Cognitive, Affective, & Behavioral Neuroscience (2018) 18:273–283
https://doi.org/10.3758/s13415-018-0568-6

Neural correlates of three cognitive processes involved in theory of mind and discourse comprehension

Nan Lin[1,2] · Xiaohong Yang[1,2] · Jing Li[1,2] · Shaonan Wang[3,4] · Huimin Hua[1,2] · Yujun Ma[5] · Xingshan Li[1,2]

Abstract

Neuroimaging studies have found that theory of mind (ToM) and discourse comprehension involve similar brain regions. These brain regions may be associated with three cognitive components that are necessarily or frequently involved in ToM and discourse comprehension, including social concept representation and retrieval, domain-general semantic integration, and domain-specific integration of social semantic contents. Using fMRI, we investigated the neural correlates of these three cognitive components by exploring how discourse topic (social/nonsocial) and discourse processing period (ending/beginning) modulate brain activation in a discourse comprehension (and also ToM) task. Different sets of brain areas showed sensitivity to discourse topic, discourse processing period, and the interaction between them, respectively. The most novel finding was that the right temporoparietal junction and middle temporal gyrus showed sensitivity to discourse processing period only during social discourse comprehension, indicating that they selectively contribute to domain-specific semantic integration. Our finding indicates how different domains of semantic information are processed and integrated in the brain and provides new insights into the neural correlates of ToM and discourse comprehension.

Keywords Mentalizing · Discourse processing · Social concept · Semantic integration · fMRI

Introduction

Theory of mind (ToM) and discourse comprehension are both important cognitive functions of the human brain. ToM refers to the cognitive processes that attribute independent mental states to self and others in order to predict and explain behavior (Premack & Woodruff, 1978). Discourse comprehension refers to the cognitive processes that create a coherent representation of the meanings of a discourse, which is characterized by the process of integrating the current semantic contents with prior sentential contexts (van Berkum, Zwisterlood, Hagoort, & Brown, 2003). Although ToM and discourse comprehension have been investigated separately, recent neuroimaging studies have indicated that they involve very similar brain regions (Ferstl, Neumann, Bogler, & von Cramon, 2008; Mar, 2011; Xu, Kemeny, Park, Frattali, & Braun, 2005). Mar (2011) conducted a meta-analysis of previous neuroimaging studies on ToM and discourse comprehension. The results showed that the ToM network overlaps with the discourse comprehension network in several brain regions, including the medial prefrontal cortex (MPFC), the temporoparietal junction (TPJ), the anterior temporal lobe (ATL), and the inferior frontal gyrus (IFG).

Why do ToM and discourse comprehension recruit similar brain regions? One natural interpretation is that these brain regions support cognitive processes that are necessarily or frequently involved in both ToM and discourse comprehension (Mar, 2011). To explore this possibility, we decompose ToM and discourse processing along two dimensions. The first dimension is semantic domain. Specifically, we focus on the dichotomy of "social" versus "nonsocial" because

---

Electronic supplementary material The online version of this article (https://doi.org/10.3758/s13415-018-0568-6) contains supplementary material, which is available to authorized users.

✉ Nan Lin
linn@psych.ac.cn

[1] CAS Key Laboratory of Behavioral Science, Institute of Psychology, Beijing 100101, China

[2] Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China

[3] National Laboratory of Pattern Recognition, CASIA, Beijing 100190, China

[4] University of Chinese Academy of Sciences, Beijing 100049, China

[5] School of Business Administration, Zhengzhou University of Aeronautics, Zhengzhou 450046, China

ToM has been proposed as a domain-specific ability of social cognition (Leslie, 1994; Saxe, 2010) and because both ToM and discourse comprehension recruit large areas of the social cognitive network (Mar, 2011). The second dimension is processing type. On this dimension, we focus on the dichotomy of "processing of single concept" versus "processing of combination of concepts.. These two types of processing correspond to two basic semantic processing components that have been specified in several theoretical models (e.g., Hagoort, 2013; Jung-Beeman, 2005): the first component retrieves information from conceptual system and the second component constructs combinatorial representations that are not already available in memory. For the ease of description, we will refer to the first component as "concept representation and retrieval" and the second component as "semantic integration."[1]

According to the two above dimensions, the common neural correlates of ToM and discourse comprehension might be associated with three cognitive components. The first cognitive component is social concept representation and retrieval. Neuroimaging studies have indicated that the representation and retrieval of social concepts may be selectively supported by the anterior superior temporal sulcus (ASTS) and possibly also by the MPFC, TPJ, posterior cingulate gyrus (PC), and precuneus (Lin, Bi, Zhao, Luo, & Li, 2015; Lin et al., 2017; Mason, Banfield, & Macrae, 2004; Mitchell, Heatherton, & Macrae, 2002; Olson, McCoy, Klobusicky, & Ross, 2013; Rice, Ralph, & Hoffman, 2015; Ross & Olson, 2010; Zahn et al., 2007). Zahn et al. (2007) found that the activation level of the ASTS is modulated by the richness of social semantic information and thus proposed that the ASTS represents social concepts. Lin et al. (2015) found that during a verb comprehension task, social action verbs (e.g., "embrace") evoked stronger activation than private action verbs (e.g., "walk") and nonhuman verbs (e.g., "burn") in the ASTS, TPJ, MPFC, PC and precuneus, indicating that these regions may support semantic information associated with social interactions. With respect to ToM processing, thinking about other people's mental states necessarily involves processing of some basic social concepts such as "belief," "desire," "self," and "others." Therefore, social concept representation and retrieval is a necessary component of ToM processing. In discourse comprehension, the involvement of concept representation and retrieval is modulated by the discourse topic. Understanding discourses describing social events necessarily requires processing of social concepts whereas understanding discourses describing nonsocial (e.g., physical or chemistry) events may not. In fact, the stimuli in studies of discourse comprehension often contain descriptions of the mental states or behaviors of the characters (Mar, 2011; Mason & Just, 2009). Therefore, social concept representation and retrieval is a frequently-employed process in studies of discourse comprehension.

