nature human behaviour

Article

https://doi.org/10.1038/s41562-023-01704-8

A social-semantic working-memory account for two canonical language areas

Received: 1 February 2023

Accepted: 22 August 2023

Published online: 21 September 2023

Check for updates

Guangyao Zhang[1,2,3,4,5], Yangwen Xu[6], Xiuyi Wang[1,2], Jixing Li[7], Weiting Shi[1,2], Yanchao Bi[3,4,5,8] & Nan Lin[1,2]

Language and social cognition are traditionally studied as separate cognitive domains, yet accumulative studies reveal overlapping neural correlates at the left ventral temporoparietal junction (vTPJ) and the left lateral anterior temporal lobe (lATL), which have been attributed to sentence processing and social concept activation. We propose a common cognitive component underlying both effects: social-semantic working memory. We confirmed two key predictions of our hypothesis using functional MRI. First, the left vTPJ and lATL showed sensitivity to sentences only when the sentences conveyed social meaning; second, these regions showed persistent social-semantic-selective activity after the linguistic stimuli disappeared. We additionally found that both regions were sensitive to the socialness of non-linguistic stimuli and were more tightly connected with the social-semantic-processing areas than with the sentence-processing areas. The converging evidence indicates the social-semantic working-memory function of the left vTPJ and lATL and challenges the general-semantic and/or syntactic accounts for the neural activity of these regions.

Language and social cognition are two fundamental abilities of the human species. They are deeply interrelated with each other in cognitive development[1,2], daily communication[3] and evolution[4]. At the brain level, overlaps of regions underlying language and social cognition have been found in the left ventral temporoparietal junction (vTPJ; consisting of the ventral portion of the angular gyrus and its adjacent temporal cortex) and the left lateral anterior temporal lobe (lATL)[5–7]. Understanding the function of these regions will indicate how language and social cognition are associated with each other in the brain.

In the field of social neuroscience, the left vTPJ and lATL have been found to be involved in multiple social cognitive tasks[6,8,9]. Recent studies have indicated that these regions may support a very basic component of social cognition—that is, social concept representation and processing[10–14]. These regions are sensitive to a wide range of social concepts (concepts associated with people and their interactions), including traits (for example, brave[15]), mental states (for example, distrust[16]), stereotypes (for example, women[17]), social backgrounds (for example, having a good salary[18]), social actions (for example, chase[19–21]) and social artefacts (for example, telephone[22]).

In the field of language neuroscience, the left vTPJ and lATL have both been found to be critical for sentence processing. Lesions in these regions can damage sentence comprehension[23]. Neuroimaging studies have found that, in both regions, sentences induce stronger activation than word lists or other unintelligible linguistic stimuli[24–27]. These two regions, along with the inferior frontal gyrus (IFG) and posterior superior temporal sulcus (pSTS), were therefore identified as the "classic high-level language-processing regions"[28] and together form the sentence-processing network[23,29,30].

The traditional explanations for the roles of the left vTPJ and lATL in sentence processing are all based on the encoding and integration of general semantics and/or syntactic information[25,30–32]. For example, Price[32] reviewed 100 functional MRI (fMRI) studies of speech

[1]CAS Key Laboratory of Behavioral Science, Institute of Psychology, Beijing, China. [2]Department of Psychology, University of Chinese Academy of Sciences, Beijing, China. [3]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China. [4]IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China. [5]Beijing Key Laboratory of Brain Imaging and Connectomics, Beijing Normal University, Beijing, China. [6]Center for Mind/Brain Sciences, University of Trento, Trento, Italy. [7]Department of Linguistics and Translation, City University of Hong Kong, Hong Kong SAR, China. [8]Chinese Institute for Brain Research, Beijing, China. e-mail: linn@psych.ac.cn

https://doi.org/10.1038/s41562-023-01704-8

comprehension and production and proposed that in sentence comprehension, the left lATL and angular gyrus (overlapping with the vTPJ) support amodal semantic combination and task-dependent semantic retrieval, respectively. In an influential fMRI study, Pallier et al.[25] found that, in contrast to the left IFG and pSTS, which responded even to abstract syntactic structures containing no content words, the left vTPJ and lATL responded only to meaningful sentences. They therefore proposed that the left vTPJ and lATL may bind syntactic roles to lexico-semantic representations to form high-level representations of semantic constituent structure.

If the language functions of the left vTPJ and lATL are primarily associated with general semantic and/or syntactic processing, it is difficult to understand why these areas are especially sensitive to social semantics. Mellem et al.[7], in addition to replicating the sentence effect discovered by Pallier et al.[25], found that the left lATL is more sensitive to social–emotional semantics than to object semantics. Bzdok et al.[5] identified through meta-analysis that the left vTPJ has been associated with both social-cognitive functions and semantic/syntactic processing. These findings provide direct evidence for the neural correlates common to social cognition and sentence processing in the left vTPJ and lATL.

Here, we propose a hypothesis that integrates the sentence and social-semantic sensitivity of the left vTPJ and lATL into one cognitive function—that is, social-semantic working memory. Our hypothesis is motivated by two views in the literature. First, multiple brain regions ranging from sensory to association cortex can represent particular contents of working memory[33–36]. According to this view, the short-term retention of semantic information can be supported by the temporary activation of long-term memories[37]. Therefore, the left vTPJ and lATL (which are assumed to support social concept representation) can support social-semantic working memory. Second, sentences have more stable and durable semantic representation than fragmented linguistic stimuli, such as word lists. It has been found that when a series of words are presented very rapidly, recall is poor for word lists but near perfect for sentences[38,39]. On the basis of this and related findings, Potter[39,40] proposed that word meaning is activated rapidly, but the initial activation is highly unstable and will be forgotten within a few hundred milliseconds unless incorporated into a structure. The sensitivity of brain regions to sentences may therefore reflect either the semantic/syntactic integration itself or the semantic working memory consolidated by the semantic/syntactic integration. Given that most stimuli used in the previous studies of sentence processing are about people, the activation of the left vTPJ and lATL in these studies may reflect social-semantic working memory.

In this Article, we examined the social-semantic working-memory hypothesis using six fMRI experiments. First, we examined whether the sensitivity of the left vTPJ and lATL to sentences is selectively associated with social-semantic comprehension (Experiments 1 and 2). The social-semantic working-memory hypothesis predicts that the sensitivity of these regions to sentences can be found only if the sentences convey social meaning; the traditional general-semantic and/or syntactic accounts, however, predict that the sensitivity to sentences should be found even when the sentence conveys no social meaning. Second, we examined a key prediction of the social-semantic working-memory hypothesis—namely, that the left vTPJ and lATL would show persistent social-semantic-selective activity even after the stimuli disappeared (demonstrating the key signature property of working memory; Experiments 3 and 4). Finally, we supplemented the above key examinations with two additional experiments, aiming to investigate whether the left vTPJ and lATL were involved in the processing of non-linguistic, high-socialness stimuli (Experiment 5) and whether the left vTPJ and lATL had stronger intrinsic connectivity to the social-semantic-processing areas or to the sentence-processing areas (Experiment 6).

Results

Social meaning drives sentence effects in the left vTPJ and lATL

The aim of Experiments 1 and 2 was to examine whether the sensitivity of the left vTPJ and lATL to sentences is selectively associated with social-semantic comprehension. In both experiments, the participants were asked to read sentences and word lists during fMRI scanning. Following previous neuroimaging studies[25,27,29,31], we measured the sensitivity to sentences by subtracting the neural responses to word lists from those to sentences. For both sentence and word-list stimuli, we manipulated their socialness, leading to four conditions: the high-socialness-sentence (HSS), high-socialness-word-list (HSWL), non-social-sentence (NSS) and non-social-word-list (NSWL) conditions. In both experiments, the stimuli of the HSWL and NSWL conditions were constructed by pseudorandomly combining the constituent words of the HSS and NSS conditions, respectively (Methods).

Experiments 1 and 2 were mainly different in the lengths and structures of the sentences. In Experiment 1, we used short sentences with the noun–verb–noun structure, which has two advantages. First, we can easily control syntactic complexity by consistently using the noun–verb–noun structure in both sentence conditions. Second, we can manipulate the socialness of the stimuli word by word to maximize the social-semantic effect because there are only content words. Therefore, in Experiment 1, for the HSS condition, all constituent words of the sentences have high socialness (for example, '歹徒抢劫商店', meaning '(the) gangsters robbed (the) shops'); by contrast, for the NSS condition, all constituent words of the sentences refer to natural and non-human entities and events (for example, '洪水淹没草原', meaning '(the) flood inundated (the) grassland'). See Fig. 1a,b for sample stimuli and trials.

Despite the advantages of using short sentences, natural sentences often consist of both content and function words and are often longer than three words. We therefore conducted Experiment 2 to examine whether the findings of Experiment 1 could be generalized to sentences with more natural structures and lengths. In Experiment 2, the sentences were approximately eight words long for both conditions. The sentences of the HSS condition were all related to interpersonal interactions or relationships (for example, '班主任要负责管理好班级秩序', meaning 'The head teacher should be responsible for managing the class order'), and the sentences of the NSS condition were all about natural and non-human entities and events (for example, '长期缺水使这里的灌木渐渐枯萎', meaning 'Long-term water shortage has caused the shrubs here to gradually wither'). See Fig. 2a,b for sample stimuli and trials.

The whole-brain analysis in Experiment 1 showed both social-semantic and sentence effects in the left vTPJ and lATL (Supplementary Fig. 1 and Supplementary Table 1), replicating findings in the literature[20,27,41,42]. Importantly, the left TPJ and lATL showed significant interactions between social-semantic and sentence effects (Fig. 1c and Supplementary Table 1). We then examined the sentence effect using the high-socialness and non-social stimuli separately. The comparison between the HSS and HSWL conditions revealed significant sentence effects in the left vTPJ, the left lATL and other canonical regions of the sentence-processing network; however, the comparison between the NSS and NSWL conditions did not reveal any significant sentence effect in either the left vTPJ or the left lATL. The whole-brain results of Experiment 2 replicated the main findings of Experiment 1, except that the significant interaction effect was found only in the left vTPJ (Fig. 2c and Supplementary Table 2).

We further examined the results for the target brain regions (that is, the left vTPJ and lATL) using region-of-interest (ROI) analysis. To reveal the robustness of the results, we used two different ways to define the target ROIs. The first way was to define the ROIs on the basis of a meta-analysis of the previous neuroimaging studies that compared sentences with word lists[27]; the second way was to define the ROIs on the basis of the individual result of the sentence effect ('HSS + NSS' > 'HSWL + NSWL') on half of the data within predefined

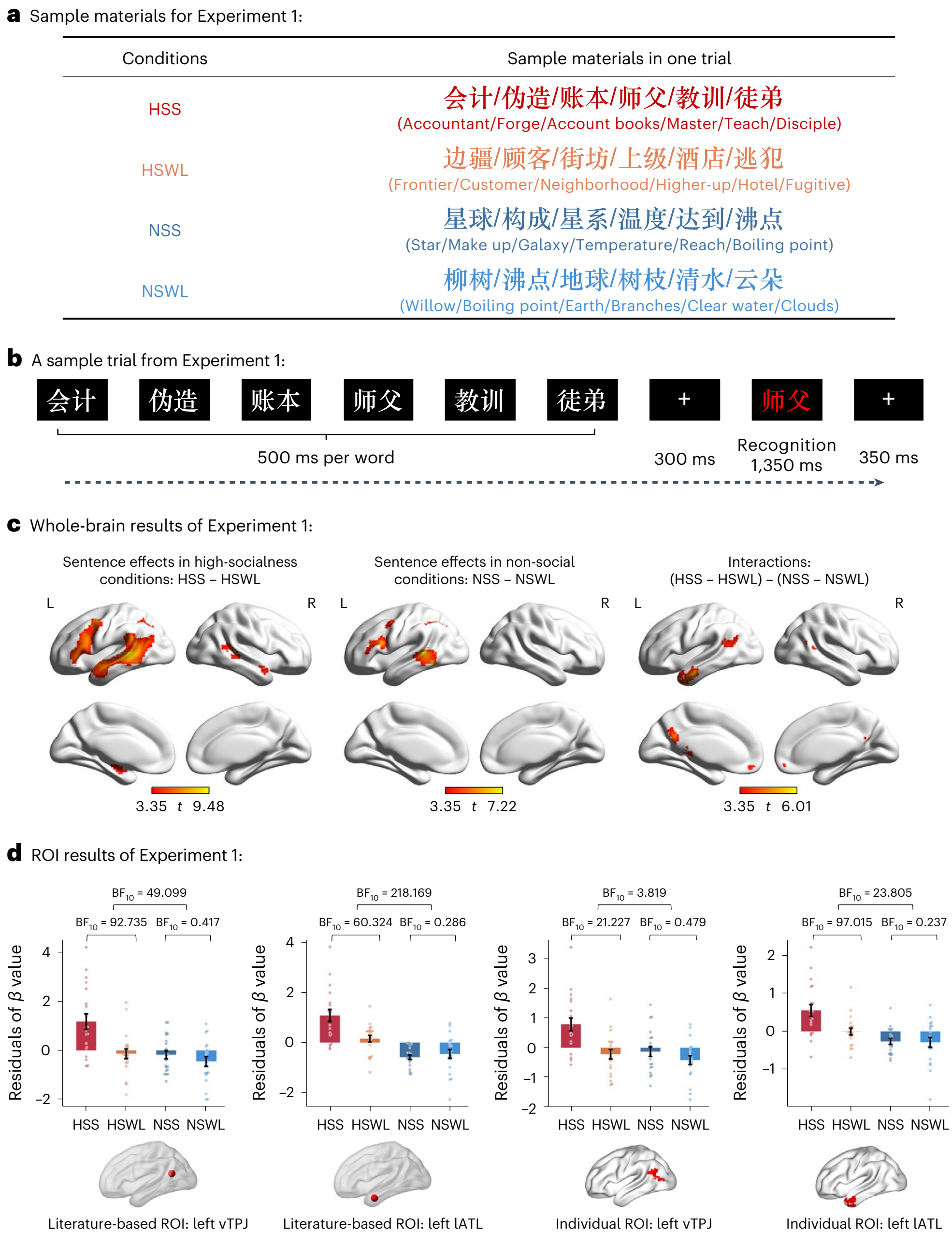

Fig. 1 | Social meaning drives sentence effects in the left vTPJ and lATL (Experiment 1). **a**, Sample materials for each critical experimental condition. For each trial, the stimuli were six words of either high or no socialness. The words formed two short sentences or an unconnected word list. The slashes indicate the word boundaries. Note that although the two sample word lists provided in the figure are both noun lists, one third of the word-list stimuli were verb lists. **b**, The trial structure of the task. In each trial, the participants saw the stimuli word by word and then performed a word recognition task. The experiment used a block design, with each block consisting of four trials of the same condition. **c**, The whole-brain results (for the full results, see Supplementary Table 1). An interaction between social-semantic and sentence effects was found in the left vTPJ and lATL. The left vTPJ and lATL showed stronger activation to sentences than to word lists in the high-socialness conditions (HSS versus HSWL) but not in the non-social conditions (NSS versus NSWL). **d**, The ROI results (sample size, $n = 20$; for the results of classical parametric tests, see Supplementary Tables 3 and 4). The two plots on the left show the results in the target ROIs defined on the basis of a meta-analysis for the contrast between sentences and word lists[27]. The two plots on the right show the results in the target ROIs defined on the basis of the contrast between sentences and word lists on half of the individual data. The bars show the mean residuals of the β values with the IES being regressed out, the error bars show the standard errors, and each point shows the data of a participant. The BF_{10} values represent the ratios of the likelihood of the data under the alternative hypothesis compared to the null hypothesis. The brain maps below the bar graphs show the locations of the literature-based ROIs and the group-constrained masks for individual ROIs. All target ROIs show a strong sentence effect in high-socialness conditions, no sentence effect in non-social conditions, and an interaction between social-semantic and sentence effects.

