et al., 2006). This is a multi-echo EPI: images are acquired at multiple TE's following a single excitation. The TR was 2398 ms and each volume consisted of 31 slices of 3 mm thickness with a slice-gap of 17%. The voxel size was 3.5 mm × 3.5 mm × 3 mm and the field of view was 224 mm. Functional scans were acquired at multiple TEs following a single excitation (TE_1 at 9.4 ms, TE_2 at 21.2 ms, TE_3 at 33 ms, TE_4 at 45 ms, and TE_5 at 56 ms with echo spacing of 0.5 ms) so that there was a broadened T_2^* coverage. Because T_2^* mixes into the five echoes in a different way, the estimate of T_2^* is improved. Accelerated parallel imaging reduces image artifacts and thus is a good method to acquire data when participants are producing sentences in the scanner (causing motion and susceptibility artifacts). The number of slices did not allow acquisition of a full brain volume in most participants. We always made sure that the entire temporal and frontal lobes were scanned because these were the areas where the fMRI adaptation effects of interest were expected. This meant however that data from the superior posterior frontal lobe and the anterior superior parietal lobe (thus data from the top of the head) were not acquired in several participants. The functional scans of the first and second runs were aligned using AutoAlign. A whole-brain high resolution structural T1-weighted MPRAGE sequence was performed to characterize participants' anatomy (TR = 2300 ms, TE = 3.03 ms, 192 slices with voxel size of 1 mm^3, FOV = 256), accelerated with GRAPPA parallel imaging (Griswold et al., 2002).

For the behavioral data of the speaking experiment, to separate participants' speech from the scanner sound and extract RTs, the speech recordings were bandpass filtered with a frequency band of 250–4000 Hz and smoothed with a width half the sampling rate. Response onsets and durations were determined through thresholding of these filtered recordings (basically, a *post hoc* voicekey) and linked to the stimulus presentation times to extract the RTs and total speaking times. Trials with errors or prevoicing were discarded from the analysis. The planning times, speaking times and total response times for correct responses to the target items were analyzed in a repeated measures ANOVA using SPSS.

The fMRI data were preprocessed and analyzed with SPM5 (Friston et al., 1995). The first 5 images were discarded to allow for T_1 equilibration. Then the five echoes of the remaining images

were realigned to correct for motion artifacts (estimation of the realignment parameters was done for one echo and then copied to the other echoes). Subsequently the five echoes were combined into one image with a method designed to filter task-correlated motion out of the signal (Buur et al., 2009). First, echo two to five (i.e., TE_2, TE_3, TE_4, and TE_5) were combined using a weighting vector dependent on the measured differential contrast to noise ratio per voxel. The time course of an image acquired at a very short echo time (i.e., TE_1) was used as a voxelwise regressor in a linear regression for the combined image of TE_2, TE_3, TE_4, and TE_5. Weighting of echoes was calculated based on 25 volumes acquired before the actual experiment started. The resulting images were coregistered to the participants' anatomical scan, normalized to MNI space and spatially smoothed using a 3D isotropic Gaussian smoothing kernel (FWHM = 8 mm).

We then performed first- and second-level statistics. For the first level general linear model (GLM), we modeled the individual start time of the picture. The events of our model were convolved with the canonical hemodynamic response function included in SPM5. In the speaking experiment, the first level model included verbs, filler pictures, prime pictures, the nine conditions and errors. Erroneous responses were therefore put in a separate regressor, leaving only correct responses in the actual analyses. For the control experiment, the first level model included filler pictures, prime pictures, the nine conditions, and catch trials. Both models included the six motion parameters as event-related regressors of no interest. The second-level model consisted of a 9 (condition) × 2 (experiment) factorial design. All effects were then tested by computing the appropriate contrasts for the model. We performed two types of analyses to test our hypotheses: to find intersections between different effects, we conducted conjunction analyses. In these analyses multiple different contrasts are tested, and only areas showing an effect in all tested contrasts under a conjunction null hypothesis result in a significant conjunction (Friston et al., 2005). To look for areas sensitive to one factor but not the other, we applied exclusive masking. In such an analysis, the significant clusters for one factor are overlaid with a low-threshold mask for the other factor ($p < 0.20$ uncorrected voxelwise), and only clusters that survive the masking procedure are reported. Note that due to the very nature of the type of statistical framework we employ, we cannot prove that an effect does *not* exist. However, if an effect does not survive thresholding at $p < 0.20$ uncorrected voxelwise, it may be said to be very weak at the very least. For all tests, the cluster size at voxelwise threshold $p < 0.001$ uncorrected was used as the test statistic and only clusters significant at $p < 0.05$ corrected for multiple non-independent comparisons are reported. Local maxima are also reported for all clusters with their respective voxelwise family wise error (FWE) corrected p-values. The effects for repetition of words and actors are reported in the tables, but since the aim of the study is to distinguish reference and sense we focus on those two factors in discussing the results.