The second cognitive component is domain-general semantic integration, through which one can combine simple semantic contents into more complex representations. It is assumed that the engagement of the semantic integration system is sensitive to the richness of contexts and the congruence between contexts and current semantic contents (Hagoort & Indefrey, 2014). For example, the effect of domain-general semantic integration can be examined by comparing the activation evoked by a sentence with that evoked by a word list (Humphries, Binder, Medler, & Liebenthal, 2007), by comparing the activation evoked by a congruent sentence pair with that evoked by a semantically unrelated sentence pair (Ferstl & von Cramon, 2001), or by comparing the activation evoked by the discourse endings whose contextual information is rich with that evoked by the discourse beginnings that have no contextual information (Xu et al., 2005). Neuroimaging studies have indicated that several brain areas may contribute to domain-general semantic integration, including the IFG, TPJ, ATL, MPFC, and middle temporal gyrus (MTG) (Bemis & Pylkkanen, 2011, 2013; Hagoort & Indefrey, 2014; Jung-Beeman, 2005; Price, Bonner, Peelle, & Grossman, 2015; Zhu et al., 2009). Domain-general semantic integration is an essential component of discourse comprehension (Ferstl et al., 2008). It is also necessary for the completion of most ToM tasks because ToM tasks typically require event comprehension, and domain-general semantic integration plays an important role in constructing a coherent representation of an event (Sitnikova, Holcomb, Kiyonaga, & Kuperberg, 2008).

The third cognitive component is domain-specific integration of social semantic contents, through which one can construct domain-specific compositional semantic representations of social events. Similar to social concept representation and retrieval, domain-specific semantic integration should be required for all ToM tasks and for comprehension of discourses describing social events. The existence of domain-specific semantic integration mechanisms has been indicated by developmental psychology studies. Leslie and colleagues proposed that the development of ToM ability relies on a domain-specific mechanism, called "theory-of-mind mechanism" (ToMM) (Leslie, 1994; Leslie, Friedman, & German, 2004). They proposed that ToMM integrates four aspects of semantic information into an agent-centered meta-representation, which includes an "agent" (e.g., "mother"), an "attitude" describing an informational relation (e.g., "believes"), an "anchor" describing an aspect of real situation (e.g., "speaks to the banana"), and a "description" describing an "imagery" situation (e.g., "it is a telephone"). According to this hypothesis,

---

[1] In some articles, the term "semantic integration" specifically refers to the process of integrating word meaning into semantic context (e.g., van Berkum, Hagoort, & Brown, 1999). In the other studies, this term has a broader sense, referring to all types of combinatorial semantic processes, such as discourse-level integration (St George, Kutas, Martinez, & Sereno, 1999), phrase-level conceptual combination (Baron & Osherson, 2011), and cross-modal information integration (Loftus, Miller, & Burns, 1978). Our usage of the term "semantic integration" is based on its broad sense but not the narrow one.

ToMM should be selectively activated when one tries to integrate these four aspects of semantic information. Several neuroimaging studies have indicated that the semantic representation system is organized in a domain-specific manner (for comprehensive reviews, see Mahon & Caramazza, 2009; Martin, 2007). However, to our knowledge, no neuroimaging study has strictly examined the domain-specificity of the semantic integration system. Although a large body of studies observed that the ASTS, TPJ, MPFC, PC, and precuneus showed stronger activation in ToM tasks than in non-ToM tasks (Schurz, Radua, Aichhorn, Richlan, & Perner, 2014), this observation could be explained either as social concept representation/retrieval or as social semantic integration. On the basis of the existing evidence, the right TPJ is the most plausible candidate brain region that supports domain-specific semantic integration. In a series of studies, Saxe and colleagues found that the activation of the right TPJ is highly specific to descriptions about people's thoughts (Saxe, 2010; Saxe & Kanwisher, 2003; Saxe & Wexler, 2005). Saxe and Wexler (2005) further found that the right TPJ is sensitive not only to the domain of semantic contents but also to the congruence of the social background and thought of an agent. The activation level of the right TPJ was higher when the agent's background and thought were incongruent than when the background and thought were congruent. Therefore, the right TPJ showed not only the semantic domain effect but also the semantic integration effect. However, because Saxe and Wexler (2005) did not examine whether the same right TPJ region also contributes to the semantic integration of nonsocial semantic contents, it remains unclear whether this region is specific to social semantic integration. In addition, their analysis was restricted to the bilateral TPJ, MPFC, and PC, leaving it unclear whether there are other areas contributing to social semantic integration.

In the present fMRI study, we differentiated between the brain activations associated with the three abovementioned cognitive components during a discourse comprehension (and also ToM) task by manipulating two factors—discourse topic (social/nonsocial) and discourse processing period (ending/beginning). We examined the effect of discourse processing period because this effect is tightly associated with semantic integration processes in discourse comprehension and has been observed in most areas that are recruited by both ToM and discourse processing (Xu et al., 2005). Our assumptions were as follows: brain regions specific to social concept representation and retrieval should be sensitive to discourse topic but not to discourse processing period; brain regions that are sensitive to discourse processing period in both topics of discourse are associated with domain-general semantic integration; brain regions that are sensitive to discourse processing period in social discourses but not in nonsocial ones are associated with domain-specific integration of social semantic contents.

Materials and methods

Participants

A total of 39 healthy undergraduate and graduate students (18 women, 21 men) participated in the fMRI experiment. The average age of the participants was 22.2 years (SD = 2.7 years). All participants were right-handed and were native Chinese speakers. The participants neither suffered from psychiatric or neurological disorders nor had ever sustained a head injury. Prior to the experiment, each participant read and signed an informed consent form issued by the Institutional Review Board of the Magnetic Resonance Imaging Research Center, the Institute of Psychology of the Chinese Academy of Science.

Materials and procedure

The materials were adapted from a publicly available false-belief localizer (Dodell-Feder et al., 2011). They included ten false-belief discourses and ten false-picture discourses, all of which described an outdated representation (a false belief or a false picture). It has been well demonstrated that the classic ToM network can easily be localized by contrasting these two sets of discourses (Dodell-Feder et al., 2011; Jacoby, Bruneau, Koster-Hale, & Saxe, 2016; Spunt & Adolphs, 2014). The effect of semantic integration was examined by manipulating the extent to which the comprehension of the current sentence is influenced by its preceding contextual semantic information. To maximize the effect, we chose the ending and beginning sentences of discourses as the stimuli of interest (Xu et al., 2005). Therefore, the beginning and ending sentences of the two topics of discourse fits a 2 × 2 design with the factors discourse topic ("social vs. nonsocial" as represented by "false belief vs. false picture") and discourse processing period (ending vs. beginning).