group-level masks of the left vTPJ and lATL (for details, see the Methods). Because recent studies have found that the left vTPJ and lATL showed deactivation to the general task effort in ROI analysis[14,43], we regressed out the task-effort effect as reflected by the average inverse efficiency score (IES, which is defined as the mean reaction time divided by accuracy[44]) of each condition and participant (Methods). The ROI analyses found consistent patterns in the two experiments (Figs. 1d and 2d and Supplementary Tables 3–6). In all target ROIs,

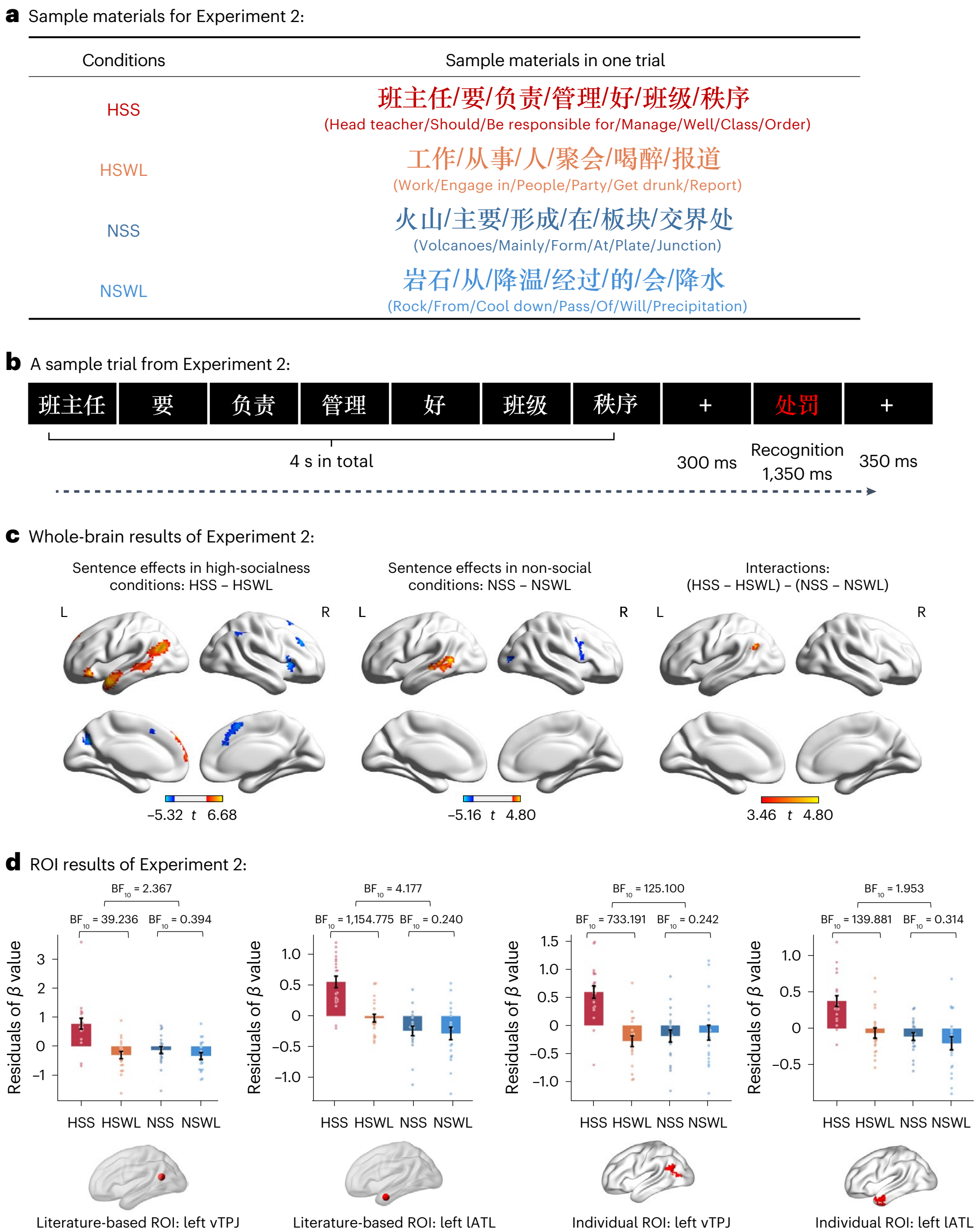

Fig. 2 | Replicating the findings of Experiment 1 using longer and more natural sentential stimuli (Experiment 2). **a**, The sample materials for each experimental condition. For each trial, the stimuli were a series of words that formed either a sentence or a word list of either high or no socialness. The slashes indicate the word boundaries. **b**, The trial structure of the task. The task was the same as that used in Experiment 1. **c**, The whole-brain results (for the full results, see Supplementary Table 2). An interaction between social-semantic and sentence effects was found in the left vTPJ. As in Experiment 1, the left vTPJ and lATL showed stronger activation to sentences than to word lists in the high-socialness conditions (HSS versus HSWL) but not in the non-social conditions (NSS versus NSWL). **d**, The ROI results (sample size, $n = 20$; for the results of classical parametric tests, see Supplementary Tables 5 and 6). The layouts are identical to those of Fig. 1d. The results replicate the findings in Fig. 1d: all target ROIs show a strong sentence effect in high-socialness conditions, no sentence effect in non-social conditions, and an interaction between social-semantic and sentence effects.

both Bayesian and classical parametric tests found an interaction between social-semantic and sentence effects (for the literature-based left vTPJ in Experiment 1: $t_{19} = 4.031$, d (Cohen's d, which represents the difference in means between conditions that is standardized by the pooled standard deviation) = 0.904; 95% confidence interval (CI), (0.50, 1.58); $P < .001$; for the literature-based left lATL in Experiment 1: $t_{19} = 4.776$; $d = 1.071$; 95% CI, (0.60, 1.52); $P < .001$; for the individual ROIs of the left vTPJ in Experiment 1: $t_{19} = 2.698$; $d = 0.605$; 95% CI, (0.16, 1.28); $P = 0.014$; for the individual ROIs of the left lATL in Experiment 1: $t_{19} = 3.666$; $d = 0.815$; 95% CI, (0.23, 0.83); $P = 0.002$; for the literature-based left vTPJ in Experiment 2: $t_{19} = 2.422$; $d = 0.54$; 95% CI, (0.12, 1.62); $P = 0.026$; for the literature-based left lATL in Experiment 2: $t_{19} = 2.748$; $d = 0.607$; 95% CI, (0.12, 0.96); $P = 0.013$; for the individual ROIs of the left vTPJ in Experiment 2: $t_{19} = 4.498$; $d = 1.011$; 95% CI, (0.50, 1.38); $P < .001$; for the individual ROIs of the left lATL in Experiment 2: $t_{19} = 2.306$; $d = 0.522$; 95% CI, (0.04, 0.66); $P = 0.033$), and simple effect analysis showed sentence effects in the contrast between the HSS and HSWL conditions (HSS > HSWL; for the literature-based left vTPJ in Experiment 1: $t_{19} = 4.349$; $d = 0.978$; 95% CI, (0.69, 1.95); $P < .001$; for the literature-based left lATL in Experiment 1: $t_{19} = 4.134$; $d = 0.92$;

95% CI, (0.45, 1.39); $P < .001$; for the individual ROIs of the left vTPJ in Experiment 1: $t_{19} = 3.608$; $d = 0.808$; 95% CI, (0.42, 1.60); $P = 0.002$; for the individual ROIs of the left lATL in Experiment 1: $t_{19} = 4.372$; $d = 0.982$; 95% CI, (0.29, 0.83); $P < .001$; for the literature-based left vTPJ in Experiment 2: $t_{19} = 3.919$; $d = 0.878$; 95% CI, (0.50, 1.66); $P = 0.001$; for the literature-based left lATL in Experiment 2: $t_{19} = 5.617$; $d = 1.255$; 95% CI, (0.37, 0.81); $P < .001$; for the individual ROIs of the left vTPJ in Experiment 2: $t_{19} = 5.385$; $d = 1.192$; 95% CI, (0.53, 1.21); $P < .001$; for the individual ROIs of the left lATL in Experiment 2: $t_{19} = 4.554$; $d = 1.023$; 95% CI, (0.24, 0.64); $P < .001$) but not in the contrast between the NSS and NSWL conditions (for the literature-based left vTPJ in Experiment 1: $t_{19} = 1.155$; $d = 0.259$; 95% CI, (−0.23, 0.79); $P = 0.263$; for the literature-based left lATL in Experiment 1: $t_{19} = 1.291$; $d = 0.287$; 95% CI, (−0.18, 0.76); $P = 0.212$; for the individual ROIs of the left vTPJ in Experiment 1: $t_{19} = -0.678$; $d = 0.148$; 95% CI, (−0.54, 0.28); $P = 0.506$; for the individual ROIs of the left lATL in Experiment 1: $t_{19} = 0.201$; $d = 0.051$; 95% CI, (−0.25, 0.31); $P = 0.843$; for the literature-based left vTPJ in Experiment 2: $t_{19} = 1.096$; $d = 0.25$; 95% CI, (−0.18, 0.60); $P = 0.287$; for the literature-based left lATL in Experiment 2: $t_{19} = 0.262$; $d = 0.056$; 95% CI, (−0.29, 0.37); $P = 0.796$; for the individual ROIs of the left vTPJ in Experiment 2: $t_{19} = -0.305$; $d = 0.067$; 95% CI, (−0.48, 0.36); $P = 0.764$; for the individual ROIs of the left lATL in Experiment 2: $t_{19} = 0.823$; $d = 0.18$; 95% CI, (−0.14, 0.32); $P = 0.421$). Across all tests comparing the NSS and NSWL conditions, the Bayesian factors were consistently lower than 1/3 or 1/2, indicating evidence in favour of the null hypothesis. Notably, in two of these tests, the β values for the NSS condition were even slightly lower than those for the NSWL condition. These findings suggest that the left vTPJ and lATL demonstrate negligible or no sensitivity to non-social sentences.

We further examined the social-semantic and sentence effects in three other regions. The first two regions were the dorsal medial prefrontal cortex (dmPFC) and right vTPJ. Some studies have found social-semantic and sentence effects in these regions[20,41], but they are not viewed as classic regions of the sentence-processing network[23,30,45]. We therefore defined them as two supplementary ROIs. These ROIs showed very similar patterns of results to the left vTPJ and lATL (Supplementary Fig. 2 and Supplementary Tables 7 and 8; for the interaction of the dmPFC in Experiment 1: $t_{19} = 2.61$; $d = 0.588$; 95% CI, (0.10, 0.84); $P = 0.017$; for the 'HSS > HSWL' contrast of the dmPFC in Experiment 1: $t_{19} = 3.366$; $d = 0.754$; 95% CI, (0.17, 0.75); $P = 0.003$; for the 'NSS > NSWL' contrast of the dmPFC in Experiment 1: $t_{19} = -0.095$; $d = 0.018$; 95% CI, (−0.27, 0.25); $P = 0.925$; for the interaction of the right vTPJ in Experiment 1: $t_{19} = 2.406$; $d = 0.541$; 95% CI, (0.06, 0.86); $P = 0.026$; for the 'HSS > HSWL' contrast of the right vTPJ in Experiment 1: $t_{19} = 3.586$; $d = 0.797$; 95% CI, (0.23, 0.87); $P = 0.001$; for the 'NSS > NSWL' contrast of the right vTPJ in Experiment 1: $t_{19} = 0.652$; $d = 0.138$; 95% CI, (−0.21, 0.39); $P = 0.522$; for the interaction of the dmPFC in Experiment 2: $t_{19} = 2.819$; $d = 0.631$; 95% CI, (0.18, 1.22); $P = 0.011$; for the 'HSS > HSWL' contrast of the dmPFC in Experiment 2: $t_{19} = 3.982$; $d = 0.892$; 95% CI, (0.35, 1.13); $P = 0.001$; for the 'NSS > NSWL' contrast of the dmPFC in Experiment 2: $t_{19} = 0.275$; $d = 0.059$; 95% CI, (−0.28, 0.36); $P = 0.786$; for the interaction of the right vTPJ in Experiment 2: $t_{19} = 2.001$; $d = 0.442$; 95% CI, (−0.02, 0.70); $P = 0.06$; for the 'HSS > HSWL' contrast of the right vTPJ in Experiment 2: $t_{19} = 2.193$; $d = 0.49$; 95% CI, (0.01, 0.47); $P = 0.041$; for the 'NSS > NSWL' contrast of the right vTPJ in Experiment 2: $t_{19} = -0.812$; $d = 0.175$; 95% CI, (−0.37, 0.17); $P = 0.427$), indicating that their sensitivity to sentences is also associated with social-semantic comprehension. The third region was the left dorsal IFG. The sentence effect in this region has been found to be driven by syntactic processing[25,46,47] and thus should not be influenced by the socialness of sentence meaning. We therefore defined this region as a control ROI. As expected, in this ROI, both Bayesian and classical parametric tests revealed a strong sentence effect (for Experiment 1: $t_{19} = 4.583$; $d = 1.025$; 95% CI, (0.86, 2.30); $P = < 0.001$; for Experiment 2: $t_{19} = 3.238$; $d = 0.724$; 95% CI, (0.18, 0.84); $P = 0.004$) but no effect of socialness (for Experiment 1: $t_{19} = 0.397$; $d = 0.089$; 95% CI, (−0.64, 0.93); $P = 0.696$; for Experiment 2: $t_{19} = -0.347$; $d = -0.078$; 95% CI, (−0.46, 0.33); $P = 0.732$) or interaction (for Experiment 1: $t_{19} = 0.578$; $d = 0.129$; 95% CI, (−0.28, 0.49); $P = 0.57$; for Experiment 2: $t_{19} = 0.188$; $d = 0.042$; 95% CI, (−0.50, 0.60); $P = 0.853$; Supplementary Fig. 3). This is in sharp contrast to the results found for the left vTPJ and lATL, which indicates that not all areas of the sentence-processing network are sensitive to social semantics.