RESULTS

BEHAVIORAL DATA

For the speaking experiment, we performed repeated measures GLMs on the planning times (RTs), speaking times (the duration of the response), and the total planning + speaking times. The model included one factor (condition) with 9 levels, and the three dependent variables. The effects reported were computed through custom hypothesis tests within this model, using contrasts much like for the fMRI analyses. The data are reported in **Figure 2**. For planning times, repetition of sense, actors and reference produced significant priming effects [words: $F < 1$; sense: $F(1,23) = 109.53$, $p < 0.001$; actors: $F(1,23) = 22.95$, $p < 0.001$; reference: $F(1,23) = 94.60$, $p < 0.001$]. For speaking times, repetition of reference and sense significantly affected the duration of the response [words: $F(1,23) = 1.52$, $p < 0.232$; sense: $F(1,23) = 12.31$ $p < 0.002$; actors: $F(1,23) = 3.71$, $p < 0.066$; reference: $F(1,23) = 9.50$, $p < 0.005$]. However, the direction of these effects was reversed. Priming led to shorter planning times but longer speaking times. Analyses on the total time the participants took to complete the response (so planning plus speaking time) again revealed significant effects for reference and sense [words: $F(1,23) < 1$; sense: $F(1,23) = 13.41$, $p < 0.001$; actors: $F(1,23) = 2.78$, $p < 0.11$; reference: $F(1,23) = 33.307$, $p < 0.001$]. The total response time mirrored the planning time pattern: when primed, subjects were faster to complete the response. There were no significant interactions in any of the analyses, in so far as these could be computed given the design. In the control experiment, the average d-prime was 0.7, indicating that participants did pay attention.

fMRI RESULTS

All results are reported in **Tables 1** and **2**, and depicted in **Figure 3**. **Table 2** lists the main effects for all factors in the design; we limit the discussion to the more specific results for reference and sense as listed in **Table 1**.

Speaking experiment

The first step in semantic encoding, the input, is to compute a non-linguistic representation underlying the sentence to be produced. We therefore looked for areas exhibiting fMRI adaptation to reference, while not showing an effect of sense repetition. The BOLD-response of the bilateral temporo-parietal junctions (BA 39/37/19) and the precuneus decreased after repetition of reference. In the right middle frontal and inferior gyrus (BA 45/46) the BOLD-response increased after repetition of reference.

The next step in semantic encoding, is to map the reference onto a linguistic semantic structure that can be expressed, the sense. This is the actual process of semantic encoding. We therefore tested for areas sensitive to repetition of both reference and sense. The bilateral superior parietal lobes (BA 7), fusiform gyrus (BA 37) and posterior middle temporal areas showed suppression in the conjunction analysis. The left calcarine sulcus (BA 17) exhibited suppression as well. Finally, three frontal clusters in the left middle frontal gyrus, left SMA and left precentral gyrus (all BA 6) also showed repetition suppression. We also tested for increased responses upon repetition (enhancement), but found no areas exhibiting this pattern.

Finally, the mapping process produces an output, the sense. This should be reflected in regions showing fMRI adaptation to sense, without showing an effect for reference. The BOLD-response of the pars triangularis in the left inferior frontal gyrus (BA 45) was reduced after repetition of sense. The response in the left angular

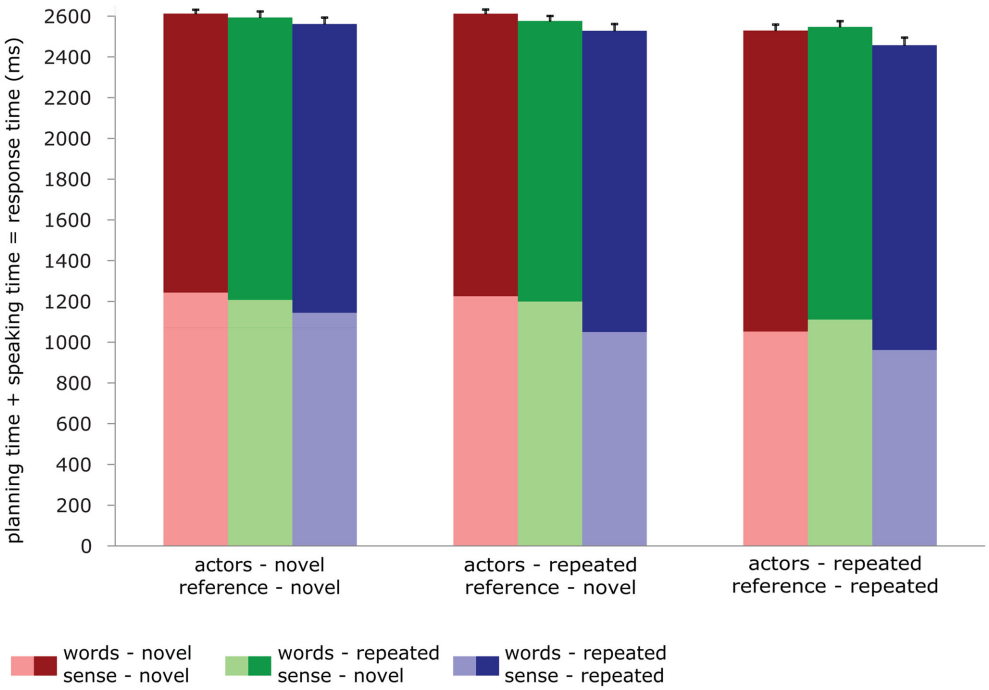

FIGURE 2 | Behavioral data in the speaking experiment: reaction times (light shades), durations (dark shades), and total speaking times (total bar length) for all conditions. Error bars represent SE of the mean of the total speaking times.