We translated the discourses into Chinese and made a few modifications to adapt to the cultural differences (e.g., changing the English names into Chinese names) and to match the length and sentence number between the two topics of discourse. Examples of our stimuli and their translations are shown in Table 1 and a complete list of our stimuli and their translations are shown in Table S1 in the Supplementary Material. The characteristics of our stimuli are shown in Table 2. All modified discourses consisted of two or three sentences. For both topics of discourse, the average sentence number was 2.7 per discourse. The length (character number) of the beginning, middle (if applicable), and ending sentences and that of the whole discourse were matched between the two topics of discourse (*t*s [18] < 1.320, *p* > .204). The two topics of discourse have no significant differences in character stroke number (*t* [439] = 1.85, *p* = .064) or word frequency (*t* [522] <

Table 1 Examples of the experimental stimuli

	Stimulus	Translation
Social Discourse (BELIEF Item)		
Sentence 1	李明和小芳在房间里找不到钥匙。	Li Ming and Xiao Fang searched the house for their keys with no luck.
Sentence 2	小芳就到外面车里去找了。	Then Xiao Fang went outside to look in the car.
Sentence 3	她出去后, 明发现钥匙在沙发后面。	Suddenly Li Ming noticed the keys behind the sofa.
Question	小芳回来的时候, 明不知道钥匙在哪。	*By the time Xiao Fang comes in, Li Ming doesn't know where the keys are.*
Nonsocial Discourse (PICTURE Item)		
Sentence 1	这幅1885年的名画描绘了河的南岸的景色。	Sargent famously painted the south bank of the river in 1885.
Sentence 2	1910年建成了一座水坝, 没了整个流域, 灭了古老的森林。	In 1910 a huge dam was built, flooding out the whole river basin, killing the old forests.
Sentence 3	现在整片地区都在水下。	Now the whole area is under water.
Question	在画中, 的南岸树木繁茂。	*In the painting the south bank of the river is wooded.*

1; word frequency was obtained from a corpus with news articles from The People's Daily, which has 23 million words).

During the scan, each discourse was presented for 10 s, followed by a true/false statement (presented for 4 s) and a 12-s fixation. Different from the original version of false-belief localizer (Dodell-Feder et al., 2011), the discourses were not presented as a whole but sentence by sentence. We kept the total presentation time of each discourse at 10 s to make our results comparable with the original version of false-belief localizer. Therefore, within a discourse, the presentation time of a sentence (ranging from 1.6 s to 5.6 s) was linearly dependent on the quotient of its length divided by the discourse length. All discourses were presented within a single run lasting 8 min 50 s, with the first 10 s of the run being a fixation. Presentation orders of the discourses were counterbalanced across participants.

Acquisition and analysis of magnetic resonance imaging data

Structural and functional data were collected using a GE Discovery MR750 3 T scanner at the Magnetic Resonance Imaging Research Center, Institute of Psychology of the Chinese Academy of Science. Functional blood-oxygenation-level-dependent data were obtained in 3.0-mm isotropic voxels (TR = 2 s; TE = 30 ms) in 42 near-axial slices. After the task-fMRI scan, T1-weighted structural images were collected in 176 sagittal slices with 1.0-mm isotropic voxels.

All fMRI data were preprocessed using the Statistical Parametric Mapping software (SPM8; http://www.fil.ion.ucl.ac.uk/spm/) and the advanced edition of DPARSF V2.3 (Yan & Zang, 2010). The first five volumes (10 s) of each functional run were discarded for signal equilibrium. Slicing timing and 3-D head motion correction were then performed, and a mean functional image was obtained for each participant. The structural image of each participant was coregistered to the mean functional image and subsequently segmented using the unified segmentation VBM module (Ashburner & Friston, 2005) implemented in DPARSFA. The parameters obtained during segmentation were used to normalize the functional images of each participant onto the Montreal Neurological Institute space. The functional images were then spatially smoothed using a 6-mm FWHM Gaussian kernel.

Statistical analyses on the task fMRI data were conducted according to two-level, mixed-effects models implemented in SPM8. At the first level, a general linear model was applied to explore the fixed effect of each subject. The beginning, middle, and ending sentences and the following statements of the

Table 2 Characteristics of the experimental stimuli

	Discourse length (sentences)	Discourse length (characters)	Sentence length			Word frequency (per million)	Visual complexity of characters (stroke number)
			Beginning (characters)	Middle (characters)	Ending (characters)		
Social discourse	2.7 ± 0.5	49.0 ± 6.3	17.4 ± 4.7	16.9 ± 6.3	19.8 ± 6.6	3278 ± 9506	7.2 ± 3.3
Nonsocial discourse	2.7 ± 0.5	49.3 ± 6.9	19.9 ± 3.7	18.7 ± 7.3	17.5 ± 5.4	4083 ± 10886	6.8 ± 3.3
Effect size for the difference between social and nonsocial discourses (Cohen's *d*)	0	0.05	0.59	0.27	0.38	0.08	0.13

two topics of discourse were set as eight covariates and were all modeled as epoch-related responses according to their onsets and durations. Six head motion parameters, which were obtained by head motion correction, were included as nuisance regressors. In addition, the length and presentation speed (sentence length/presentation time) of each sentence were included as two additional nuisance covariates, which were constructed by modulating the amplitude of the predicted neural response for each sentence by the demeaned (normalized by removing mean value) sentence length and presentation speed. Time series data were subjected to a high-pass filter (128 Hz). After the estimation of model parameters, subject-specific statistical maps were generated. There were four conditions of interest, which correspond to the covariates of the beginning and ending sentences of the two topics of discourse. Contrasts between the four conditions and fixation were created and computed for every subject and inputted into a second-level, random-effects analysis, in which a flexible factorial design was applied to accommodate a 2 × 2 within-subject design. The two main effects and the interaction effect were examined. A whole-brain conjunction analysis (Nichols, Brett, Andersson, Wager, & Poline, 2005) was further conducted to locate areas that were sensitive to discourse topic across beginning and ending sentences ([beginning: social > nonsocial] ∩ [ending: social > nonsocial]) and those that showed stronger activation in discourse endings compared to beginnings across both topics of discourse ([social: ending > beginning] ∩ [nonsocial: ending > beginning]). For single-contrast and conjunction analyses, the false positive rate of a given contrast was controlled at $\alpha < .05$ using voxel-level FWE correction implemented in SPM8, combined with a cluster threshold of 10 voxels. The peak voxels were localized by using xjView toolbox (http://www.alivelearn.net/xjview). The results were shown using the Brainnet Viewer software (Xia, Wang, & He, 2013).