In addition to the sentence and word-list conditions, Experiment 1 included two supplementary baseline conditions consisting of high-socialness and non-social character lists (HSCL and NSCL conditions). (Most Chinese characters have unique meanings, and the meaning of a Chinese word is often related to those of its constituent characters. The constituent characters of high-socialness words, which form the stimuli for the corresponding character-list conditions, generally have high socialness in their meanings. We therefore refer to the two character-list conditions as the high-socialness and non-social character-list conditions.) The results using these supplementary baseline conditions are similar to those using the word-list baselines (for the interaction of the literature-based left vTPJ: $t_{19} = 3.337$; $d = 0.748$; 95% CI, (0.41, 1.79); $P = 0.003$; for the 'HSS > HSCL' contrast of the literature-based left vTPJ: $t_{19} = 3.344$; $d = 0.747$; 95% CI, (0.47, 2.07); $P = 0.003$; for the 'NSS > NSCL' contrast of the literature-based left vTPJ: $t_{19} = 0.56$; $d = 0.128$; 95% CI, (−0.48, 0.84); $P = 0.582$; for the interaction of the literature-based left lATL: $t_{19} = 4.208$; $d = 0.936$; 95% CI, (0.52, 1.54); $P < .001$; for the 'HSS > HSCL' contrast of the literature-based left lATL: $t_{19} = 3.667$; $d = 0.819$; 95% CI, (0.37, 1.35); $P = 0.002$; for the 'NSS > NSCL' contrast of the literature-based left lATL: $t_{19} = -0.933$; $d = 0.205$; 95% CI, (−0.56, 0.22); $P = 0.362$; for the interaction of the individual ROIs of the left vTPJ: $t_{19} = 2.669$; $d = 0.595$; 95% CI, (0.20, 1.68); $P = 0.015$; for the 'HSS > HSCL' contrast of the individual ROIs of the left vTPJ: $t_{19} = 3.26$; $d = 0.724$; 95% CI, (0.33, 1.51); $P = 0.004$; for the 'NSS > NSCL' contrast of the individual ROIs of the left vTPJ: $t_{19} = -0.059$; $d = 0.016$; 95% CI, (−0.61, 0.57); $P = 0.954$; for the interaction of the individual ROIs of the left lATL: $t_{19} = 3.7$; $d = 0.823$; 95% CI, (0.28, 1.02); $P = 0.002$; for the 'HSS > HSCL' contrast of the individual ROIs of the left lATL: $t_{19} = 4.011$; $d = 0.887$; 95% CI, (0.22, 0.72); $P = 0.001$; for the 'NSS > NSCL' contrast of the individual ROIs of the left lATL: $t_{19} = -1.488$; $d = 0.327$; 95% CI, (−0.44, 0.08); $P = 0.153$; Supplementary Fig. 4 and Supplementary Tables 9–11).

Taken together, in both Experiments 1 and 2, we found that the left vTPJ and lATL showed sensitivity to sentences only when the sentences conveyed social meaning, which is robust to different sentence lengths and structures. This finding is consistent with the prediction of the social-semantic working-memory hypothesis but not that of the general-semantic and/or syntactic accounts for the neural activity of the left vTPJ and lATL.

Persistent social-semantic activation in the left vTPJ and lATL

Working memory is characterized by persistent neural activity during the maintenance of information[33,48]. Experiments 3 and 4 therefore examined whether the left vTPJ and lATL showed persistent social-semantic-selective neural activity after the linguistic stimuli conveying the social meanings disappeared.

Persistent social-semantic activity during the delay period. In Experiment 3, we examined whether the left vTPJ and lATL showed persistent social-semantic-selective neural activity as reflected by the amplitude of the blood-oxygenation-level-dependent (BOLD) signals. Note that persistent neural activity is not always associated with the amplitude of BOLD signals; rather, in many cases, it reflects as multiple-voxel activation patterns[49]. However, the social-semantic working-memory hypothesis assumes that in the left vTPJ and lATL, the increased BOLD signals in high-socialness-sentence comprehension reflect social-semantic working memory. According to this assumption, similar effects should also be found during the maintenance of

https://doi.org/10.1038/s41562-023-01704-8

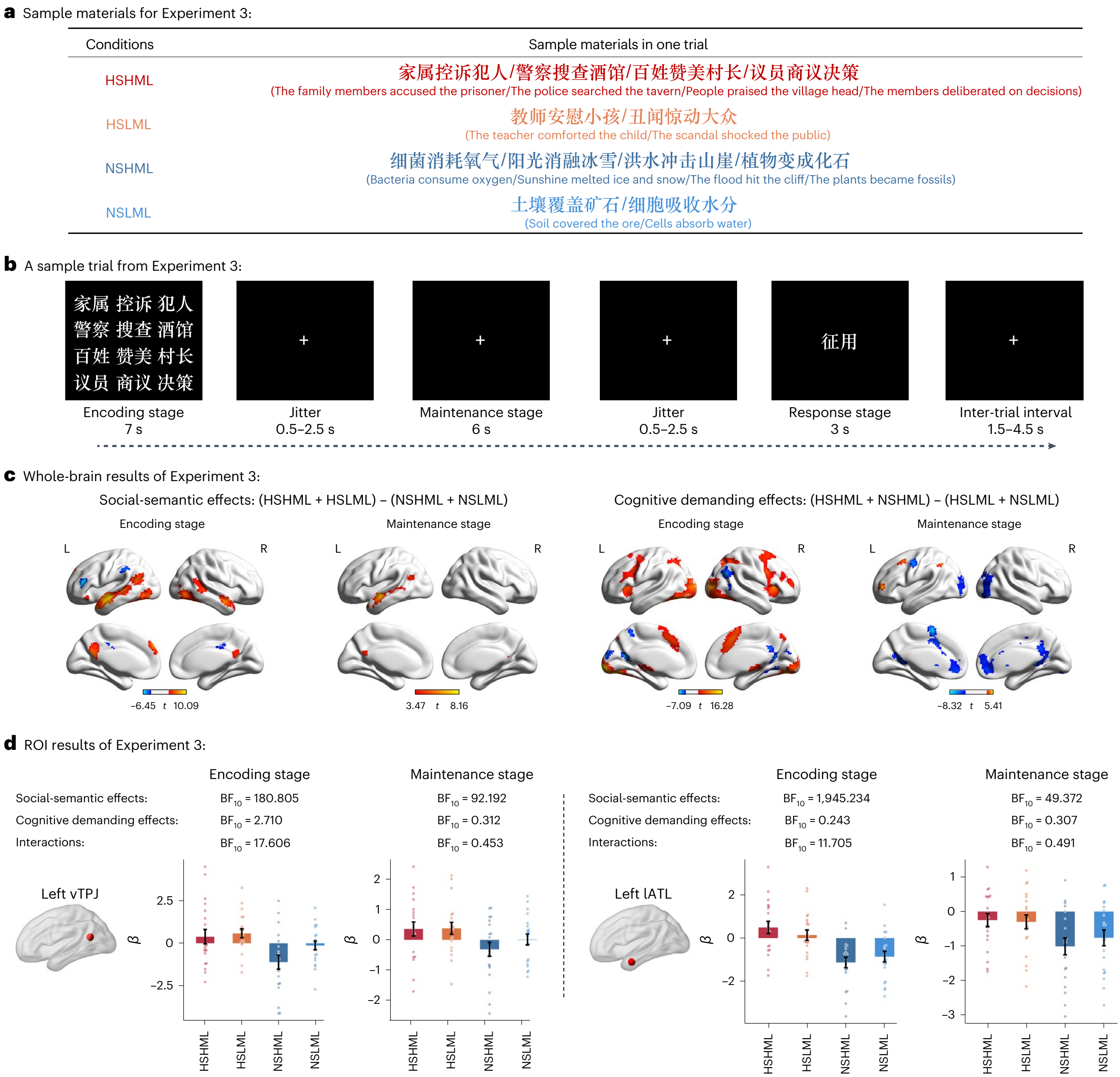

Fig. 3 | Persistent social-semantic-selective neural activity during the delay period (Experiment 3). a, The sample materials for each experimental condition. For each trial, the stimuli were two or four short sentences consisting of words of either high or no socialness. The slashes indicate the sentence boundaries. **b**, The trial structure of the task. In each trial, the participants were asked to read and maintain the words. After a delay period, they performed a word recognition task. **c**, The whole-brain results (for more information on the results, see Supplementary Tables 12 and 13). The left vTPJ and lATL showed social-semantic activation in both the encoding and maintenance stages, which is consistent with the key prediction of the social-semantic working-memory hypothesis. Encoding more sentences invoked stronger activation in the core fronto-parietal working-memory network and visual cortex; maintaining more sentences invoked stronger activation in a few areas in the left lateral frontal and parietal cortex. An interaction between social-semantic and demanding effects was found only in the posterior cingulate in the encoding stage, which is not shown in the figure (Supplementary Table 12). **d**, The ROI results (sample size, $n = 20$; for the results of classical parametric tests, see Supplementary Table 14). The bars show the mean β values, the error bars show the standard errors and each point shows the data of a participant. As predicted by the social-semantic working-memory hypothesis, in both ROIs, social-semantic effects were found in both the encoding and maintenance stages. An interaction between social-semantic and demanding effects was found in both ROIs in the encoding stage but not the maintenance stage. HSHML, high-socialness and high-memory-load; HSLML, high-socialness and low-memory-load; NSHML, non-social and high-memory-load; NSLML, non-social and low-memory-load.

the sentences. To examine this prediction, we gave the participants a delayed recognition task during fMRI scanning. Each trial contained three stages: in the encoding stage, the participants read two or four sentences consisting of high-socialness or non-social words; in the maintenance stage, the participants maintained the sentences for a period; and in the response stage, the participants made responses to a forced-choice word recognition task. The sentential stimuli were identical to those used in Experiment 1, each consisting of three words. See Fig. 3a,b for sample stimuli and trials.

We varied the number of sentences to manipulate the memory load. Memory load is a classic factor in working-memory studies, and its effect has been reliably found in the core fronto-parietal working-memory network[50,51]. In the brain regions that are assumed to selectively represent particular types of contents (such as objects, faces and mental states), previous findings on load effects have been inconsistent: some studies have found load effects for working memory of specific stimuli[52–54], but others have not[55–58]. We therefore explored the interaction between social-semantic and load effects in the left vTPJ and lATL.

The whole-brain results of Experiment 3 are shown in Fig. 3c and Supplementary Tables 12 and 13. The left vTPJ and lATL showed the persistent social-semantic-selective activity predicted by the social-semantic working-memory hypothesis: both regions showed significant social-semantic effects (high-socialness > non-social) at both the encoding and maintenance stages, as reflected by the amplitude of the BOLD signals. These regions, however, showed no significant response to the sentence number or interaction between the two factors at either the encoding or maintenance stage.

As in Experiments 1 and 2, we conducted ROI analyses to further examine the results within the left vTPJ and lATL. To remain consistent with Experiments 1 and 2, the ROIs were defined on the basis of the results of Zaccarella et al.[27]. Unlike in Experiments 1 and 2, we were not able to define individual ROIs because Experiment 3 did not include any localizer tasks; we also did not regress out the effect of task efforts because we explicitly manipulated the task demands (that is, memory load) in this experiment.

The results of the ROI analysis are shown in Fig. 3d and Supplementary Table 14. Both Bayesian and classical parametric tests showed social-semantic effects in both ROIs at both the encoding (for the left vTPJ: $t_{19} = 4.682$; $d = 1.047$; 95% CI, (1.22, 3.20); $P < .001$; for the left lATL: $t_{19} = 5.885$; $d = 1.317$; 95% CI, (1.69, 3.55); $P < .001$) and maintenance stages (for the left vTPJ: $t_{19} = 4.346$; $d = 0.972$; 95% CI, (0.54, 1.54); $P < .001$; for left the lATL; $t_{19} = 4.034$, $d = 0.904$; 95% CI, (0.59, 1.85); $P = 0.001$), confirming the key prediction of the social-semantic working-memory hypothesis. In addition, at the encoding stage, both ROIs showed the interaction between the social-semantic and sentence-number effects (for the left vTPJ: $t_{19} = 3.512$; $d = 0.782$; 95% CI, (0.32, 1.26); $P = 0.002$; for the left lATL: $t_{19} = 3.302$; $d = 0.733$; 95% CI, (0.23, 1.03); $P = 0.004$): larger social-semantic effects were found in the four-sentence conditions than in the two-sentence conditions; however, no such effect was found at the maintenance stage (for the left vTPJ: $t_{19} = 1.239$; $d = 0.275$; 95% CI, (−0.21, 0.81); $P = 0.231$; for the left lATL: $t_{19} = 1.314$; $d = 0.297$; 95% CI, (−0.17, 0.77); $P = 0.204$). These findings indicate that the left vTPJ and lATL can show persistent social-semantic-selective neural activity; in addition, these regions are selectively sensitive to the encoding load of social-semantic information but insensitive to the maintenance load of it. (As mentioned above, previous findings have been inconsistent in the load effects in the areas that are assumed to selectively represent particular types of contents. One possibility is that the emergence of such effects relies on task-dependent modulation by the core working-memory areas[58]. According to this view, at the maintenance stage, the lack of an interaction effect in the left vTPJ and lATL might be associated with the weakness of the load effect in the core fronto-parietal working-memory network as shown in Fig. 3c.)