gyrus (BA 39/19) and in the left middle frontal gyrus (BA 44/9), on the other hand, increased after repetition of sense.

Control experiment

In the process of semantic encoding, the input is the conceptual representation that has to be transformed into a preverbal message. To some extent at least, such a representation should also be constructed when we are not speaking. The only significant effect in the control experiment was indeed a main effect of repetition of reference in bilateral posterior middle temporal gyri/inferior parietal lobe (BA 37/39). This effect survived masking with sense in the control experiment, and was also the same as the main effect of reference in the speaking experiment, as demonstrated by a conjunction analysis (**Table 3**). The right middle temporal gyrus (BA 21) showed enhancement upon repetition of reference.

DISCUSSION

In this sentence production study, we aimed to distinguish brain areas sensitive to reference (the mental representation an utterance refers to) and the sense (the linguistic structure that interfaces meaning with linguistic form). The behavioral data in the speaking experiment showed that both reference and sense priming affect the responses, and that these two effects do not interact. This shows that both processes are psychologically real and distinct, and that priming them affects the speed with which a sentence can be produced.

In speaking, constructing an utterance is an incremental process, involving several steps (Levelt, 1989). The first is to construct a *preverbal message*. In the present experiment, this requires encoding the situation we want to talk about (MAN_a hitting $WOMAN_b$) into a thematic role structure which can be described as HIT(MAN(YELLOW), WOMAN(BLUE)): there is a HIT event, performed by a MAN, who has the property of being YELLOW, at the expenses of a WOMAN who has the property of being BLUE (perhaps reasonable given that she is being HIT). As outlined in the introduction, the input is the reference, the output the sense. We wanted to find out which areas in the brain are involved in this mapping process. In the following, we will trace step by step how, based on our results, we think the brain comes to encode an utterance.

The first step is to build a representation of a situation we are going to talk about – the reference. This representation forms the input to semantic encoding, and is non-linguistic (conceptual) in nature. As outlined in the introduction, in the case of a concrete referent this representation is the result of perceptual processes within the perceptual system – in the present case, the visual system. Presumably, such a representation is, at least to some extent, built for what we perceive independently of whether we are going to talk about it or not. In the present paradigm, this step should be independent of the sense of the final utterance. Areas showing suppression to repetition of reference but not sense were the bilateral occipito-temporo-parietal junctions (BA 37/39/19) and the precuneus. Data from the control experiment corroborate the idea that the role of these areas in reference in the present experiment is primarily to build a perceptual representation: the same bilateral areas at the junction of the occipital, temporal and parietal lobes show suppression to repetition of reference in the absence

Table 1 | Overlap and segregation of reference and sense.

Effect	Cluster	BA	Anatomical label	Global and local maxima			Cluster-level		Voxel-level	
				x	*y*	*z*	*K*	*p*	*T*	*p* (FWE)
Sense and reference	1	37	Fusiform_L	−42	−60	−12	4162	0.000	5.95	0.000
		7	Parietal_Sup_L	−24	−62	54			5.57	0.001
		7	Parietal_Sup_L	−22	−72	44			5.33	0.004
	2	6	Frontal_Mid_L	−26	−8	50	728	0.000	5.73	0.001
		6	Precentral_L	−40	0	50			4.99	0.018
	3	19	Occipital_Inf_R	44	−76	−2	1323	0.000	4.90	0.026
		19	Occipital_Inf_R	38	−72	−8			4.86	0.030
		37	Temporal_Mid_R	46	−66	10			4.41	0.161
	4	6	Supp_Motor_Area_L	−4	10	56	191	0.035	4.76	0.045
	5	7	Parietal_Sup_R	26	−58	54	461	0.000	4.74	0.048
		7	Parietal_Sup_R	24	−72	50			4.27	0.247
		7	n/a	24	−48	48			3.94	0.583
	6	6/44	Precentral_L	−44	6	20	299	0.005	4.42	0.153
		6	Precentral_L	−42	−4	32			4.27	0.253
	7	17	Calcarine_L	−8	−92	6	329	0.003	4.08	0.423
		17	Calcarine_L	−10	−82	8			3.77	0.772
		17	Calcarine_L	−14	−68	8			3.74	0.807
Sense-suppression	1	n/a	Frontal_Inf_Tri_L	−34	18	24	186	0.039	4.71	0.054
(masked for reference)		45	Frontal_Inf_Tri_L	−50	28	26			4.46	0.136
		45	Frontal_Inf_Tri_L	−38	26	26			4.10	0.409
Sense-enhancement	1	39	Angular_L	−54	−58	36	497	0.000	5.85	0.000
(masked for reference)		19	Occipital_Mid_L	−42	−74	38			4.23	0.283
	2	9	Frontal_Mid_L	−26	24	44	251	0.012	4.03	0.476
		44	Frontal_Mid_L	−42	18	40			4.01	0.501
Reference-suppression	1	39	Temporal_Mid_R	50	−66	20	805	0.000	9.44	0.000
(masked for sense)		39	Occipital_Mid_R	42	−74	26			8.09	0.000
		37	Temporal_Mid_R	60	−60	12			7.81	0.000
	2	39	Temporal_Mid_L	−42	−68	20	617	0.000	8.99	0.000
		39	Temporal_Mid_L	−54	−66	18			7.37	0.000
		19	Occipital_Mid_L	−30	−78	32			5.35	0.002
	3	23	Precuneus_R	4	−58	24	1389	0.000	5.56	0.001
		n/a	Precuneus_R	2	−54	40			5.24	0.006
		n/a	Precuneus_L	−10	−50	52			5.16	0.009
Reference-enhancement	1	45	Frontal_Mid_R	38	46	10	416	0.001	4.79	0.040
(masked for sense)		46	Frontal_Inf_Orb_R	44	48	−4			4.31	0.218