Results

The main effect of discourse topic (social > nonsocial) was observed in all classic regions of the ToM network, including the bilateral MPFC, TPJ, ASTS, precuneus, and PC; in addition, it was also observed in the bilateral IFG and precentral gyrus (Table 3 and Fig. 1). We further conducted a more stringent conjunction analysis to obtain brain areas showing consistent discourse topic effect across two single contrasts ([beginning: social > nonsocial] ∩ [ending: social > nonsocial]). Four clusters located in the bilateral ASTS and the ventral TPJ showed the conjunction effect. These four regions showed very reliable sensitivity to discourse topic. To further indicate their functional specificity, we conducted region-of-interest (ROI) analyses to examine whether they were also sensitive to discourse processing period (Table S2 in the Supplementary Material). As we have mentioned, the discourse topic effect may reflect either social concept representation/retrieval or social semantic integration. The brain regions specific to social concept representation and/or retrieval should be sensitive to discourse topic and insensitive to discourse processing period. In comparison to their activation levels in discourse beginnings, the activation levels of the bilateral ASTS clusters were not enhanced in social discourse endings and even declined in nonsocial discourse endings. This resulting pattern indicates that the bilateral ASTS are not sensitive to semantic integration in discourse processing. Therefore, as assumed by previous studies (Zahn et al., 2007), the bilateral ASTS may support the representation and/or retrieval of social concepts. The decline of its activation level in nonsocial discourse comprehension could be explained as a top-down modulation effect of semantic access. Participants were more aware of the topic of the discourses in discourse endings and thus, they accessed less social semantic information in nonsocial discourse endings than in beginnings. In contrast, the bilateral ventral TPJ clusters showed stronger activation in social discourse endings compared to beginnings but did not show such activation differences in nonsocial discourses. Therefore, the bilateral ventral TPJ clusters may contribute to social semantic integration and the discourse topic effect observed in these regions may reflect social semantic integration rather than social concept representation or retrieval.

The main effect of discourse processing period (ending > beginning) was observed in the bilateral TPJ, IFG, middle frontal gyrus (MFG), insula, MTG, MPFC, supplementary motor area (SMA), and the left precuneus (Table 3 and Fig. 1). In a more stringent conjunction analysis, the discourse processing period effect ([social: ending > beginning] ∩ [nonsocial: ending > beginning]) was observed in the bilateral IFG, MFG, and insula, as well as the left TPJ and MPFC (and adjacent SMA). We further inspected the activation patterns of these clusters to examine whether these regions also showed activation differences between the two topics of discourse. We found that nine of the ten clusters showed the discourse topic effect (social > nonsocial) in ending sentences and three of them showed the discourse topic effect in beginning sentences (Table S3). This resulting pattern indicates that the comprehension of false-belief (social) discourses may involve more domain-general semantic integration processing than the comprehension of false-picture discourses, especially in the endings of discourses.

The brain areas showing the "discourse topic × discourse processing period" interaction effect are shown in Fig. 2. The interaction effect was observed in the bilateral TPJ, MTG, MPFC, the right anterior cingulate (AC), the left angular gyrus, and the left precuneus. The peak coordinates and cluster sizes of these clusters are listed in Table 3. To inspect the activation patterns of the brain regions showing the interaction effect, we conducted a ROI analysis. To avoid the double dipping problem, we conducted a split-half analysis in which the participants

Table 3 Results of whole-brain fMRI data analysis

Contrast	Anatomical region	Cluster size (voxels)	MNI coordinates of peak voxel (x, y, z)			Peak t value
Main effect of discourse topic: Social > Nonsocial	Right superior temporal sulcus and temporoparietal junction	949	57	3	-21	12.88
	Left superior temporal sulcus and inferior frontal gyrus	455	-51	12	-30	11.88
	Left temporoparietal junction	467	-42	-60	18	11.59
	Bilateral medial frontal gyrus and superior frontal gyrus	596	-9	54	33	9.65
	Bilateral precuneus and posterior cingulate gyrus	169	-6	-54	39	8.86
	Right inferior frontal gyrus	135	57	24	3	7.92
	Right superior frontal gyrus	38	24	24	45	7.38
	Right precentral gyrus	41	39	6	42	6.56
	Right medial frontal gyrus	30	6	51	-12	6.46
	Left precentral gyrus	40	-33	3	48	6.22
	Right inferior frontal gyrus	10	42	21	21	5.62
Main effect of semantic integration demands: High > Low	Left temporoparietal junction	431	-54	-54	24	11.55
	Left inferior frontal gyrus	243	-48	42	-6	10.74
	Left middle frontal gyrus	373	-48	21	42	10.54
	Bilateral medial frontal gyrus and supplementary motor area	304	-12	21	63	10.41
	Right middle frontal gyrus	511	48	24	33	10.08
	Right temporoparietal junction	245	51	-51	27	9.75
	Right insula	61	30	24	-3	9.74
	Left insula	44	-30	21	-6	9.64
	Left inferior frontal gyrus	85	-54	18	6	9.46
	Left middle temporal gyrus	141	-57	-27	-6	8.94
	Right inferior frontal gyrus	47	48	42	-12	7.55
	Right middle temporal gyrus	96	57	-30	-9	7.03
	Left precuneus	10	-6	-60	45	6.17
“.iscourse topic × semantic integration demands” interaction	Right temporoparietal junction	213	54	-51	24	8.74
	Right middle temporal gyrus	158	60	-18	-12	8.51
	Left middle temporal gyrus	56	-60	-21	-12	7.42
	Left temporoparietal junction	136	-51	-54	24	6.83
	Left medial frontal gyrus	35	-9	45	27	6.26
	Right medial frontal gyrus and anterior cingulate	31	9	48	12	6.14
	Left precuneus	36	-6	-54	42	5.85
	Left angular gyrus	20	-42	-69	36	5.49

were separated into two groups, with Group 1 containing 20 participants and Group 2 containing 19 participants. In the analysis of Group 1, the ROIs were defined using the data of Group 2; in the analysis of Group 2, the ROIs were defined using the data of Group 1. In the analysis to define the ROIs, the false positive rate was controlled at $\alpha < .05$ using cluster-level FWE correction implemented in SPM8 (with the individual voxel threshold probability setting of $p < .001$). In both groups, the interaction effect was observed in the bilateral TPJ, MTG, MPFC, and AC (Fig. 2). The precuneus cluster was only observed in Group 2 and thus, it was not included in the ROI analysis. The results of the ROI analysis are shown in Fig. 2. Across the two groups, consistent result patterns were observed in the left TPJ and in the right MTG and TPJ. The right MTG and TPJ showed the discourse processing period effect (ending > beginning) only in social discourses, indicating that these regions may be specific to social semantic integration. The left TPJ showed the discourse processing period effect in both topics of discourse, indicating that this region may contribute to both social and nonsocial semantic integration.