We performed dynamic causal modelling (DCM) analysis to further explore the direction of influences between the left vTPJ and lATL and whether the directional connections between the two regions are modulated by social-semantic encoding and maintenance. We constructed and compared nine models with the two regions, considering all possible combinations of directional connections and modulations (Supplementary Fig. 5a,c; for details, see the Methods). For both the encoding and maintenance stages, we found that the winning models were those with bidirectional intrinsic connections and bidirectional modulations by social-semantic inputs (Supplementary Fig. 5b,d). These findings indicate that the left vTPJ and lATL are functionally interdependent during both social-semantic encoding and maintenance, with bidirectional transmission of social-semantic information between them. This is consistent with previous research indicating that these two areas form a subsystem of the social-cognition network[59,60].

Social-semantic working memory of the last sentence. Language comprehension often requires processing successive sentences, during which the meaning of the context sentences must be maintained while the current sentence is processed. In Experiment 4, we examined whether the left vTPJ and lATL showed persistent social-semantic-selective neural activity during the processing of successive sentences. We conducted multivariate pattern analyses (MVPA) to reveal the semantic contents of the neural representation, which allows us to decode the maintained semantic representations of the context sentence from the neural activity associated with the presentation of the current sentence.

In Experiment 4, the participants accomplished a 'mental portrait' task. In each trial, the participants read two successive sentences describing two features of a person; then, they saw two photos of different people and decided which photo was more consistent with the preceding sentences by pressing buttons. In half of the trials, people read sentences about two trait dimensions, dominance and trustworthiness, which are the major trait dimensions associated with face evaluation[61]. In the other half of the trials, people read sentences about two physical facial dimensions, the size of the face and the length of the eyebrows. Although the trait and physical facial dimensions are both person-related, the trait dimensions should have higher socialness than the physical ones because they are more directly related to interactions between people. To dissociate the brain activities associated with the presentation of the two sentences and the pictures, the sentences and pictures were separated by jitters (for task procedure, see Fig. 4a).

The key prediction of the experiment is that if the left vTPJ and lATL represent social-semantic working memory during the processing of successive sentences, then they should maintain the social meaning of the first sentence while the second sentence is being processed. We therefore performed MVPA to decode the poles of each dimension (for example, high dominance versus low dominance) expressed in the first sentence from the neural activity associated with the presentation of the second sentence. We also conducted MVPA to decode the poles of each dimension expressed in each sentence from the neural activity associated with the presentation of the sentence itself. The analyses were performed at both the whole-brain and ROI levels (Methods). The whole-brain searchlight analysis failed to reveal any significant results. The ROI analysis was based on the same ROIs as in Experiment 3. As shown in Fig. 4c, in the left vTPJ, the poles of dominance expressed by the sentences could be decoded from both their concomitant and delayed neural activity (with average decoding accuracies of 3.22% and 6.05% above the 50% chance level, with lower bounds of one-sided 95% CIs of 51.14% and 52.73%, and with P values of 0.010 and 0.001, respectively); in the left lATL, the poles of dominance expressed by the sentences could be decoded from their delayed neural activity (with an average decoding accuracy of 3.52% above the 50% chance level, with a lower bound of the one-sided 95% CI of 50.31% and with a P value of 0.037), but the effect did not survive the Bonferroni correction for the number of dimensions decoded ($n = 4$). No other dimension could be decoded from either the concomitant or the delayed neural activity of the sentences. Our results therefore partially confirmed the prediction of the social-semantic working-memory hypothesis by showing that at least social-semantic information associated with dominance can be maintained in the left vTPJ (and possibly the left lATL) during the processing of successive sentences.

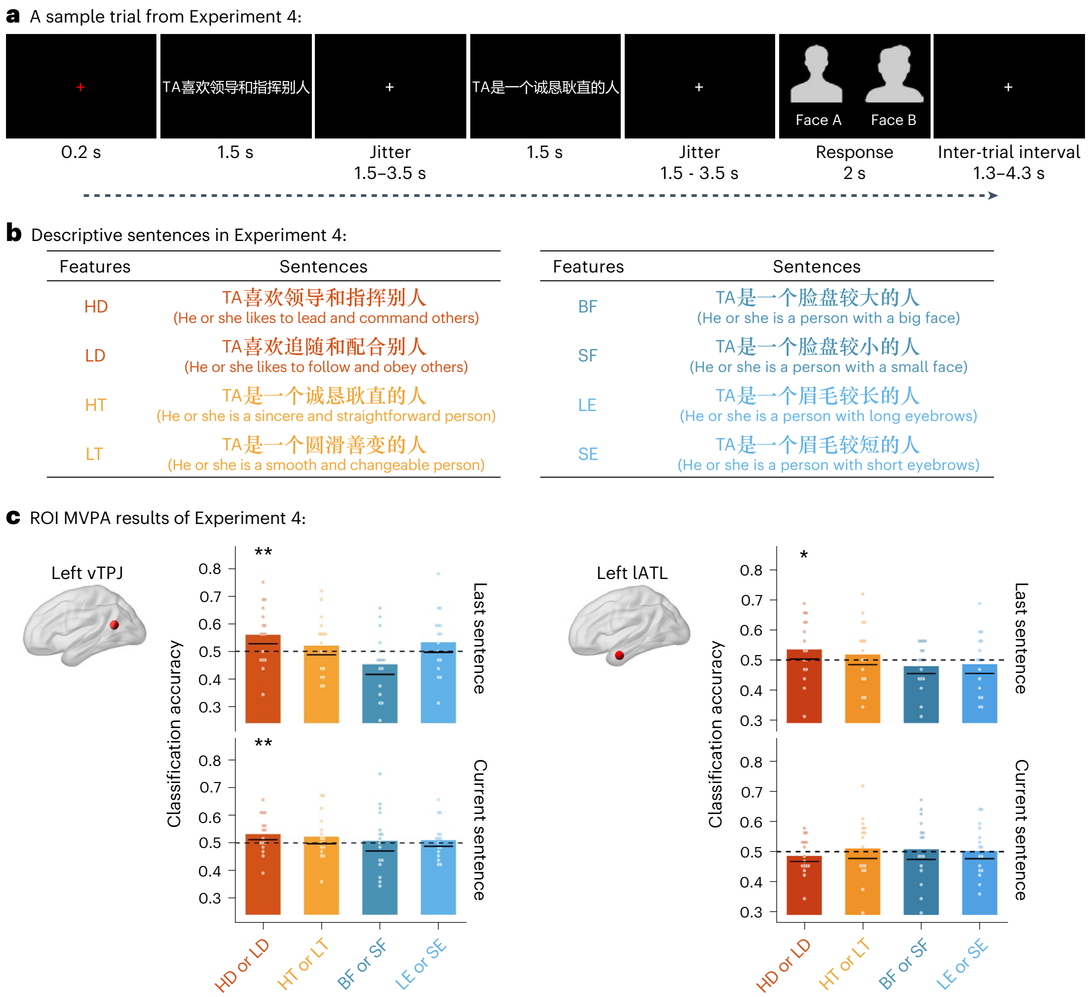

Fig. 4 | Persistent social-semantic-selective neural activity during the processing of successive sentences (Experiment 4). a, The trial structure of the task. In each trial, the participants read two sentences describing either two trait features (dominance and trustworthiness) or two physical facial features (big/small face and long/short eyebrows) of a person. They were then asked to choose a photograph from two to match the contents of the sentences. In the experiment, the images labelled Face A and Face B were authentic portrait photos from the CAS-PEAL Chinese face database. **b**, The descriptive sentences for each trait or physical facial dimension. **c**, The results of MVPA in ROIs (sample size, $n = 16$). The bars show the average classification accuracies on the poles of the four dimensions described in the last (top) and current sentences (bottom). The black solid line in each bar shows the lower bound of the 95% CI for the classification accuracy obtained using bootstrapping, and each point shows the data of a participant. The dashed line shows the chance level (50%). In the left vTPJ, the features of dominance described in the current (accuracy, 53.22%; $P_{\text{uncorrected}} = 0.01$) and last sentences (accuracy, 56.05%; $P_{\text{uncorrected}} = 0.001$) can both be decoded from the activation patterns; in the left lATL, the features of dominance described in the last sentence can be decoded from the activation patterns (accuracy, 53.52%; $P_{\text{uncorrected}} = 0.037$), but the effect did not survive the Bonferroni correction for the number of dimensions decoded ($n = 4$). No other features can be decoded from the activation of the two ROIs. All statistical tests were one-sided. HD, high dominance; LD, low dominance; HT, high trustworthiness; LT, low trustworthiness; BF, big face; SF, small face; LE, long eyebrows; SE, short eyebrows. $^{*}P < 0.05$; $^{**}P < 0.01$ (uncorrected). Icons in **a** made by Freepik and Pixel perfect from www.flaticon.com.

Non-linguistic social stimuli activate the left vTPJ and lATL

The above experiments indicate that the neural activity of the left vTPJ and lATL during sentence processing is associated with social-semantic working memory rather than linguistic processes. However, because all the above experiments used sentence stimuli to induce social-semantic processing, it was unclear whether the social-semantic sensitivity of the left vTPJ and lATL was specific to language comprehension. If the left vTPJ and lATL represent social-semantic working memory, then they should be sensitive to the social meaning of both linguistic and non-linguistic stimuli.

In Experiment 5, we examined the sensitivity of the left vTPJ and lATL to the social information conveyed by non-linguistic stimuli. The participants were asked to watch short silent videos and rate their pleasantness (see Fig. 5b for the task procedures). There were three kinds of videos corresponding to three conditions: the high-socialness (HS) videos were about interactions between people, the single-person (SP) videos contained only one person and the non-social (NS) videos contained only natural and non-human entities.

The whole-brain results are shown in Fig. 5c and Supplementary Table 15. In the left vTPJ, significant differences were found between each two of the three conditions, with the HS condition evoking the strongest neural activity and the NS condition evoking the weakest neural activity. The left lATL showed stronger activation to the HS condition than to the NS condition. A small cluster in the left dorsal ATL and temporal pole showed stronger activation to the SP condition than to the NS condition. Significant socialness effects were also found in many other brain regions of the social-cognitive system (Fig. 5c). The ROI results are shown in Fig. 5d. The ROI analysis was

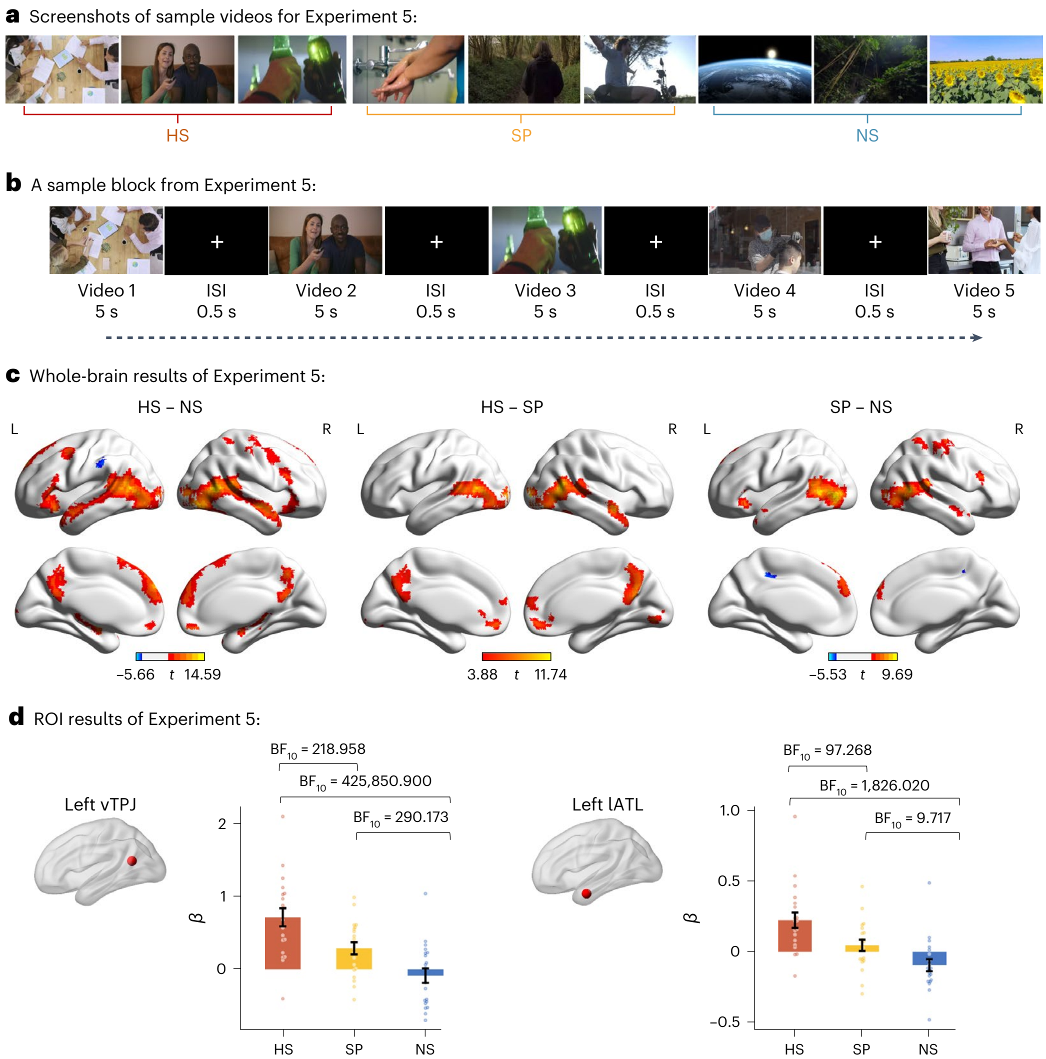

Fig. 5 | The left vTPJ and lATL are sensitive to the socialness of non-linguistic stimuli (Experiment 5). a, Screenshots of sample videos in the HS (left), SP (middle) and NS (right) conditions. **b**, The block structure. Each block contained five videos from the same condition. **c**, The whole-brain results (for more information on the results, see Supplementary Table 15). The left vTPJ showed sensitivity to socialness in all contrasts between conditions (HS > SP > NS). The left lATL showed sensitivity to socialness in the contrast between the HS and NS conditions. **d**, The ROI results (sample size, $n = 20$). The bars show the mean β values, the error bars show the standard errors and each point shows the data of a participant. Both ROIs showed sensitivity to socialness in all contrasts between conditions (HS > SP > NS). ISI, inter-stimulus interval. Images in **a**,**b** reproduced from stock footage provided by Videvo, downloaded from www.videvo.net.

based on the same ROIs used in Experiments 3 and 4. In both ROIs, both Bayesian and classical parametric tests showed differences between each two of the three conditions, with the HS condition evoking the strongest neural activity and the NS condition evoking the weakest neural activity (for the 'HS > SP' contrast of the left vTPJ: $t_{19} = 4.778$; $d = 1.075$; 95% CI, (0.24, 0.62); $P < 0.001$; for the 'HS > SP' contrast of the left lATL: $t_{19} = 4.373$; $d = 1$; 95% CI, (0.10, 0.26); $P < 0.001$; for the 'HS > NS' contrast of the left vTPJ: $t_{19} = 8.921$; $d = 2$; 95% CI, (0.61, 0.99); $P < 0.001$; for the 'HS > NS' contrast of the left lATL: $t_{19} = 5.852$; $d = 1.333$; 95% CI, (0.21, 0.43); $P < 0.001$; for the 'SP > NS' contrast of the left vTPJ: $t_{19} = 4.918$; $d = 1.118$; 95% CI, (0.22, 0.54); $P < 0.001$; for the 'SP > NS' contrast of the left lATL: $t_{19} = 3.204$; $d = 0.7$; 95% CI, (0.05, 0.23); $P = 0.005$). The left vTPJ and lATL are therefore sensitive to social information, whether conveyed by linguistic or non-linguistic stimuli, in line with the social-semantic working-memory hypothesis.