Listed are the MNI-coordinates for the first three local maxima for each significant cluster in the relevant comparisons ($p < 0.05$ corrected cluster-level, threshold $p < 0.001$ uncorrected voxelwise, exclusive masks $p < 0.20$ uncorrected voxel-wise). Anatomical labels are derived from the Automatic Anatomical Labeling atlas (Tzourio-Mazoyer et al., 2002) and from Brodmann's atlas. Cluster-level statistics are listed for each cluster, voxel-level statistics also for local maxima.

of a linguistic task. The finding that these areas are involved in generating the non-linguistic representation to refer to now also allows us to further specify a previous finding on semantic encoding in sentence production: in a previous study, we found that part of the superior right MTG is sensitive to sentence- but not word-level meaning (Menenti et al., submitted). This effect overlaps with the area sensitive to repeated reference but not sense, and therefore was presumably due to the encoding of the referent as well. These same regions have also been found sensitive to subsequent memory for short stories (Hasson et al., 2007), a further indication that they are involved in constructing a representation of what linguistic material refers to. Repetition of reference did not just elicit suppression: in right inferior frontal gyrus the response increased upon repeated presentation. The repetition enhancement effect for reference in right inferior frontal cortex was particularly striking since large parts of contralateral left inferior frontal cortex showed repetition suppression for reference. Repetition enhancement has been postulated to be caused, among other things, by novel network formation due to the construction of new representations (Henson et al., 2000; Conrad et al., 2007; Gagnepain et al., 2008; Segaert et al., submitted). In speech comprehension, right inferior frontal cortex has previously been implicated in the construction of a situation model (Menenti et al., 2009; Tesink et al., 2009), a mental representation of text containing

Table 2 | Main effects for all factors in the design.