Discussion

We explored the brain activations associated with three cognitive components involved in ToM and discourse comprehension,

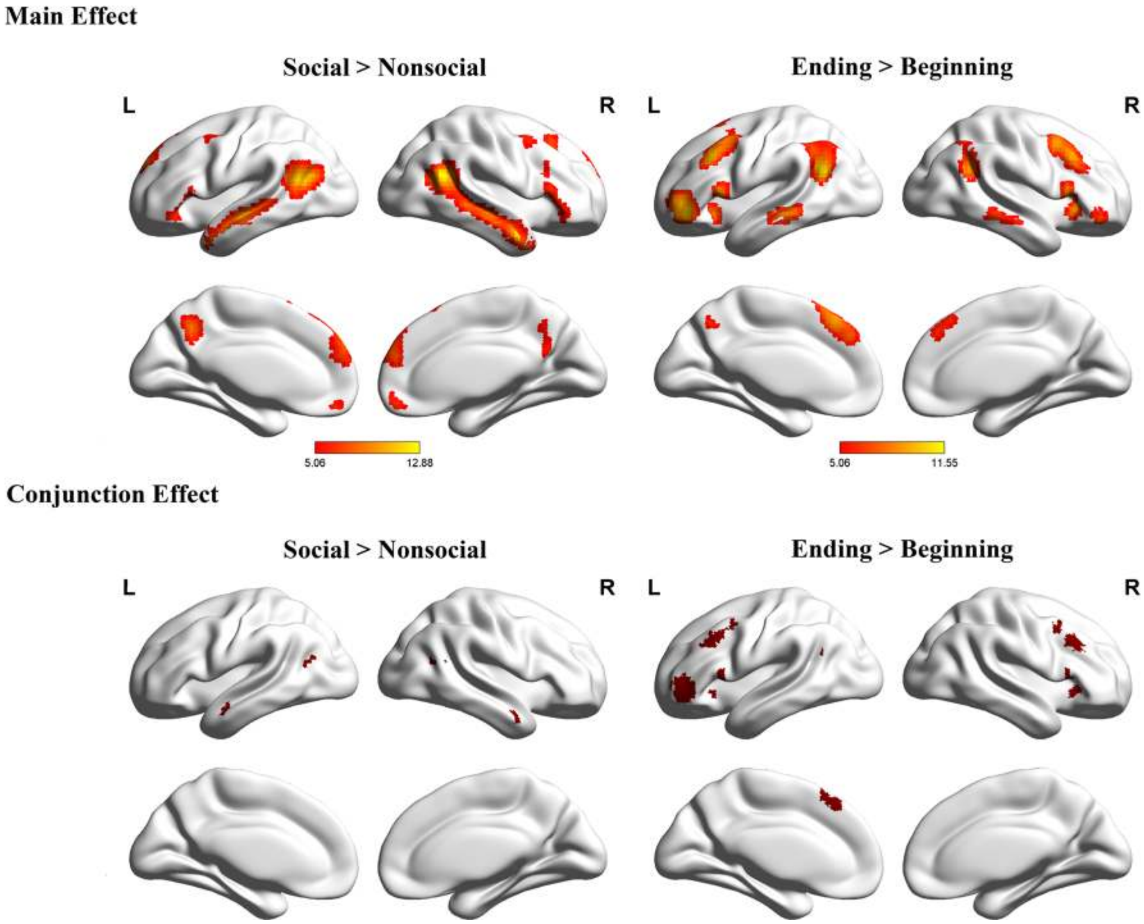

Fig. 1 Brain regions showing the effects of discourse topic (social > nonsocial) and discourse processing period (ending > beginning) in whole-brain functional MRI (fMRI) data analysis

which are social concept representation and retrieval, domain-general semantic integration, and domain-specific integration of social semantic contents. The brain activity associated with the three cognitive processes was investigated by examining the effects of discourse topic, discourse processing period, and the interaction between the two in a discourse comprehension task. The results showed that the three cognitive processes were supported by three different groups of brain regions.

The most novel finding of our study pertains to the interaction effect of the discourse topic and discourse processing period. This effect was observed in the bilateral TPJ, MTG, MPFC, and AC, in the analysis of the whole participant group as well as in the split-half analysis. A further ROI analysis was conducted by using a split-half approach. Across the two groups of participants, consistent activation patterns were found in the left TPJ and in the right MTG and TPJ. The right MTG and TPJ showed the discourse processing period effect only in social discourses, indicating that they may specifically contribute to social semantic integration. Although previous studies have indicated the domain-specificity of the right TPJ by comparing different topics of discourses (Dodell-Feder et al., 2011; Saxe & Kanwisher, 2003), we for the first time reported the interaction effect of the social domain and discourse processing period in this area. More interestingly, a similar interaction effect was observed in the right MTG. The involvement of this area in social semantic integration has not been reported in previous studies and the specific function of this area should be further investigated in future studies. The left TPJ showed the discourse processing period effect (ending > beginning) across the two topics of discourse, indicating that it may play a role in domain-general semantic integration. This speculation is consistent with the findings of previous lesion studies, which have indicated that the left TPJ is involved in not only the false belief task but also the false picture task (Apperly, Samson, Chiavarino, Bickerton, & Humphreys, 2007). However, it should be noted that the left TPJ also showed strong interaction effect. It showed a much stronger semantic integration effect in social discourses than in nonsocial ones (see Fig. 2). Therefore, we propose that this region may serve as an interface between the domain-general and domain-specific semantic integration systems. The other regions showing the interaction effect (the left MTG, the MPFC, and the AC) did not show consistent result patterns in the two groups of participants. This may be due to lack of statistical power in the split-half analysis, individual differences in the anatomical distribution of functional regions, or