The left vTPJ and lATL belong to the social-semantic network

According to the social-semantic working-memory hypothesis, the function of the left vTPJ and lATL is intrinsically more tightly associated with social-semantic working memory than with linguistic processing. These regions would therefore have stronger intrinsic functional connectivity to the social-semantic-processing areas than to the sentence-processing areas. Experiment 6 examined this prediction using resting-state fMRI.

We defined the seed ROIs on the basis of two published meta-analyses (Zaccarella et al.[27] for the key ROIs and sentence-processing ROIs, and Zhang et al.[42] for the social-semantic-processing ROIs; Fig. 6a). Both meta-analyses identified areas in the left vTPJ and lATL, and their locations were adjacent. To remain consistent with the ROI analyses of Experiments 1 to 5, we defined the key seed ROIs of the left vTPJ and lATL on the basis of Zaccarella et al.[27] and did not include the left vTPJ and lATL area identified by Zhang et al.[42] as seed ROIs. According to our prediction, the left vTPJ and lATL should have stronger intrinsic connectivity to the social-semantic-processing areas than to the sentence-processing areas even when they were defined on the basis of sentence-processing tasks.

We first examined whether the left vTPJ and lATL have stronger resting-state functional connectivity (RSFC) to the remaining

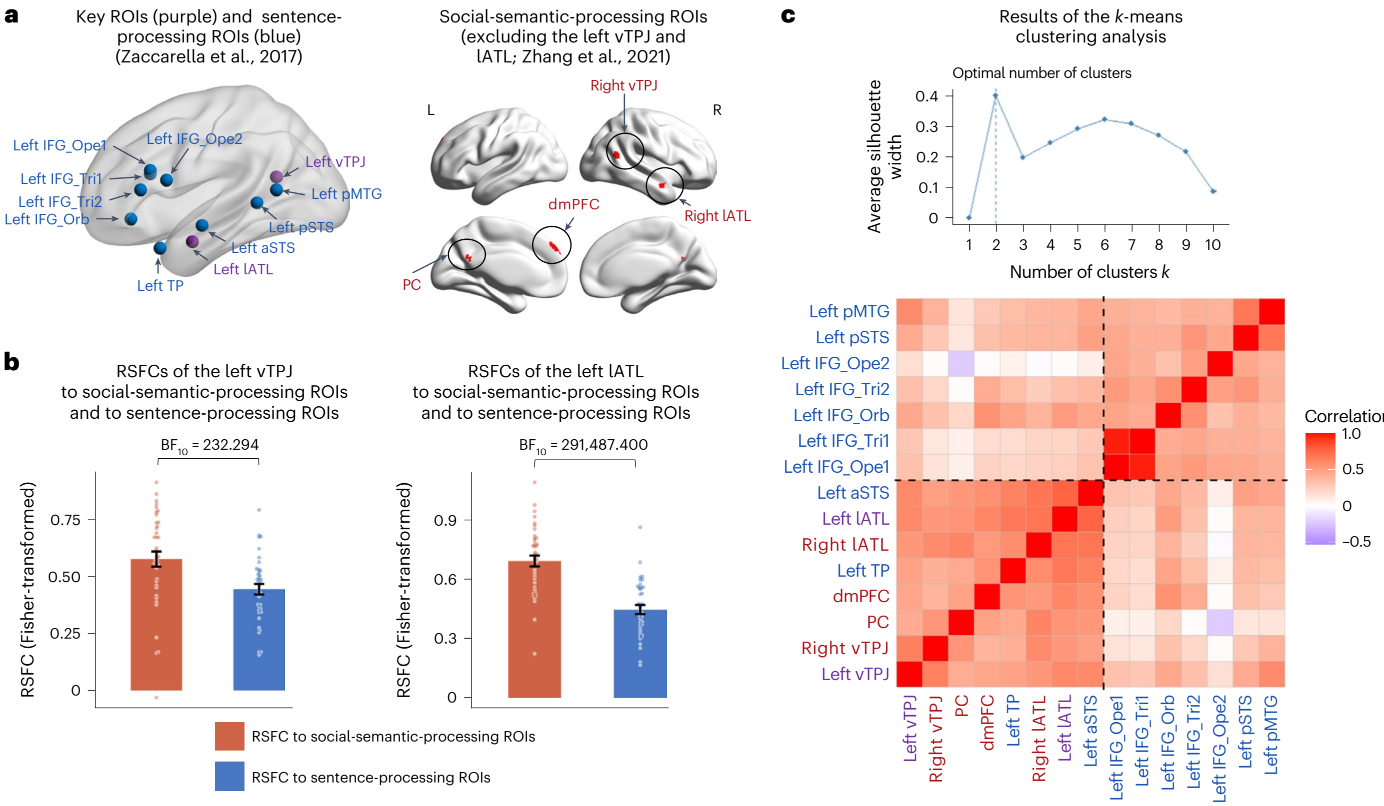

Fig. 6 | The left vTPJ and lATL have stronger intrinsic connectivity to the social-semantic-processing areas than to the sentence-processing areas (Experiment 6). **a**, The locations of the seed ROIs. The ROIs were defined on the basis of two meta-analyses by Zaccarella et al.[27] and Zhang et al.[42]. To remain consistent with the ROI analyses of Experiments 1 to 5, we defined the key ROIs (that is, the left vTPJ and lATL) according to Zaccarella et al.[27]. The left vTPJ and lATL found in the meta-analysis of social-semantic-processing tasks[42] were thus not included in the analysis. The sentence-processing ROIs and social-semantic-processing ROIs were defined according to Zaccarella et al.[27] and Zhang et al.[42], respectively. According to the prediction of the social-semantic working-memory hypothesis, the left vTPJ and lATL should have stronger intrinsic connectivity to the social-semantic-processing ROIs than to the sentence-processing ROIs even when they were defined on the basis of sentence-processing tasks. **b**, Mean RSFCs of the key ROIs to social-semantic-processing and sentence-processing ROIs. For both the left vTPJ and the left lATL, their average RSFCs to the social-semantic-processing ROIs were stronger than those to the sentence-processing ROIs (sample size, $n = 39$). **c**, The results of the *k*-means clustering analysis on all ROIs. The left vTPJ, lATL, aSTS and TP clustered together with the social-semantic-processing ROIs rather than the other sentence-processing ROIs, even though they were defined on the basis of the meta-analysis results of sentence-processing studies[27]. Top, the averaged silhouette scores of the *k*-means clustering analysis. Bottom, the best clustering solution shown by the dashed lines in the RSFC matrix of the seed ROIs. IFG_Orb, orbital part of the IFG; IFG_Ope, opercular part of the IFG; IFG_Tri, triangular part of the IFG; aSTS, anterior superior temporal sulcus; TP, temporal pole; pMTG, posterior middle temporal gyrus; PC, posterior cingulate.

sentence-processing ROIs or to the social-semantic-processing ROIs. Both regions showed stronger RSFC to the social-semantic-processing ROIs than to the sentence-processing ones (for the left vTPJ: $t_{38} = 4.327$; $d = 0.684$; 95% CI, (0.07, 0.19); $P < 0.001$; for the left lATL: $t_{38} = 6.783$; $d = 1.087$; 95% CI, (0.18, 0.32); $P < 0.001$; Fig. 6b). We then conducted a *k*-means cluster analysis based on the correlation matrix between all pairs of seed ROIs (Methods). The silhouette score indicated that these nodes could be best grouped into two clusters (Fig. 6c). The results of the two-cluster solution are shown in Fig. 6c. Four seed ROIs defined by the sentence-processing task, including the left vTPJ, the left lATL and two ROIs close to the lATL (that is, the anterior superior temporal sulcus and temporal pole), clustered together with the social-semantic-processing ROIs, while the other sentence-processing ROIs clustered together. These results confirm our prediction that the left vTPJ and lATL have stronger intrinsic connectivity to the social-semantic-processing areas than to the sentence-processing areas.

Discussion

We examined the function of the left vTPJ and lATL in sentence processing and social-semantic working memory. Two key findings indicate that these regions engage in sentence processing through social-semantic working memory: first, they are more sensitive to sentences than to word lists only if the sentences convey social meaning (Experiments 1 and 2); second, they show persistent social-semantic-selective activity after the linguistic stimuli disappear (Experiments 3 and 4). Two additional findings also indicate that these regions are more tightly associated with social-semantic processing than with linguistic processing: they are sensitive to the socialness of non-linguistic stimuli (Experiment 5) and are intrinsically more tightly connected to the social-semantic-processing areas than to the sentence-processing areas (Experiment 6). Taken together, our results provide converging evidence for the social-semantic working-memory hypothesis of the left vTPJ and lATL and challenge the general-semantic and/or syntactic accounts for the neural activity of these regions in sentence processing.

Our results indicate that during sentence processing, the stronger neural responses of the left vTPJ and lATL to sentences than to word lists are selectively associated with social-semantic comprehension, which is probably due to more durable working memory for coherent social meanings than for incoherent ones in these regions. It is notable that without controlling for the socialness of the stimuli, the activation of these regions in sentence processing has consistently been reported in the literature[25–27,31]. Why is the activity of the left vTPJ and lATL so frequently observed in previous studies of sentence processing? The primary reason could be that language use is a social behaviour and sentences are thus naturally dominated by social-semantic information[62]. For example, it has been found that approximately 2/3 of natural conversations are on social topics[63]. In addition, the left vTPJ and lATL are sensitive to the social meaning of a very broad range of concepts[14]. For example, they are even sensitive to the social meaning of non-living objects[22]. The comprehension of the vast majority of our daily sentences may therefore naturally require the involvement of social-semantic working memory.

Our results provide an alternative explanation for the results of previous neuroimaging studies that compared sentences with fragmented linguistic stimuli. The stronger brain activity in response to sentences than to non-sentential stimuli has been viewed as a classic neural signature for linguistic processing. It has been used for localizing the language network[24,26,29,41], examining the linguistic functions of the brain networks defined by resting-state fMRI data processing[64] and revealing the recruitment of the occipital cortex of congenitally blind individuals in linguistic processing[65]. However, our results show that, without controlling for the socialness of the stimuli, this classic effect may reflect social-semantic working memory. One limitation of this study is the absence of evidence indicating whether sentence effects outside the left vTPJ and lATL may reflect working memory for non-social semantics. However, the findings of a previous study have suggested that this speculation may be valid. Humphries et al.[31] used sentences describing concrete events (which should be rich in visual semantics) as the target stimuli and found that when compared with meaningless sentences and word lists, these sentences induced stronger activation not only in the classic sentence-processing regions but also in the bilateral middle occipital gyri and left fusiform gyrus. This finding indicates that the activation to sentences may also reflect non-social types of semantic working memory (for example, visual-semantic working memory), especially when the non-social-semantic dimensions of sentences are manipulated.

Our results indicate that the left vTPJ and lATL may connect language comprehension with social cognition through social-semantic working memory. Most previous studies on the relationship between language and social cognition focused on the ability to reason about mental states, which is known as theory of mind (ToM)[1,45,66–68]. The key regions supporting ToM are the right TPJ and dmPFC[18,69], which are often engaged in the comprehension of stories and non-literal meanings[6,66,70]. We assume that in comparison with ToM, social-semantic working memory is a more general and basic social-cognitive component that connects language comprehension with social cognition: it is not specific to mental states but is involved in the processing of a wide range of social concepts, and it forms the basis of social-semantic manipulation and integration, which in turn supports higher-order social cognition such as ToM. Consistent with our view, in the field of social neuroscience, the left vTPJ and lATL are associated with not only ToM[71,72] but also other social functions[18,62,73]; in the field of language comprehension, the left vTPJ and lATL are involved in not only the comprehension of stories[6] and non-literal meanings[74] but also social-semantic comprehension of sentences[42], phrases[13,75] and words[19,20,22,76].

Recent research has indicated that social cognition may play more important roles in multiple cognitive domains than previously believed[77]. To explore the functions of the left vTPJ and lATL in broader cognitive domains, we conducted location-based analyses of Neurosynth (neurosynth.org)[78] to identify the cognitive terms associated with these brain regions (see Supplementary Information section B for the details). We identified 15 cognitive terms that are commonly associated with both areas and 14 cognitive terms that are unique to either the vTPJ or the lATL (Supplementary Table 16). All of the 15 common terms pertain to the domains of social, language, semantics and autobiographical memory, and the 14 unique terms are also highly correlated with these domains. The relationships between social-semantic working memory and the first three domains have already been indicated by the current study. The processing of autobiographical memory may also engage social-semantic working memory because relationships and interactions with others are a crucial part of individual experiences. Therefore, the social-semantic working-memory function of the left vTPJ and lATL may account for the majority of their activation observed in previous studies.

Although the left vTPJ and lATL showed highly similar results in the current and previous studies, some functional differences between the two areas have also been indicated by the literature. It has been found that the left lATL is more stably involved in social concept retrieval and word-level social-semantic processing[15,42,73], while the left vTPJ is more sensitive to discourse-level social-semantic processing[42,68]. The left lATL and vTPJ may therefore play greater roles in social concept retrieval and integration, respectively. The left vTPJ also plays a role in cross-modal social-semantic integration: it is sensitive to both speeches and gestures that convey communicative intent[79] and is especially sensitive to co-speech gestures[80].