Effect	Cluster	BA	Anatomical label	Global and local maxima			Cluster-level		Voxel-level	
				x	*y*	*z*	*K*	*p*	*T*	*p* (FWE)
Main effect words	No significant clusters									
Main effect sense	1	6	Supp_Motor_Area_L	−4	8	54	1389	0.000	7.71	0.000
		n/a	n/a	−10	20	22			4.44	0.146
		n/a	Cingulum_Mid_L	−10	16	32			4.12	0.387
	2	7	Occipital_Mid_L	−26	−60	42	5208	0.000	6.63	0.000
		2	Parietal_Inf_L	−46	−38	50			6.07	0.000
		37	Fusiform_L	−42	−60	−12			5.95	0.000
	3	6	Precentral_L	−46	−2	34	3488	0.000	6.53	0.000
		6	Precentral_L	−46	0	42			6.23	0.000
		44	Frontal_Inf_Oper_L	−46	8	22			6.18	0.000
	4	n/a	Thalamus_L	−10	−14	10	983	0.000	5.33	0.004
		n/a	Thalamus_L	−4	−20	12			4.99	0.018
		n/a	Thalamus_R	4	−18	6			4.61	0.079
	5	19	Occipital_Inf_R	44	−76	−2	2342	0.000	4.90	0.026
		18	Vermis_6	3	−72	−8			4.86	0.030
		7	Parietal_Sup_R	26	−58	54			4.74	0.048
	6	17	Calcarine_L	−8	−92	6	378	0.000	4.08	0.423
		19	Calcarine_L	−20	−66	6			3.97	0.556
		17	Calcarine_L	−12	−82	6			3.91	0.617
Main effect actors	1	7	Parietal_Sup_L	−28	−60	46	374	0.002	4.96	0.020
	2	44	Frontal_Inf_Oper_R	42	10	28	471	0.000	4.89	0.027
		45	Frontal_Inf_Tri_R	42	24	22			4.01	0.500
		45	Frontal_Inf_Tri_R	46	34	16			3.97	0.548
	3	37	Fusiform_R	40	−44	−18	396	0.001	4.88	0.028
		37	Temporal_Inf_R	42	−58	−10			4.76	0.045
	4	44	Frontal_Inf_Oper_L	−36	8	28	252	0.012	4.33	0.209
		n/a	Frontal_Inf_Tri_L	−38	20	24			3.66	0.875
		45	Frontal_Inf_Tri_L	−42	28	18			3.44	0.979
	5	37	Ternporal_Mid_R	36	−58	14	332	0.003	4.11	0.395
		40	n/a	30	−56	34			4.04	0.469
		19	n/a	32	−64	28			3.94	0.587
Main effect reference	1	37	Ternporal_Mid_R	54	−58	6	21483	0.000	12.62	0.000
		37	Ternporal_Mid_R	48	−64	12			11.49	0.000
		19	Occipital_Mid_L	−48	−74	4			11.27	0.000
	2	6	Frontal_Sup_L	−26	−6	55	851	0.000	5.97	0.000
	3	20	Ternporal_Mid_R	50	−10	−18	307	0.005	5.08	0.012
		20	Fusiform_R	42	−16	−20			3.62	0.903
	4	6	Frontal_Sup_R	30	−6	60	598	0.000	5.00	0.017
		n/a	Frontal_Mid_R	24	12	42			4.56	0.093
		6	Precentral_R	40	0	46			4.01	0.504
	5	6	Supp_Motor_Area_L	−4	10	56	236	0.016	4.76	0.045
		6	Frontal_Sup_L	−14	10	50			3.83	0.716
	6	27	Lingual_R	6	−34	−4	479	0.000	4.64	0.072
		27	n/a	−8	−28	−4			4.31	0.222
		27	Thalamus_L	−16	−30	6			3.84	0.702
	7	44	Precentral_L	−44	6	20	299	0.005	4.42	0.153
		6	Precentral_L	−42	−4	32			4.27	0.253

Listed are the MNI-coordinates for the first three local maxima for each significant cluster in the relevant comparisons ($p < 0.05$ corrected cluster-level, threshold $p < 0.001$ uncorrected voxelwise). Anatomical labels are derived from the Automatic Anatomical Labeling atlas (Tzourio-Mazoyer et al., 2002) and from Brodmann's atlas. Cluster-level statistics are listed for each cluster, voxel-level statistics also for local maxima.

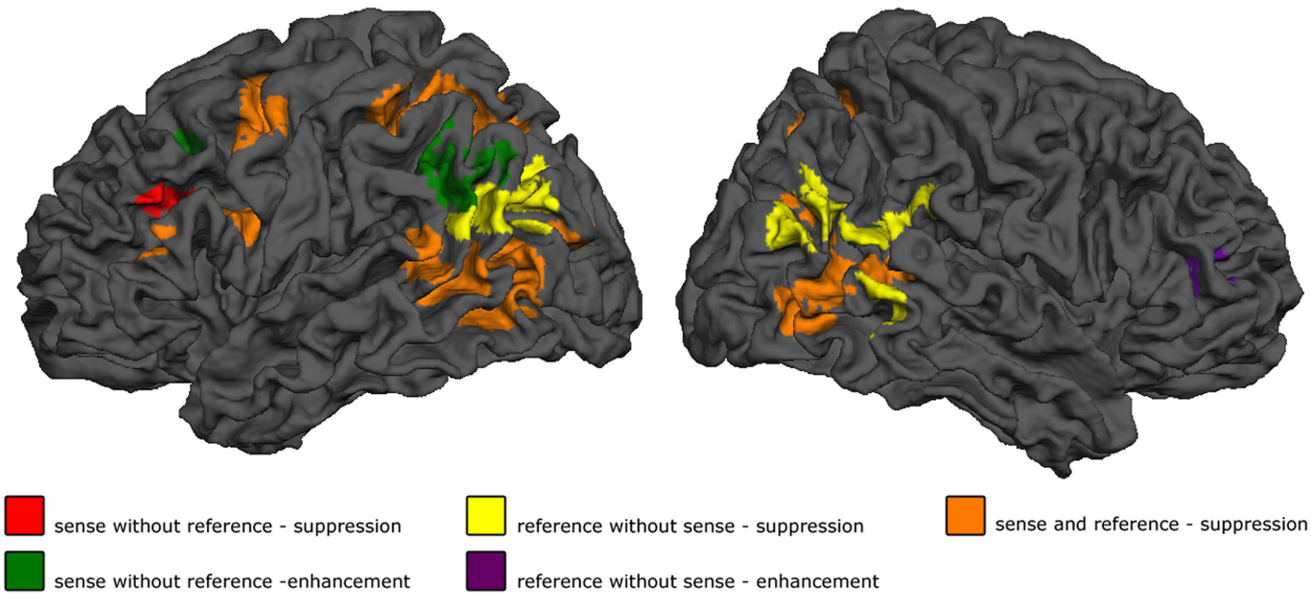

FIGURE 3 | Results. Shown are the effect of sense masked exclusively for reference, for reference masked exclusively for sense, and the conjunction of the two. $p < 0.05$ cluster-level FWE corrected for simple effects, maps thresholded at $p < 0.001$ voxelwise uncorrected, exclusive masks thresholded at $p < 0.20$ voxelwise uncorrected.