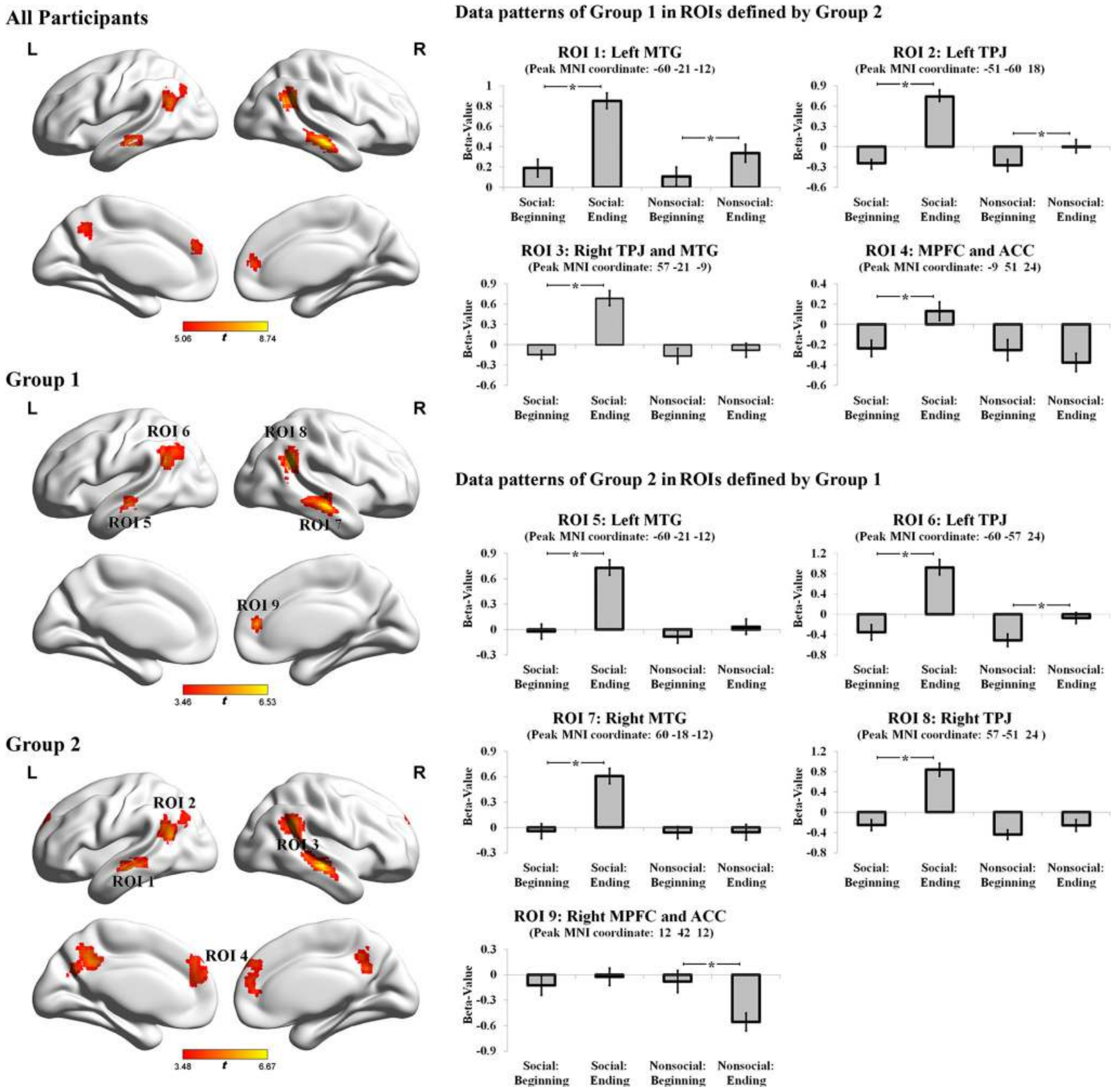

Fig. 2 Brain regions showing the interaction effect of discourse topic and discourse processing period in whole-brain fMRI data analysis and the results of the split-half region of interest (ROI) analysis

individual differences in cognitive strategies employed in discourse reading. Therefore, the functions of these regions could not be clearly inferred by the data from the present study.

We also investigated the main effects of discourse topic and discourse processing period. The main effect of discourse topic replicated the classic findings of the ToM network (Saxe & Kanwisher, 2003). Brain activation was observed in the bilateral MPFC, TPJ, ASTS, precuneus, and PC. Brain areas contributing to social concept representation and retrieval should be sensitive to social semantic contents but not discourse processing period. Thus, we conducted conjunction analysis to strictly localize the brain areas showing consistent discourse topic effect across discourse beginnings and endings. Conjunction effect was observed in the bilateral ASTS and ventral TPJ. Further ROI analysis showed that the bilateral ASTS, but not the bilateral ventral TPJ, were insensitive to discourse processing period in social discourses. Therefore, as has been proposed by previous studies of social concept processing, the bilateral ASTS may support social concept representation/retrieval (Zahn et al., 2007)[2].

---

[2] Some studies have reported semantic integration effects in the ATL (Baron & Osherson, 2011; Bemis & Pylkkanen, 2011; Humphries, Binder, Medler, & Liebenthal, 2006). However, there is a lack of evidence whether the ASTS regions supporting social semantic processing and the ATL regions supporting semantic integration overlap with each other.

The main effect of discourse processing period was observed in the bilateral TPJ, IFG, MFG, insula, MTG, MPFC, SMA, and left precuneus. These observations are highly similar to the findings of the previous studies that have examined the semantic integration effect in discourse comprehension (Ferstl & von Cramon, 2002; Xu et al., 2005). Brain areas that contribute to domain-general semantic integration should be sensitive to discourse processing period in both social and nonsocial discourses. Thus, we conducted conjunction analysis to locate the brain areas showing consistent discourse processing period effect across the two topics of discourse. The conjunction effect was observed in large areas of the bilateral lateral frontal cortex and in the left TPJ and superior part of the MPFC (and adjacent SMA). It should be noted that considering that all discourses used in the present study described outdated representations (false beliefs or false pictures), the comprehension of discourse endings may additionally require inhibitory control to dissociate the true representations from the false representations (Carlson & Moses, 2001). Based on existing knowledge from literature, we speculated that the observed activation of the bilateral IFG and the left TPJ should be associated with semantic integration (Bemis & Pylkkanen, 2013; Hagoort & Indefrey, 2014; Humphries et al., 2007; Price et al., 2015; Zhu et al., 2009), whereas the effect observed in the bilateral MFG and the left superior MPFC and adjacent SMA should be associated with inhibitory control (Cole & Schneider, 2007).

It is important to note that almost all regions showing the discourse processing period effect in the conjunction analysis also showed relatively stronger activation in response to social discourse endings than to nonsocial discourse endings in ROI analysis. This observation is consistent with the finding of a previous study (Ferstl & von Cramon, 2002) that the ToM effects and the semantic coherence effects (which mainly reflect semantic integration processes) overlap with each other in several brain regions of the ToM network. A related observation was that the main effect of discourse topic was observed in the lateral frontal cortex, which largely overlapped with those areas showing the discourse processing period effect. The discourse topic effect in the lateral frontal cortex is not a consistent observation across studies using the false belief task (Schurz et al., 2014), although it has been observed in some studies (Dodell-Feder et al., 2011). Our findings indicate that such effects may be associated with the unbalanced semantic integration processes between conditions. Thus, studies employing stimuli with rich semantic contents (e.g., discourses, cartoons, and movies) should more carefully consider this potential confounding effect in future.