As indicated by previous social neuroscience studies, the representation of social concepts relies on fine-grained social-semantic dimensions that have distinct neural correlates[81–83]. In Experiment 4, we examined the social-semantic working memory of two specific trait dimensions: dominance and trustworthiness. Only dominance could be decoded from the neural activity of the left vTPJ and lATL. This finding is consistent with the previous finding that dominance is the most salient and conserved across the trait–state divide according to neural representation[83]. It also indicates that the left vTPJ and lATL may not represent all kinds of social-semantic dimensions. However, one limitation of this study is that it does not provide evidence as to whether the left vTPJ and lATL also represent other dimensions of social semantics beyond dominance. In addition, many regions outside the left vTPJ and lATL have been found to represent specific social-semantic subdimensions[81–83]. It remains to be investigated whether these regions support working memory on specific social-semantic subdimensions.

The finding that the left vTPJ is involved in social-semantic working memory can be linked to the previous finding that several functional subdivisions of the left TPJ support working-memory processes. In the field of language processing, the left supramarginal and angular gyri have been found to buffer phonological and semantic information, respectively[84–86]. In the field of social cognition, Meyer and Collier[87] found that the bilateral dorsal TPJs are involved in working memory of mental states of specific individuals such as characters from a television show. These findings, together with ours, indicate that the left TPJ as a whole may play important roles in working memory, with its different subdivisions supporting working memory of different types of information.

Our findings seem to contradict the findings of the classic studies by Bemis and Pylkkänen[88,89], which revealed neural activity associated with phrase-level composition (for example, 'red boat') in the left ATL and angular gyrus. These seemingly inconsistent findings may be explained by two reasons. First, Bemis and Pylkkänen[88,89] did not fully control for the socialness of their stimuli. Some words they used have direct or symbolic social meanings, such as religion (cross), emotions (heart), secrets (lock and key), signals for communication (bell and flag), property (house and car) and politics (red and blue). Moreover, combinations of colour words and nouns could convey further social meanings (for example, 'red heart' and 'black heart' can convey the meanings of love and evil, respectively). Second, both the left TPJ and ATL contain multiple functional subdivisions that support distinct cognitive processes[13,14,20,76,90–93]. Given the broad definitions of ATL and angular gyrus used by Bemis and Pylkkänen[88,89] and the relatively low spatial resolution of magnetoencephalography, it is uncertain whether the composition effects were localized to the vTPJ and lATL or to other subregions of the TPJ and ATL. Several fMRI studies have also shown that phrase-level composition activates the left TPJ[13,94,95]. In a recent study, Lin et al.[13] manipulated both the socialness and the plausibility of phrases and found that in the left TPJ, the region sensitive to the plausibility of phrases is dorsal to the region sensitive to social semantics, with no overlap between the two regions. More recently, Yang and Bi[75] investigated phrase-level composition using representation similarity analysis. They found that the bilateral ATLs are sensitive to social-semantic composition but not to non-social-semantic composition. Therefore, the only two studies that have examined both social-semantic and phrase-composition effects found that the left vTPJ and lATL are selectively sensitive to social-semantic processing rather than to general composition processes.

To conclude, we examined whether the sentence and social-semantic effects observed in the left vTPJ and lATL both reflect social-semantic working memory. We found that the stronger responses of these regions to sentences than to word lists are selectively associated with social-semantic comprehension and that these regions are involved in social-semantic working memory during and after sentence processing, which supports the social-semantic working-memory hypothesis. Our findings provide insights into the function of the left vTPJ and lATL in language comprehension and indicate that these regions may connect language with social cognition through social-semantic working memory.

Methods

Ethic approval

All protocols and procedures of the current study were approved by the Institutional Review Board of the Institute of Psychology of the Chinese Academy of Sciences (IPCAS2019006, IPCAS2020003 and IPCAS2021004). Each participant read and signed the informed consent form before taking part in the experiments. All experiments were conducted in accordance with the Declaration of Helsinki and all relevant ethical regulations.

Participants

The participants were all right-handed and native Chinese speakers. None of them had experienced psychiatric or neurological disorders or had sustained a head injury. The sample sizes of Experiments 1, 2, 3, 4, 5 and 6 were 20 (16 women; mean age, 22.3 years; s.d., 2.3 years), 20 (13 women; mean age, 23.5 years; s.d., 1.9 years), 20 (11 women; mean age, 21.8 years; s.d., 2.4 years), 16 (9 women; mean age, 24.0 years; s.d., 2.4 years), 20 (14 women; mean age, 22.8 years; s.d., 2.7 years) and 39 (28 women; mean age, 22.9 years; s.d., 2.2 years), respectively. These sample sizes were determined by referencing those of previous fMRI studies on social-semantic and sentence effects, which have been summarized in two meta-analyses, conducted by Zhang et al.[42] and Zaccarella et al.[27], respectively. There were 82 participants in total (56 women; mean age, 22.7 years; s.d., 2.4 years). For Experiments 1 to 5, 70 participants took part in only one of the five experiments, 10 participants took part in two of them and 2 participants took part in three of them. The participants of Experiment 6 were all from Experiments 1 and 2. The participants received payments of 120, 120, 120, 150 and 50 RMB for Experiments 1, 2, 3, 4 and 5, respectively.

Designs and procedures

Experiment 1. Experiment 1 included six conditions (the HSS, NSS, HSWL, NSWL, HSCL and NSCL conditions). Each of the six conditions contained 96 trials. For the HSS and NSS conditions, each trial consisted of two sentences. For both conditions, the stimuli consisted of 96 different sentences, with each sentence being presented twice in different pairs. Five independent rating experiments (each recruiting 16 participants who did not participate in the fMRI experiment) were conducted to obtain the socialness, imageability, semantic familiarity and semantic plausibility of the sentences and the socialness of the constituent words. The HSS and NSS conditions were significantly different in both word-level and sentence-level socialness (HSS > NSS) and were matched on all the other ratings. The two conditions were also matched on word frequency (Chinese Linguistic Data Consortium[96]). See Supplementary Table 17 for the statistics of the manipulated and controlled variables of Experiment 1. We then segmented the sentences of the HSS and NSS conditions into words and characters. The constituent words and characters of the HSS condition were used to form the stimuli of the HSWL and HSCL conditions; the constituent words and characters of the NSS condition were used to form the stimuli of the NSWL and NSCL conditions. Each trial of the word-list conditions consisted of six nouns or six verbs. Each trial of the character-list conditions consisted of six character pairs that did not form words. Character pairs are often used as non-words in Chinese reading research (for example, ref. 97). However, given that almost all Chinese characters have their own meanings and that many of them can function as single-character words, it is impossible to rule out semantic or lexical processing from Chinese-character reading. Therefore, the character-string conditions were included only as supplementary baselines.

The fMRI experiment included six runs of 9.9 min each, employing a block design. In the first and last 10 s of each run, the participants were shown a fixation. Each run contained four blocks for each condition, with interblock intervals of 10 s. The order of blocks of different conditions was counterbalanced across runs and participants. Each block contained a cue and four trials, lasting 20 s in total. The cue was presented for 1 s, indicating whether the following stimuli were sentences, word lists or character lists. The structure of a trial is shown in Fig. 1b. Words and character pairs were presented one by one for 500 ms each. A fixation of 300 ms appeared after the last word or character pair, followed by a probe word or character pair appearing for 1.35 s. As soon as the participants saw the probe word or character pair, they were asked to judge whether the probe stimulus had been presented within the current trial quickly and accurately. Each trial ended with a fixation of 350 ms.

Experiment 2. Experiment 2 included two sentence conditions (the HSS and NSS conditions) and two word-list conditions (the HSWL and NSWL conditions). The HSS and NSS conditions each contained 60 different sentences. Five independent rating experiments (each recruiting 16 participants who did not participate in the fMRI experiment) were conducted to obtain the socialness, imageability, semantic familiarity, semantic plausibility and syntactic plausibility of the sentences. The HSS and NSS conditions were significantly different in the socialness of sentence meaning (HSS > NSS) and were matched on all the other ratings (Supplementary Table 18). For each sentence, the maximum depth and mean depth of the syntactic nodes were calculated on the basis of the bottom-up node tree obtained from the Chinese Stanford Parser (https://nlp.stanford.edu/software/lex-parser.shtml), serving as two measures of syntactic complexity. The HSS and NSS conditions were matched on both measures (Supplementary Table 18). In addition, the HSS and NSS conditions were matched on character number, word number and mean log-transformed word frequency (Supplementary Table 18). For both the HSS and NSS conditions, the 60 sentences were grouped into 15 blocks (each containing 4 sentences). The constituent words of each block were then shuffled and rearranged into four word lists to constitute the stimuli for the HSWL and NSWL conditions.

The fMRI experiment included four runs of 7.1 min each, employing a block design. In the first and last 10 s of each run, the participants were shown a fixation. Each run contained four blocks for three conditions and three blocks for the other conditions, with interblock intervals of 10 s. The number and order of blocks for different conditions were counterbalanced across runs and participants. Each block contained a cue and four trials. The cue was presented for 1 s, indicating whether the following stimuli were sentences or word lists. The trial structure was the same as that of Experiment 1, except that each sentence or word list was presented for 4 s, with each constituent word within a trial having an equal length of presentation time (Fig. 2b).

Experiment 3. Experiment 3 employed a delayed word-recognition task, in which the participants were asked to read sentences, maintain them for a period and then perform word-recognition judgement by pressing buttons. We manipulated the socialness and number of stimuli (two or four sentences) to create four experimental conditions: the HSHML, HSLML, NSHML and NSLML conditions. Each condition contained 32 trials. The stimuli were the 96 high-socialness and 96 non-social sentences used in Experiment 1. Each sentence appeared in two different trials. As in Experiment 1, we matched the imageability, semantic familiarity and semantic plausibility of the sentences and the log-transformed word frequency across conditions (Supplementary Table 19).

The fMRI experiment included four runs of 12 min each, employing an event-related design. Each run included 32 trials, with 8 trials for each condition. The numbers and orders of the trials for different conditions were counterbalanced across runs and participants. In the first and last 10 s of each run, the participants were shown a fixation. In each trial, the encoding, maintenance and recognition stages lasted 7, 6 and 3 s, respectively. The three stages were separated by two jitter intervals, each 0.5 to 2.5 s, with an average duration of 1.5 s (Fig. 3b).

Experiment 4. Experiment 4 employed a 'mental portrait' task, in which the participants read two sentences describing either two trait features (dominance and trustworthiness) or two physical facial features (big/small face and long/short eyebrows) of a person successively and then chose a photograph from two to match the contents of the sentences. For the trait dimension of dominance, the participants saw either the sentence 'TA喜欢领导和指挥别人' (meaning 'He or she likes to lead and command others'), indicating high dominance, or 'TA喜欢追随和配合别人' (meaning 'He or she likes to follow and obey others'), indicating low dominance. For the trait dimension of trustworthiness, the participants saw either the sentence 'TA是一个诚恳耿直的人' (meaning 'He or she is a sincere and straightforward person'), indicating high trustworthiness, or 'TA是一个圆滑善变的人' (meaning 'He or she is a smooth and changeable person'), indicating low trustworthiness. For the two physical facial dimensions, the participants saw 'TA是一个脸盘较大/小的人' (meaning 'He or she is a person with a big/small face') and 'TA是一个眉毛较长/短的人' (meaning 'He or she is a person with long/short eyebrows'). We chose the two physical facial dimensions on the basis of the findings of Vernon et al.[98] and the results of rating experiments on our stimuli (see below), both of which indicate that the correlations between dominance, trustworthiness and the two selected physical facial features are very low. Note that for each of the four dimensions, the two sentences describing the different poles of the dimension were identical in syntactic structure, avoiding confounding between the semantic and syntactic differences.

The different orders and contents of the sentences resulted in eight trait and eight physical facial conditions. The social conditions were labelled HDHT, HDLT, LDHT, LDLT, HTHD, HTLD, LTHD and LTLD, in which the letters HD, LD, HT and LT indicate high dominance, low dominance, high trustworthiness and low trustworthiness, respectively. The physical facial conditions were labelled BFLE, BFSE, SFLE, SFSE, LEBF, LESF, SEBF and SESF, in which BF, SF, LE and SE indicate big face, small face, long eyebrows and short eyebrows, respectively. The picture stimuli were 128 photographs selected from the CAS-PEAL Chinese face database[99,100]. Each picture shows an authentic portrait photo of a person facing forwards, with a neutral expression and without a hat, glasses or any facial accessories. Four rating experiments (each recruiting 16 participants who did not participate in the fMRI experiment) were conducted to rate each photograph on the two traits and two physical facial dimensions using 1–100 scales. The results showed that the correlations between the z-transformed scores of the photographs on any two of the four dimensions were low ($|r| < 0.15$). The photographs were then grouped into 64 pairs. Each pair of photographs was included in two trait conditions that were opposite in both trait dimensions (for example, HDHT and LDLT) and two physical facial conditions that were opposite in both physical facial dimensions (for example, BFLE and SFSE). Each condition therefore contained 16 pairs of photographs, corresponding to 16 different trials of the condition.

The fMRI experiment included eight runs of 436 s each, employing an event-related design. In the first and last 10 s of each run, the participants were shown a fixation. Each run included 32 trials, with 2 trials for each condition. The orders of the trials of different conditions were counterbalanced across runs and participants. The trial structure is shown in Fig. 4a. Each trial started with a red fixation of 0.2 s. Then, the first and second sentences appeared in turn, each lasting for 1.5 s, followed by a jitter fixation of 1.5 to 3.5 s (mean, 2.5 s). The photographs were shown for 2 s, during which the participants were asked to make their judgement by pressing buttons. Each trial ended with a jitter fixation of 1.3 to 4.3 s (mean, 2.8 s).

Experiment 5. Experiment 5 included three conditions, each of which contained 30 short silent videos. The videos were obtained from online resources[101]. All videos were cut to 5 s long. We selected the stimuli for the three conditions on the basis of the number of people shown in the video; the HS videos contained multiple people who are interacting, the SP videos contained a single person and the NS videos contained no people.

The fMRI experiment included a single run of 11 min 16 s, employing a block design. In the first and last 10 s of each run, the participants were shown a fixation. There were six blocks for each condition. Each block contained five videos, each lasting 5 s, and four inter-stimulus fixations of 500 ms (see Fig. 5b for the block structure). The participants were asked to rate their pleasantness while watching the video by pressing one of four number buttons (1, very unpleased; 4, very pleased).