information on, for instance, space, time, intentionality, causation and protagonists (Zwaan and Radvansky, 1998). These are integrated and updated over several sentences and also contain all inferences that were not explicitly stated but are necessary for comprehension (Zwaan and Radvansky, 1998). The difference between reference as discussed above and situation models is that the latter pertain to the integration of referents of several utterances into one mental model and also contain unstated information, arrived at through inferences. A similar distinction is likely in production: the situation model may contain any information that the speaker knows pertains to the situation, but that he does not mention. Right inferior frontal gyrus has repeatedly been found to be involved in generating inferences (Mason and Just, 2004; Kuperberg et al., 2006). The first presentation of a referent may therefore induce the start of situation model construction. This same area did not show enhancement in the control experiment, supporting the idea that the process in which this region is involved is language-related. We do not currently have an explanation for the enhancement effect found in right middle temporal gyrus in the control experiment.

The second main step in semantic encoding is to map the representation that we want to talk about onto a linguistic structure that can be syntactically encoded – the actual process of encoding. This would presumably involve areas sensitive to both reference and sense, interfacing between the mental representation of the situation that will be described and the linguistic representation describing it. What is perhaps most striking about our data, is the great extent to which these two processes are neurally intertwined: bilateral posterior middle temporal gyri (BA 37), superior parietal areas (BA 7), precentral gyrus (BA 6) and LIFG (BA 44/6) all show largely overlapping suppression effects for reference and sense. Our data show that large parts of the language network are involved in processing reference, and that reference therefore presumably is important throughout much of the task of building an utterance. But what is the contribution of all these areas to semantic encoding? Due to the proximity of areas coding the perceptual representation of the referent and some of the areas involved in processing both reference and sense, we hypothesize that the bilateral temporal areas sensitive to reference and sense are primarily involved in mapping one onto the other. Such mapping requires the retrieval of the relevant lexical items from the mental lexicon, which indeed has often been postulated to involve the posterior middle temporal gyrus (Hagoort, 2005; Jung-Beeman, 2005). The bilateral superior parietal lobes also showed suppression to the repetition of both reference and sense. These parietal areas have previously been found involved in studies investigating linguistic inference (Nieuwland et al., 2007; Monti et al., 2009). In the sense/reference fMRI study discussed in the introduction, the parietal areas were more strongly activated for both referentially ambiguous and anomalous conditions compared to coherent conditions, but this effect was more pronounced for the ambiguous condition (Nieuwland et al., 2007). In a study on linguistic and logical inference, this area was found to be common to both types of inference compared to detection of grammatical violations (Monti et al., 2009). Our suppression effect in this area may reflect that in a situation where sense, reference, or both are repeated, less inferences are required than in a situation where that is not the case. The superior LIFG (BA6) also showed suppression both to repetition of sense and of reference. On the hypothesis that IFG is involved in unifying different elements into a coherent representation (Hagoort, 2005), this means that the reference of an utterance is also kept active in the working space of language. The fact that

Table 3 | Results from control experiment.

Effect	Cluster	BA	Anatomical label	Global and local maxima			Cluster-level		Voxel-level	
				x	*y*	*z*	*K*	*p*	T	*p* (FWE)
Main effect reference suppression	1	39	Temporal_Mid_L	−44	−64	16	885	0.000	6.42	0.000
		37	Temporal_Mid_L	−54	−54	4			4.97	0.015
		19	Occipital_Mid_L	−34	−76	26			3.26	0.996
	2	39	Temporal_Mid_R	40	−62	18	1379	0.000	5.66	0.001
		21	Temporal_Mid_R	46	−56	12			5.34	0.009
		37	Temporal_Mid_R	50	−66	6			5.31	0.009
Main effect reference enhancement	1	21	Termporal_Mid_R	56	−24	−8	275	0.010	5.22	0.005
Reference masked for sense	1	39	Temporal_Mid_L	−44	−64	16	695	0.000	6.42	0.000
		37	Temporal_Mid_L	−54	−54	4			4.97	0.015
		39	Occipital_Mid_L	−34	−76	26			3.26	0.996
	2	39	Temporal_Mid_R	40	−62	18	1313	0.000	5.66	0.001
		39	Temporal_Mid_R	46	−56	12			5.34	0.009
		37	Temporal_Mid_R	50	−66	6			5.31	0.009
Conjunction reference speaking and reference control	1	39	Temporal_Mid_L	−44	−64	16	885	0.000	6.42	0.000
		37	Temporal_Mid_L	−54	−54	4			4.97	0.015
		19	Occipital_Mid_L	−34	−76	26			3.26	0.996
	2	39	Temporal_Mid_R	40	−62	18	1346	0.000	5.66	0.001
		21	Temporal_Mid_R	46	−56	12			5.34	0.003
		37	Temporal_Mid_R	50	−66	6			5.31	0.003
Sense	No significant clusters									
Actors	No significant clusters									
Words	No significant clusters									