The present study has indicated a dissociation between domain-general and domain-specific semantic integration systems. Thus, a further interesting question is how the two systems interact with each other and how semantic information is transferred from one to the other. Interestingly, a recent study (Lavoie, Vistoli, Sutliff, Jackson, & Achim, 2016) indicated that the domain-general semantic integration system, but not the domain-specific semantic integration system, may be involved in a delayed stage of discourse comprehension. In Lavoie et al. (2016), participants read social and nonsocial discourses consisting of two sentences. Then they saw two statements about the discourse and were asked to select the correct one. For both types of discourses, the first sentence could induce an inference about the forthcoming choices and the second sentence provided either a congruent cue that confirmed the initial inference or an incongruent cue that led to an adjustment of the initial inference. Brain activations during the reading of the second sentence and those during the response period were analyzed. Discourse topic effect (social > nonsocial) was observed in both periods, and their distributions were consistent with the typical distributions of the classic ToM network. The congruency effect (incongruent > congruent) was observed only in the response period, and their distributions were highly similar to those of the main effect of domain-general semantic integration observed in the present study. Lavoie et al. (2016) interpreted the observed congruency effect during the response period as reflecting a delayed stage of semantic integration, which may be induced by task demands. Interestingly, at this delayed stage, although the effects of discourse topic and semantic integration were both observed, their interaction was not significant. Combining our findings with the observation of Lavoie et al. (2016), it seems that domain-specific integration of social semantic contents may occur only in the on-line stage of language comprehension (which could not be detected in Lavoie et al.'s study because no on-line semantic integration effect was induced by their manipulation). Afterwards, the obtained social semantic representations would be transferred into the domain-general semantic system for delayed retrieval and manipulations.

Finally, it should be noted that the two dimensions that we used to decompose ToM and discourse processing, i.e. "social versus nonsocial" and "processing of single concept versus processing of combination of concepts," are very broad dimensions. Our findings may associate with some more detailed dimensions that we cannot specify in the present study. The semantic dimension of sociality can be decomposed into several sub-dimensions, such as warmth, competence, communal sharing, authority ranking, equality matching, and market pricing (Fiske, 1992; Fiske et al., 2002). Previous studies have also indicated that sociality has natural correlations with several other semantic dimensions such as valence, arousal, and imageability (Binder et al., 2016; Tamir, Thornton, Contreras, & Mitchell, 2016; Troche, Crutch, & Reilly, 2014). Therefore, the observed effect of discourse topic may reflect the mixture of some or all of these detailed semantic dimensions and their contributions need to be further classified in future studies. Similarly, the effect of discourse processing period may also be associated with several detailed integration processes and factors. At the textbase level (Kintsch & van Dijk, 1978), the processing of the ending sentence involves rich semantic contextual information while the processing of the beginning sentence doesn't. At the situation-model level (van Dijk & Kintsch, 1983), at the beginning period a situation model is set up and at the ending period the contents of individual sentences are woven into a coherent and global representation (Ferstl, 2010; Yarkoni, Speer, & Zacks, 2008). With respect to processing time scale (Lerner, Honey, Silbert, & Hasson, 2011), the temporal receptive window (the length of time during which information input may affect the current processing) is longer at the ending period than at the beginning period. In addition, for the specific false-belief and false-picture stimuli used in our experiment, the ending period involves more complex inference than the beginning period does and the false-belief stories involve additional mental-state inference in comparison with the false-picture stories (Saxe & Kanwisher, 2003). The relationships between these detailed integration processes and factors and our findings should also be clarified in future studies.

In summary, the present study indicates the presence of fine-grained functional divisions of the brain network involved in ToM and discourse comprehension. The bilateral ASTS support social concept representation and/or retrieval, the bilateral lateral frontal cortex and the superior part of MPFC support domain-general semantic integration and perhaps also other related domain-general processes like inhibitory control, the right TPJ and MTG support social semantic integration, and the left TPJ serves as an interface for social semantic integration and domain-general semantic integration. These findings provide a new schema to understand the functional organization of the human semantic processing system and provide new insights into the neural correlates of ToM and discourse comprehension.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant numbers: 31300842, 31400968, 31300859, and 31571125) and the Beijing Advanced Innovation Center for Imaging Technology (BAICIT-2016018).

References

Apperly, I. A., Samson, D., Chiavarino, C., Bickerton, W. L., & Humphreys, G. W. (2007). Testing the domain-specificity of a theory of mind deficit in brain-injured patients: Evidence for consistent performance on non-verbal, "reality-unknown" false belief and false photograph tasks. *Cognition, 103*(2), 300-321.

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *Neuroimage, 26*(3), 839-851.

Baron, S. G., & Osherson, D. (2011). Evidence for conceptual combination in the left anterior temporal lobe. *Neuroimage, 55*(4), 1847-1852.

Bemis, D. K., & Pylkkanen, L. (2011). Simple Composition: A Magnetoencephalography Investigation into the Comprehension of Minimal Linguistic Phrases. *Journal of Neuroscience, 31*(8), 2801-2814.

Bemis, D. K., & Pylkkanen, L. (2013). Basic Linguistic Composition Recruits the Left Anterior Temporal Lobe and Left Angular Gyrus During Both Listening and Reading. *Cerebral Cortex, 23*(8), 1859-1873.

Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M., & Desai, R. H. (2016). Toward a brain-based componential semantic representation. *Cognitive neuropsychology, 33*(3-4), 130-174.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032-1053.

Cole, M. W., & Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *Neuroimage, 37*(1), 343-360.

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage, 55*(2), 705-712.

Ferstl, E. C. (2010). Neuroimaging of text comprehension: Where are we now. *Italian Journal of Linguistics, 22*(1), 61-88.

Ferstl, E. C., Neumann, J., Bogler, C., & von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping, 29*(5), 581-593.

Ferstl, E. C., & von Cramon, D. Y. (2001). The role of coherence and cohesion in text comprehension: an event-related fMRI study. *Cognitive Brain Research, 11*(3), 325-340.

Ferstl, E. C., & von Cramon, D. Y. (2002). What does the frontomedian cortex contribute to language processing: coherence or theory of mind? *Neuroimage, 17(3)*, 1599-1612.

Fiske, A. P. (1992). The four elementary forms of sociality: framework for a unified theory of social relations. *Psychological review, 99*(4), 689.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and social psychology, 82*(6), 878.