Experiment 6. Experiment 6 collected resting-state fMRI data using a single run lasting 8 min. During the fMRI scanning, the participants were asked to look at a white cross in the centre of a black screen.

Image acquisition and preprocessing

Structural and functional data were collected using a GE Discovery MR750 3T scanner at the Magnetic Resonance Imaging Research Center of the Institute of Psychology of the Chinese Academy of Sciences. For all experiments, T1-weighted structural images were obtained using a spoiled gradient-recalled pulse sequence in 176 sagittal slices with 1.0 mm isotropic voxels. From Experiments 1 to 5, functional BOLD data were collected using a gradient-echo echo-planar imaging sequence in 42 near-axial slices (repetition time, 2 s; echo time, 30 ms; flip angle, 70°; matrix size, 64 × 64; voxel size, 3.0 mm × 3.0 mm × 3.0 mm). In Experiment 6, functional BOLD data were collected using a gradient-echo echo-planar imaging sequence in 33 axial slices (repetition time, 2 s; echo time, 30 ms; flip angle, 90°; matrix size, 64 × 64; voxel size, 3.5 mm × 3.5 mm × 4.2 mm).

The fMRI data were preprocessed using the Statistical Parametric Mapping software (SPM12; http://www.fil.ion.ucl.ac.uk/spm/) and the advanced edition of DPARSF v.4.3 (ref. 102) implemented in DPABI v.3.0 (ref. 103). For the preprocessing of the task fMRI data, the first five volumes of each functional run were discarded to reach signal equilibrium. Slice timing and 3D head-motion correction were performed. Subsequently, a mean functional image was obtained for each participant, and the structural image of each participant was coregistered to the mean functional image. Thereafter, the structural image was segmented using a unified segmentation module[104]. Next, a custom, study-specific template was generated by applying diffeomorphic anatomical registration through exponentiated lie algebra (DARTEL[105]). The parameters obtained during segmentation were used to normalize the functional images of each participant into the Montreal Neurological Institute space by applying the deformation field estimated by segmentation. The functional images were subsequently spatially smoothed using a 6 mm full-width-half-maximum (FWHM) Gaussian kernel for univariate analysis but not for MVPA.

For the preprocessing of the resting-state fMRI data, after the same procedure for univariate analysis, linear trends were removed to reduce the effects of low-frequency drifts. The effects of nuisance variables, including 24 rigid head-motion parameters[106,107], white matter signal and cerebrospinal fluid signal, were removed by linear regression from each voxel's time course. Temporal bandpass filtering (0.01–0.1 Hz) was performed to reduce the effects of high-frequency noises.

Data analysis

For the behavioural analyses and results of Experiments 1 to 5, please see the Supplementary Information (section D) The fMRI data analyses were all conducted using SPM12 unless specifically stated.

Univariate analysis (Experiments 1, 2, 3 and 5). First-level analysis. At the first level, general linear model (GLM) analyses were performed to explore the fixed effect of each regressor for each participant. In Experiments 1, 2 and 5, which used a block design, each condition was modelled as a regressor (Experiment 1: HSS, NSS, HSWL, NSWL, HSCL and NSCL; Experiment 2: HSS, NSS, HSWL and NSWL; Experiment 5: HS, SP and NS), resulting in six, four and three regressors, respectively. In Experiment 3, for each of the four conditions (HSHML, HSLML, NSHML and NSLML), the three stages (encoding, maintenance and recognition) and the jitter intervals before and after the maintenance stage were modelled as 4 regressors of interest, resulting in 16 regressors. The above regressors were modelled with boxcar waveforms convolved with the canonical haemodynamic response function. In addition, for each GLM, six head-motion parameters obtained by head-motion correction were included as nuisance regressors, and a high-pass filter (128 s) was used to remove low-frequency signal drift for each run.

Second-level analysis. The estimated β-maps for each regressor obtained from the first-level analysis were entered into second-level (between-participant) random-effect analysis. For Experiments 1, 2 and 3, flexible factorial models were applied to accommodate their multifactor designs (Experiment 1, 2 × 3 within-participant design; Experiment 2, 2 × 2 within-participant design; Experiment 3, 2 × 2 within-participant design for each of the three trial stages). Contrasts of interest were examined using within-participant t-tests (two-tailed). For Experiments 1 and 2, the contrasts included '[(HSS + HSWL) − (NSS + NSWL)] / [(HSS + HSCL) − (NSS + NSCL)]' (the main effect of socialness), '[(HSS + NSS) − (HSWL + NSWL)] / [(HSS + NSS) − (HSCL + NSCL)]' (the main effect of sentence sensitivity), '[(HSS − HSWL) − (NSS − NSWL)] / [(HSS − HSCL) − (NSS − NSCL)]' (the interaction of socialness and sentence sensitivity) and '(HSS − HSWL) / (NSS − NSWL) / (HSS − HSCL) / (NSS − NSCL)' (the sentence sensitivity effect in the high-socialness and non-social conditions). For Experiment 3, at both the encoding and maintenance stages, the contrasts included '(HSHML + HSLML) − (NSHML + NSLML)' (the main effect of socialness), '(HSHML + NSHML) − (HSLML + NSLML)' (the main effect of cognitive demand) and '(HSHML − NSHML) − (HSLM L − NSLML)' (the interaction of socialness and cognitive demand). For Experiment 5, after fitting the GLM, contrasts between each two of the three conditions were computed for every participant. The contrast images were then entered into one-sample t-tests. For whole-brain analysis, multiple comparison corrections were conducted using cluster-level family-wise error correction ($P < 0.05$) as implemented in SPM12 (voxel-wise $P < 0.001$).

ROI analysis. We conducted ROI analysis to obtain more precise information about the involvement of the left vTPJ and lATL in social-semantic and sentence processing, addressing the issue of low statistical power caused by multiple comparison correction in whole-brain analysis. For ROI analysis, we used two methods to define the left vTPJ and lATL areas sensitive to sentences. The first way was to define the target ROIs on the basis of a previously published meta-analysis[27]. Zaccarella et al.[27] reported 11 peak Montreal Neurological Institute coordinates where sentences induced reliably stronger activity than word lists, among which we selected the coordinates of −44 −56 18 and −54 −4 −22 to represent the left vTPJ and lATL, respectively. We chose these coordinates because they are the most consistent with the anatomical positions 'vTPJ' and 'lATL'. In addition, these coordinates are very close to the peak coordinates reported by recently published meta-analyses on social-semantic processing[12,42]. For each coordinate, the ROI was defined as a 6-mm-radius sphere centred on it.

The second way to define the target ROIs was based on individual data. This method was applicable to Experiments 1 and 2 because the task used in these experiments can serve as a localizer for the brain areas sensitive to sentences. The localizing method was modified from the method proposed by Fedorenko et al.[41]. Fedorenko et al. provided a set of group-constrained masks for the areas involved in language processing (http://web.mit.edu/evlab/funcloc/). Because the masks covered broader regions than the left vTPJ and lATL, we overlapped the original group-constrained masks with a social-constrained map to obtain the neural overlaps between language and social cognition in the left vTPJ and lATL. The social-constrained map was defined by the Neurosynth meta-analysis (neurosynth.org) using the term 'social' as the keyword (association test; false discovery rate criterion of 0.01). We obtained four overlapping clusters larger than 100 voxels (voxel size, 3.0 mm × 3.0 mm × 3.0 mm), located in the left vTPJ, lATL, dmPFC and right vTPJ (which are located within Fedorenko et al.'s group-constrained masks of the left posterior temporal lobe, left anterior temporal lobe, left superior frontal gyrus and right middle-posterior temporal lobe, respectively). We therefore used these clusters as our group-constrained masks, which were used to define the two target individual ROIs in the left vTPJ and lATL and two supplementary individual ROIs in the other two regions. For each participant, the data from the first half of scanning were used to define the individual ROIs. The group-constrained masks were intersected with the participant's unthresholded t-maps of the contrast of 'sentence > word-list' (that is, HSS + NSS > HSWL + NSWL). (For the supplemental analysis using character-list conditions as the baseline conditions, the ROIs were defined using the contrasts of 'sentence > character-list' (that is, HSS + NSS > HSCL + NSCL).) For each participant, the 10% of voxels with the highest t values in each mask were chosen as the individual ROI[108]. The data from the second half of scanning were then used to conduct the statistical analyses at the ROI level. We also defined a control individual ROI in the left dorsal IFG to replicate the syntactic-driven sentence effect in this area[25,27], which should not be influenced by the socialness of sentence meaning. The left dorsal IFG is the most reliable syntax-sensitive region in the literature[46,47], and it has basically no overlap with the social brain. To eliminate the influence of social-semantic processing on the definition of individual control ROIs, we defined them by contrasting the NSS and NSWL conditions on half of the data. We also excluded 85 voxels from the group-constrained mask of the left IFG (which originally contained 1,014 voxels) due to their proximity to the social-constrained map, where the distance between them was less than 6 mm. Using an exclusion criterion of 6 mm was meant to account for the impact of spatial smoothing on fMRI data. We present the results for the target ROIs in the main text and those for the supplementary and control ROIs in the Supplementary Information (Supplementary Figs. 2 and 3 and Supplementary Tables 7 and 8).

In the ROI analysis, for each participant, the estimated β values for each regressor obtained from the GLM analysis were averaged across all voxels within each ROI. For Experiments 1 and 2, the influence of IES was regressed out from the β values for each condition and participant. Specifically, for each ROI, a linear mixed model was fit to the participant's β value using the lme4 package (version 1.1-30)[109] in R (version 4.2.1)[110]. This model included IES as the fixed effect and participant as a random factor with only a random intercept. The residuals were obtained and then entered into the contrast analysis. All contrasts of interest were identical to those of the whole-brain analysis and were examined using both Bayesian and classical parametric t-tests (two-tailed) in R. The Bayesian tests were based on the BayesFactor package (v.0.9.12-4.3)[111], with a default Cauchy prior width of $r = 0.707$ for effect size on the alternative hypothesis (H1)[112]. Classical parametric t-tests (two-tailed) were conducted as a supplementary statistic method.

DCM analysis (Experiment 3). We performed DCM analysis in Experiment 3, using DCM12 (ref. 113) in the SPM12 software. The ROIs of the left vTPJ and lATL were defined as in the ROI analysis of Experiment 3. Because these two regions are highly similar in their functional properties[59], we assumed that experimental input enters the model through both regions. We constructed nine models with the two regions, considering all possible combinations of directional connections and modulations between the two regions—that is, no connection, only vTPJ-to-lATL connection, only lATL-to-vTPJ connection or bilinear connections, combined with only intrinsic connection or connection modulated by social-semantic inputs (Supplementary Fig. 5a). The data for the encoding and maintenance stages were investigated separately. For each stage, the nine models, representing nine competing hypotheses, were compared using random-effect Bayesian model selection, and the winning model was defined as the one with the highest exceedance probability.

We conducted a further analysis by considering the possibility that experimental input enters the model through the left vTPJ, the left lATL or both. This resulted in nine model families, each of which comprised three different models that shared the same hypothesis about effective connectivity (Supplementary Fig. 5c). We calculated the sum of exceedance probabilities for the three models within each family and then compared these probabilities across the nine model families.

MVPA (Experiment 4). First-level analysis. In Experiment 4, the first-level analysis contained two steps. The first step was GLM analysis. We built 8 regressors (corresponding to the 8 different sentences of our stimuli; Fig. 4b) for the encoding stage of the first sentence, 16 regressors for the encoding stage of the second sentence (corresponding to the 16 sentence combinations—that is, HDHT, HDLT, LDHT, LDLT, HTHD, HTLD, LTHD, LTLD, BFLE, BFSE, SFLE, SFSE, LEBF, LESF, SEBF and SESF) and 16 regressors for the response stage. These regressors were all convolved with the canonical haemodynamic response function. In addition, six head-motion parameters were included as nuisance regressors, and a high-pass filter (128 s) was used to remove low-frequency signal drift for each run.

The second step was MVPA. We conducted both whole-brain searchlight MVPA and ROI-based MVPA. All classification procedures at both the whole-brain and ROI levels were implemented by the e1071 package[114] and custom script in R[110]. Whole-brain searchlight analysis was conducted within a group-based grey mask. To obtain the mask, the normalized structural image was segmented into different tissues for each participant. The resulting grey matter probabilistic images were resliced to the same spatial resolution as that of the functional image, averaged across participants and thresholded at 0.25 to generate a binary mask for searchlight mapping. For each voxel within the grey matter mask, support vector machine (SVM) decoding was conducted within a $5 \times 5 \times 5$ voxel cube centred at that voxel using the leave-one-run-out cross-validation approach[115]. For the encoding stages of both sentences, we trained four classifiers to discriminate the poles of the four dimensions (HD or LD, HT or LT, BF or SF, and LE or SE) described in the current sentence. For the encoding stages of the second sentence, we additionally trained four classifiers to discriminate the poles of the four dimensions described in the context sentence (the first sentence). Before the SVM decoding was conducted, β values within a cube were normalized to remove the common response pattern by subtracting the mean across the conditions to be discriminated. The resulting accuracy images were smoothed using a 6 mm FWHM Gaussian kernel for subsequent second-level statistical analyses.

ROI-based MVPA was conducted within the ROIs identical to those used in Experiment 3. After we fit the GLM, for each regressor of the encoding stage of the first and second sentences in each run, the estimated β values of all voxels within a given ROI mask were normalized and concatenated to form an fMRI pattern vector. On the basis of these fMRI pattern vectors, SVM decoding was conducted to discriminate the poles of the four dimensions described in the current or last sentences, just as in the whole-brain searchlight cubes.

Second-level analysis. For whole-brain searchlight MVPA, the second-level statistical analysis was conducted to examine whether the classification performance for each dimension within each cube was above the chance level using one-tailed one-sample t-tests. For ROI-level MVPA, the participant-wise bootstrapping method was conducted to obtain the statistical significance of the classification performance for each dimension. For each round of bootstrapping iteration, the dataset was resampled with replacement to create a pseudo-sample keeping the original sample size, and the mean classification accuracy of the group was calculated. This procedure was repeated 5,000 times to form a sampling distribution for each classification. The null distribution of each classification was generated by subtracting the veritable accuracy from the sampling distribution, and the veritable accuracy was then ranked against the null distribution to calculate the P value.