Listed are the MNI-coordinates for the first three local maxima for each significant cluster in the relevant comparisons ($p < 0.05$ corrected cluster-level, threshold $p < 0.001$ uncorrected voxelwise). Anatomical labels are derived from the Automatic Anatomical Labeling atlas (Tzourio-Mazoyer et al., 2002) and from Brodmann's atlas. Cluster-level statistics are listed for each cluster, voxel-level statistics also for local maxima.

none of the regions outlined above are sensitive to any of our factors in the control experiment further indicates that the process they are involved in is linguistic in nature.

The output of semantic encoding is the sense. One area showed a repetition suppression effect for sense but not reference: the left inferior IFG (BA 45). The final, linguistic, sense is apparently assembled in LIFG. This effect may, however, also be partly due to the repetition of the exact sentence, therefore by repetition of not just semantic but also both syntactic and phonological sequencing processes, which are related to actual speech output and are not part of the sense. In fact, the focus of the effect, lying at the heart of the part of LIFG most often found involved in syntactic processing (Bookheimer, 2002), suggests just that. Ventral LIFG, most commonly known to be involved in meaning processing (Bookheimer, 2002), remains sensitive to reference throughout.

Repetition of sense also elicits enhancement in two areas. The exact same left hemispheric frontal and parietal areas here showing repetition enhancement for sense have previously been found to be involved in semantic inhibition (Hoenig and Scheef, 2009), that is, inhibition of contextually inappropriate meanings. In the present paradigm, each word (MAN, BOY, WOMAN, GIRL) has two prominent possible referents. One of them has to be suppressed in mapping the intended referent onto the sense. While this would seem harder in the case where sense is not repeated (and therefore elicit suppression instead of enhancement upon repetition), this seeming incongruity can be readily explained: the BOLD-response in both areas shows consistent deactivation in any of the conditions compared to an implicit baseline. The deactivations are less strong in the conditions with repeated sense, than those where sense is novel. This mirrors activation patterns in the so-called default mode network, which shows increasing *de*activations depending on task difficulty (Greicius et al., 2003). Both areas have been shown to be part of the default mode network.

In sum, our data suggest that the bilateral temporo-parietal-occipital junctions are involved in constructing a mental representation of a percept (the reference), that the bilateral posterior middle temporal gyri map this representation onto lexical items that can be expressed, and that the final sense is unified in left inferior frontal gyrus – this can then serve as input to both syntactic and phonological encoding which also involve left inferior frontal gyrus.

Some caveats are in order: in operationalizing reference and sense for the purpose of this study, we have made some decisions that limit the generalizability of our findings. Most notably, our experiments concern visual representations of concrete events.

As we have stressed above, we consider referents to be mental representations in our mind. These mental representations are likely to differ depending on the material underlying them. They will likely be different for auditory and visual objects, for events involving people and for non-human objects, for concrete objects and for abstract concepts. But that is precisely the point: our brains need to convert non-linguistic mental (i.e., conceptual) representations, whatever they are "made of" into language. Therefore, while we believe our findings concerning sense, and the mapping of reference onto sense will at least to a large extent hold irrespective of the underlying reference, what brain areas are involved in processing reference alone will depend on the specifics of the mental representation involved.

Another constraint concerns our task. We had participants describe a long list of pictures. If these subsequent sentences were to be perceived as part of an ongoing discourse, then some unnatural situations would arise: normally, we would avoid repeating the same sentence twice in a row, let alone while using it to refer to different things. Our behavioral data, however, provide an indication that participants were not too affected by such concerns. First, the instructions specified that they had to name the people, the colors, and the action (which was given by the verb presented prior to the picture). Though this precluded using pronouns, this did not prevent participants from adding specifications such as "the other," "again," "now," etc., to specify the relation between pictures. No participants chose to do so. Second, if repeating the sentence were more difficult than not repeating it, we should have seen an inhibitory effect of priming. While we did see this in the speaking times, we did not in the planning times, and the total time taken to compete an utterance was shorter for the primed than the unprimed conditions. These are indications that our participants were happy to consider every trial an independent unit. We believe that single sentence processing is conceptually the same as discourse processing, but on a smaller scale. Therefore, we predict our general findings would hold for more natural processing of language in context.

To conclude, our data confirm that the theoretical distinction between reference and sense is psychologically real, both in terms of behavior and of neuroanatomy. The behavioral data shows that priming of both processes can affect the ease of production. The fMRI data shows that indeed some brain regions are selectively affected by one of these computations. However, the neuronal infrastructure underlying the computation of reference and sense is largely shared in the brain. This indicates that processing reference and sense is highly interactive throughout the language system.