Hagoort, P. (2013). MUC (memory, unification, control) and beyond. *Frontiers in Psychology, 4*, 416.

Hagoort, P., & Indefrey, P. (2014). The Neurobiology of Language Beyond Single Words. *Annual Review of Neuroscience, 37*, 347-362.

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of cognitive neuroscience, 18*(4), 665-679.

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2007). Time course of semantic processes during sentence comprehension: An fMRI study. *Neuroimage, 36*(3), 924-932.

Jacoby, N., Bruneau, E., Koster-Hale, J., & Saxe, R. (2016). Localizing Pain Matrix and Theory of Mind networks with both verbal and non-verbal stimuli. *Neuroimage, 126*, 39-48.

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences, 9*(11), 512-518.

Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological review, 85*(5), 363.

Lavoie, M. A., Vistoli, D., Sutliff, S., Jackson, P. L., & Achim, A. M. (2016). Social representations and contextual adjustments as two distinct processes of the Theory of Mind brain network: Evidence from the REMICS task. *Cortex, 81*, 176-191.

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *Journal of Neuroscience, 31*(8), 2906-2915.

Leslie, A. M. (1994). Pretending and believing: issues in the theory of ToMM. *Cognition, 50*(1-3), 211-238.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences, 8*(12), 528-533.

Lin, N., Bi, Y. C., Zhao, Y., Luo, C. M., & Li, X. S. (2015). The theory-of-mind network in support of action verb comprehension: Evidence from an fMRI study. *Brain and Language, 141*, 1-10.

Lin, N., Wang, X., Xu, Y., Wang, X., Hua, H., Zhao, Y., & Li, X. (2017). Fine Subdivisions of the Semantic Network Supporting Social and Sensory–Motor Semantic Processing. *Cerebral Cortex*, 1–12. https://doi.org/10.1093/cercor/bhx148.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of experimental psychology: Human learning and memory, 4*(1), 19.

Mahon, B. Z., & Caramazza, A. (2009). Concepts and Categories: A Cognitive Neuropsychological Perspective. *Annual Review of Psychology, 60*, 27-51.

Mar, R. A. (2011). The Neural Bases of Social Cognition and Story Comprehension. *Annual Review of Psychology, 62*, 103-134.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58*, 25-45.

Mason, M. F., Banfield, J. F., & Macrae, C. N. (2004). Thinking about actions: The neural substrates of person knowledge. *Cerebral Cortex, 14*(2), 209-214.

Mason, R. A., & Just, M. A. (2009). The Role of the Theory-of-Mind Cortical Network in the Comprehension of Narratives. *Language and Linguistics Compass, 3*(1), 157-174.

Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences of the United States of America, 99*(23), 15238-15243.

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage, 25*(3), 653-660.

Olson, I. R., McCoy, D., Klobusicky, E., & Ross, L. A. (2013). Social cognition and the anterior temporal lobes: a review and theoretical framework. *Social Cognitive and Affective Neuroscience, 8*(2), 123-133.

Premack, D., & Woodruff, G. (1978). Does the Chimpanzee Have a Theory of Mind. *Behavioral and Brain Sciences, 1*(4), 515-526.

Price, A. R., Bonner, M. F., Peelle, J. E., & Grossman, M. (2015). Converging Evidence for the Neuroanatomic Basis of Combinatorial Semantics in the Angular Gyrus. *Journal of Neuroscience, 35*(7), 3276-3284.

Rice, G. E., Ralph, M. A. L., & Hoffman, P. (2015). The Roles of Left Versus Right Anterior Temporal Lobes in Conceptual Knowledge: An ALE Meta-analysis of 97 Functional Neuroimaging Studies. *Cerebral Cortex, 25*(11), 4374-4391.

Ross, L. A., & Olson, I. R. (2010). Social cognition and the anterior temporal lobes. *Neuroimage, 49*(4), 3452-3462.

Saxe, R. (2010). The right temporo-parietal junction: a specific brain region for thinking about thoughts. In Handbook of Theory of Mind, ed. A Leslie, T German, pp. 1–35. Hove: Psychology Press

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people - The role of the temporo-parietal junction in "theory of mind". *Neuroimage, 19*(4), 1835-1842.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia, 43*(10), 1391-1399.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience and Biobehavioral Reviews, 42*, 9-34.

Sitnikova, T., Holcomb, P. J., Kiyonaga, K. A., & Kuperberg, G. R. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of cognitive neuroscience, 20*(11), 2037-2057.

Spunt, R. P., & Adolphs, R. (2014). Validating the Why/How contrast for functional MRI studies of Theory of Mind. *Neuroimage, 99*, 301-311.

St George, M., Kutas, M., Martinez, A., & Sereno, M. I. (1999). Semantic integration in reading: engagement of the right hemisphere during discourse processing. *Brain, 122*(7), 1317-1325.

Tamir, D. I., Thornton, M. A., Contreras, J. M., & Mitchell, J. P. (2016). Neural evidence that three dimensions organize mental state representation: Rationality, social impact, and valence. *Proceedings of the National Academy of Sciences, 113*(1), 194-199.

Troche, J., Crutch, S., & Reilly, J. (2014). Clustering, hierarchical organization, and the topography of abstract and concrete nouns. *Frontiers in psychology, 5*, 360.

Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of cognitive neuroscience, 11*(6), 657-671.

Van Berkum, J. J. A., Zwisterlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research, 17*(3), 701-718.

Van Dijk, T. A., Kintsch, W., & Van Dijk, T. A. (1983). Strategies of discourse comprehension (pp. 11-12). New York: Academic Press.

Xia, M. R., Wang, J. H., & He, Y. (2013). BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics. *Plos One, 8*(7), e68910.

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: emergent features of word, sentence, and narrative comprehension. *Neuroimage, 25*(3), 1002-1015.

Yan, C. G., & Zang, Y. F. (2010). DPARSF: A MATLAB Toolbox for "Pipeline" Data Analysis of Resting-State fMRI. *Frontiers in Systems Neuroscience, 4*, 13.

Yarkoni, T., Speer, N. K., & Zacks, J. M. (2008). Neural substrates of narrative comprehension and memory. *Neuroimage, 41*(4), 1408-1425.

Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 104*(15), 6430-6435.

Zhu, Z. D., Zhang, J. X., Wang, S. P., Xiao, Z. W., Huang, J., & Chen, H. C. (2009). Involvement of left inferior frontal gyrus in sentence-level semantic integration. *Neuroimage, 47*(2), 756-763.