ACKNOWLEDGMENTS

This research was funded by the NWO Spinoza prize awarded to Peter Hagoort. We wish to thank Geoffrey Brookshire and Josje Verhagen for their assistance in running the control experiment.

REFERENCES

Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu. Rev. Neurosci.* 25, 151–188.

Buur, P. F., Poser, B. A., and Norris, D. G. (2009). A dual echo approach to removing motion artefacts in fMRI time series. *NMR Biomed.* 22, 551–560.

Conrad, N., Giabbiconi, C.-M., Müller, M. M., and Gruber, T. (2007). Neuronal correlates of repetition priming of frequently presented objects: insights from induced gamma band responses. *Neurosci. Lett.* 429, 126–130.

Frege, G. (1892). "On sense and nominatum," in *The Philosophy of Language*, 2nd Edn, ed. A. P. Martinich (New York: Oxford University Press), 190–192.

Friston, K. J., Holmes, A., Worsley, K., Poline, J.-B., Frith, C., and Frackowiak, R. (1995). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210.

Friston, K. J., Penny, W. D., and Glaser, D. E. (2005). Conjunction revisited. *Neuroimage* 25, 661–667.

Gagnepain, P., Chetelat, G., Landeau, B., Dayan, J., Eustache, F., and Lebreton, K. (2008). Spoken word memory traces within the human auditory cortex revealed by repetition priming and functional magnetic resonance imaging. *J. Neurosci.* 28, 5281–5289.

Greicius, M. D., Krasnow, B., Reiss, A. L., and Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 100, 253–258.

Griswold, M. A., Jakob, P. M., Heidemann, R. M., Nittka, M., Jellus, V., Wang, J., Kiefer, B., and Haase, A. (2002). Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn. Reson. Med.* 47, 1202–1210.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends Cogn. Sci. (Regul. Ed.)* 9, 416–423.

Hasson, U., Nusbaum, H. C., and Small, S. L. (2007). Brain networks subserving the extraction of sentence information and its encoding to memory. *Cereb. Cortex* 17, 2899–2913.

Henson, R., Shallice, T., and Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science* 287, 1269–1272.

Hoenig, K., and Scheef, L. (2009). Neural correlates of semantic ambiguity processing during context verification. *Neuroimage* 45, 1009–1019.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution.* Oxford: Oxford University Press.

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends Cogn. Sci. (Regul. Ed.)* 9, 512–518.

Kan, I., and Thompson-Schill, S. (2004). Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cogn. Affect. Behav. Neurosci.* 4, 43–57.

Kuperberg, G. R., Lakshmanan, B. M., Caplan, D. N., and Holcomb, P. J. (2006). Making sense of discourse: an fMRI study of causal inferencing across sentences. *Neuroimage* 33, 343–361.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation.* Cambridge, MA: The MIT Press.

Mason, R. A., and Just, M. A. (2004). How the brain processes causal inferences in text. *Psychol. Sci.* 15, 1–7.

Menenti, L., Gierhan, S. M. E., Segaert, K., and Hagoort, P. (2011). Shared Language. *Psychol. Sci.* 22, 1174–1182.

Menenti, L., Petersson, K. M., Scheeringa, R., and Hagoort, P. (2009). When elephants fly: differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. *J. Cogn. Neurosci.* 21, 2358–2368.

Monti, M. M., Parsons, L. M., and Osherson, D. N. (2009). The boundaries of language and thought in deductive inference. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12554–12559.

Nieuwland, M. S., Petersson, K. M., and Van Berkum, J. J. A. (2007). On sense and reference: examining the functional neuroanatomy of referential processing. *Neuroimage* 37, 993–1004.

Parker Jones, O. I., Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., Ali, N., Lee, H. L., Ramsden, S., Gazarian, K., Prejawa, S., Seghier, M. L., and Price, C. J. (2011). Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cereb. Cortex.* doi: 10.1093/cercor/bhr16/. [Epub ahead of print].

Poser, B. A., Versluis, M. J., Hoogduin, J. M., and Norris, D. G. (2006). BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: parallel-acquired inhomogeneity desensitized fMRI. *Magn. Reson. Med.* 55, 1227–1235.

Tesink, C. M. J. Y., Petersson, K. M., Van Berkum, J. J. A., Van Den Brink, D. L., Buitelaar, J. K., and Hagoort, P. (2009). Unification of speaker and meaning in language comprehension: an fMRI study. *J. Cogn. Neurosci.* 21, 2085–2099.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289.

Zwaan, R. A., and Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychol. Bull.* 123, 162–185.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 July 2011; accepted: 06 December 2011; published online: 18 January 2012.

Citation: Menenti L, Petersson KM and Hagoort P (2012) From reference to sense: how the brain encodes meaning for speaking. Front. Psychology ***2****:384. doi: 10.3389/fpsyg.2011.00384*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

Copyright © 2012 Menenti, Petersson and Hagoort. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.