RSFC analyses (Experiment 6). The RSFC analysis included 15 seed ROIs in total, among which 2 key ROIs (that is, the left vTPJ and lATL) and 9 sentence-processing ROIs were defined on the basis of the meta-analysis results of Zaccarella et al.[27] and 4 social-semantic-processing ROIs were defined on the basis of the meta-analysis results of Zhang et al.[42]. For each pair of seed ROIs, each participant's mean time series for each seed ROI was calculated and correlated with each other. The correlation coefficients were then Fisher-transformed to represent the RSFC. We conducted two analyses to examine whether the left vTPJ and lATL have stronger RSFC to the social-semantic or sentence-processing areas. In the first analysis, for each key ROI, we compared its mean RSFC to the social-semantic-processing ROIs with that to the sentence-processing ROIs across participants using both Bayesian and classical parametric t-tests (two-tailed). In the second analysis, the mean RSFC matrix of these 15 seed ROIs was transformed back to correlation coefficients and then applied with k-means clustering to group them into two to ten clusters. The ideal number of clusters was selected on the basis of the highest silhouette score.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data that support the findings of this study are available from Psychological Science Bank (https://doi.org/10.57760/sciencedb.psych.00138).

Code availability

Custom code that supports the findings of this study is available from Psychological Science Bank (https://doi.org/10.57760/sciencedb.psych.00138).

References

1. de Villiers, J. The interface of language and theory of mind. *Lingua* **117**, 1858–1878 (2007).
2. Richardson, H. et al. Reduced neural selectivity for mental states in deaf children with delayed exposure to sign language. *Nat. Commun.* **11**, 3246 (2020).
3. Scott, S. K. From speech and talkers to the social world: the neural processing of human spoken language. *Science* **366**, 58–62 (2019).
4. Dunbar, R. I. M. Gossip in evolutionary perspective. *Rev. Gen. Psychol.* **8**, 100–110 (2004).
5. Bzdok, D. et al. Left inferior parietal lobe engagement in social cognition and language. *Neurosci. Biobehav. Rev.* **68**, 319–334 (2016).
6. Mar, R. A. The neural bases of social cognition and story comprehension. *Annu. Rev. Psychol.* **62**, 103–134 (2011).
7. Mellem, M. S., Jasmin, K. M., Peng, C. & Martin, A. Sentence processing in anterior superior temporal cortex shows a social–emotional bias. *Neuropsychologia* **89**, 217–224 (2016).
8. Diveica, V., Koldewyn, K. & Binney, R. J. Establishing a role of the semantic control network in social cognitive processing: a meta-analysis of functional neuroimaging studies. *NeuroImage* **245**, 118702 (2021).
9. Schurz, M. et al. Toward a hierarchical model of social cognition: a neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychol. Bull.* **147**, 293–327 (2021).
10. Binney, R. J. & Ramsey, R. Social semantics: the role of conceptual knowledge and cognitive control in a neurobiological model of the social brain. *Neurosci. Biobehav. Rev.* **112**, 28–38 (2020).
11. Pexman, P. M., Diveica, V. & Binney, R. J. Social semantics: the organization and grounding of abstract concepts. *Phil. Trans. R. Soc. B* **378**, 20210363 (2022).
12. Arioli, M., Gianelli, C. & Canessa, N. Neural representation of social concepts: a coordinate-based meta-analysis of fMRI studies. *Brain Imaging Behav.* **15**, 1912–1921 (2021).
13. Lin, N. et al. Dissociating the neural correlates of the sociality and plausibility effects in simple conceptual combination. *Brain Struct. Funct.* **225**, 995–1008 (2020).
14. Zhang, G., Hung, J. & Lin, N. Coexistence of the social semantic effect and non-semantic effect in the default mode network. *Brain Struct. Funct.* **228**, 321–339 (2023).
15. Zahn, R. et al. Social concepts are represented in the superior anterior temporal cortex. *Proc. Natl Acad. Sci. USA* **104**, 6430–6435 (2007).
16. Tamir, D. I., Thornton, M. A., Contreras, J. M. & Mitchell, J. P. Neural evidence that three dimensions organize mental state representation: rationality, social impact, and valence. *Proc. Natl Acad. Sci. USA* **113**, 194–199 (2016).
17. Contreras, J. M., Banaji, M. R. & Mitchell, J. P. Dissociable neural correlates of stereotypes and other forms of semantic knowledge. *Soc. Cogn. Affect. Neurosci.* **7**, 764–770 (2012).
18. Saxe, R. & Wexler, A. Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia* **43**, 1391–1399 (2005).
19. Lin, N., Bi, Y., Zhao, Y., Luo, C. & Li, X. The theory-of-mind network in support of action verb comprehension: evidence from an fMRI study. *Brain Lang.* **141**, 1–10 (2015).
20. Lin, N. et al. Fine subdivisions of the semantic network supporting social and sensory–motor semantic processing. *Cereb. Cortex* **28**, 2699–2710 (2018).
21. Spunt, R. P., Kemmerer, D. & Adolphs, R. The neural basis of conceptualizing the same action at different levels of abstraction. *Soc. Cogn. Affect. Neurosci.* **11**, 1141–1151 (2016).
22. Lin, N. et al. Coin, telephone, and handcuffs: neural correlates of social knowledge of inanimate objects. *Neuropsychologia* **133**, 107187 (2019).
23. Dronkers, N. F. et al. Lesion analysis of the brain areas involved in language comprehension. *Cognition* **92**, 145–177 (2004).
24. Fedorenko, E., Behr, M. K. & Kanwisher, N. Functional specificity for high-level linguistic processing in the human brain. *Proc. Natl Acad. Sci. USA* **108**, 16428–16433 (2011).
25. Pallier, C., Devauchelle, A.-D. & Dehaene, S. Cortical representation of the constituent structure of sentences. *Proc. Natl Acad. Sci. USA* **108**, 2522–2527 (2011).
26. Malik-Moraleda, S. et al. An investigation across 45 languages and 12 language families reveals a universal language network. *Nat. Neurosci.* **25**, 1014–1019 (2022).
27. Zaccarella, E., Schell, M. & Friederici, A. D. Reviewing the functional basis of the syntactic Merge mechanism for language: a coordinate-based activation likelihood estimation meta-analysis. *Neurosci. Biobehav. Rev.* **80**, 646–656 (2017).
28. Fedorenko, E. & Thompson-Schill, S. L. Reworking the language network. *Trends Cogn. Sci.* **18**, 120–126 (2014).
29. Labache, L. et al. A SENtence Supramodal Areas AtlaS (SENSAAS) based on multiple task-induced activation mapping and graph analysis of intrinsic connectivity in 144 healthy right-handers. *Brain Struct. Funct.* **224**, 859–882 (2019).
30. Matchin, W., Brodbeck, C., Hammerly, C. & Lau, E. The temporal dynamics of structure and content in sentence comprehension: evidence from fMRI-constrained MEG. *Hum. Brain Mapp.* **40**, 663–678 (2019).
31. Humphries, C., Binder, J. R., Medler, D. A. & Liebenthal, E. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *J. Cogn. Neurosci.* **18**, 665–679 (2006).
32. Price, C. J. The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann. N. Y. Acad. Sci.* **1191**, 62–88 (2010).
33. Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R. & Haynes, J.-D. The distributed nature of working memory. *Trends Cogn. Sci.* **21**, 111–124 (2017).
34. D'Esposito, M. & Postle, B. R. The cognitive neuroscience of working memory. *Annu. Rev. Psychol.* **66**, 115–142 (2015).
35. Fuster, J. M. Network memory. *Trends Neurosci.* **20**, 451–459 (1997).
36. Postle, B. R. Working memory as an emergent property of the mind and brain. *Neuroscience* **139**, 23–38 (2006).
37. Cowan, N. *Attention and Memory: An Integrated Framework* (Oxford Univ. Press, 1998); https://doi.org/10.1093/acprof:oso/9780195119107.001.0001
38. Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E. & Sherman, J. Pictures in sentences: understanding without words. *J. Exp. Psychol. Gen.* **115**, 281–294 (1986).
39. Potter, M. C. Very short-term conceptual memory. *Mem. Cogn.* **21**, 156–161 (1993).
40. Potter, M. Conceptual short term memory in perception and thought. *Front. Psychol.* **3**, 113 (2012).
41. Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S. & Kanwisher, N. New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *J. Neurophysiol.* **104**, 1177–1194 (2010).
42. Zhang, G., Xu, Y., Zhang, M., Wang, S. & Lin, N. The brain network in support of social semantic accumulation. *Soc. Cogn. Affect. Neurosci.* **16**, 393–405 (2021).
43. Kuhnke, P. et al. The role of the angular gyrus in semantic cognition: a synthesis of five functional neuroimaging studies. *Brain Struct. Funct.* **228**, 273–291 (2023).
44. Townsend, J. T. & Ashby, F. G. *Stochastic Modeling of Elementary Psychological Processes* (Cambridge Univ. Press, 1983).
45. Paunov, A. M. et al. Differential tracking of linguistic vs. mental state content in naturalistic stimuli by language and theory of mind (ToM) brain networks. *Neurobiol. Lang.* **3**, 413–440 (2022).

46. Hagoort, P. & Indefrey, P. The neurobiology of language beyond single words. *Annu. Rev. Neurosci.* **37**, 347–362 (2014).
47. Heard, M. & Lee, Y. S. Shared neural resources of rhythm and syntax: an ALE meta-analysis. *Neuropsychologia* **137**, 107284 (2020).
48. Sreenivasan, K. K. & D'Esposito, M. The what, where and how of delay activity. *Nat. Rev. Neurosci.* **20**, 466–481 (2019).
49. Postle, B. R. The cognitive neuroscience of visual short-term memory. *Cogn. Control* **1**, 40–46 (2015).
50. Manoach, D. S. et al. Prefrontal cortex fMRI signal changes are correlated with working memory load. *NeuroReport* **8**, 545–549 (1997).
51. Rottschy, C. et al. Modelling neural correlates of working memory: a coordinate-based meta-analysis. *NeuroImage* **60**, 830–846 (2012).
52. Druzgal, T. J. & D'Esposito, M. Activity in fusiform face area modulated as a function of working memory load. *Cogn. Brain Res.* **10**, 355–364 (2001).
53. Meyer, M. L., Taylor, S. E. & Lieberman, M. D. Social working memory and its distinctive link to social cognitive ability: an fMRI study. *Soc. Cogn. Affect. Neurosci.* **10**, 1338–1347 (2015).
54. Xu, Y. & Chun, M. M. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* **440**, 91–95 (2006).
55. Martin, R. C., Wu, D., Freedman, M., Jackson, E. F. & Lesch, M. An event-related fMRI investigation of phonological versus semantic short-term memory. *J. Neurolinguistics* **16**, 341–360 (2003).
56. Song, J.-H. & Jiang, Y. Visual working memory for simple and complex features: an fMRI study. *NeuroImage* **30**, 963–972 (2006).
57. Thornton, M. A. & Conway, A. R. A. Working memory for social information: chunking or domain-specific buffer? *NeuroImage* **70**, 233–239 (2013).
58. Zhao, Y., Kuai, S., Zanto, T. P. & Ku, Y. Neural correlates underlying the precision of visual working memory. *Neuroscience* **425**, 301–311 (2020).
59. Amft, M. et al. Definition and characterization of an extended social–affective default network. *Brain Struct. Funct.* **220**, 1031–1049 (2015).
60. Wang, Y. et al. A large-scale structural and functional connectome of social mentalizing. *NeuroImage* **236**, 118115 (2021).
61. Oosterhof, N. N. & Todorov, A. The functional basis of face evaluation. *Proc. Natl Acad. Sci. USA* **105**, 11087–11092 (2008).
62. Olson, I. R., McCoy, D., Klobusicky, E. & Ross, L. A. Social cognition and the anterior temporal lobes: a review and theoretical framework. *Soc. Cogn. Affect. Neurosci.* **8**, 123–133 (2013).
63. Dunbar, R. I. M., Marriott, A. & Duncan, N. D. C. Human conversational behavior. *Hum. Nat.* **8**, 231–246 (1997).
64. Branco, P., Seixas, D. & Castro, S. L. Mapping language with resting-state functional magnetic resonance imaging: a study on the functional profile of the language network. *Hum. Brain Mapp.* **41**, 545–560 (2020).
65. Bedny, M., Pascual-Leone, A., Dodell-Feder, D., Fedorenko, E. & Saxe, R. Language processing in the occipital cortex of congenitally blind adults. *Proc. Natl Acad. Sci. USA* **108**, 4429–4434 (2011).
66. Ferstl, E. C., Neumann, J., Bogler, C. & von Cramon, D. Y. The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Hum. Brain Mapp.* **29**, 581–593 (2008).
67. Hagoort, P. The neurobiology of language beyond single-word processing. *Science* **366**, 55–58 (2019).
68. Lin, N. et al. Neural correlates of three cognitive processes involved in theory of mind and discourse comprehension. *Cogn. Affect. Behav. Neurosci.* **18**, 273–283 (2018).
69. Schurz, M., Radua, J., Aichhorn, M., Richlan, F. & Perner, J. Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neurosci. Biobehav. Rev.* **42**, 9–34 (2014).
70. Feng, W., Yu, H. & Zhou, X. Understanding particularized and generalized conversational implicatures: is theory-of-mind necessary? *Brain Lang.* **212**, 104878 (2021).
71. Lewis, P. A., Rezaie, R., Brown, R., Roberts, N. & Dunbar, R. I. M. Ventromedial prefrontal volume predicts understanding of others and social network size. *NeuroImage* **57**, 1624–1629 (2011).
72. Samson, D., Apperly, I. A., Chiavarino, C. & Humphreys, G. W. Left temporoparietal junction is necessary for representing someone else's belief. *Nat. Neurosci.* **7**, 499–500 (2004).
73. Wang, Y. et al. Dynamic neural architecture for social knowledge retrieval. *Proc. Natl Acad. Sci. USA* **114**, E3305–E3314 (2017).
74. Rapp, A. M., Mutschler, D. E. & Erb, M. Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies. *NeuroImage* **63**, 600–610 (2012).
75. Yang, H. & Bi, Y. From words to phrases: neural basis of social event semantic composition. *Brain Struct. Funct.* **227**, 1683–1695 (2022).
76. Wang, X., Wang, B. & Bi, Y. Close yet independent: dissociation of social from valence and abstract semantic dimensions in the left anterior temporal lobe. *Hum. Brain Mapp.* **40**, 4759–4776 (2019).
77. Dingemanse, M. et al. Beyond single-mindedness: a figure–ground reversal for the cognitive sciences. *Cogn. Sci.* **47**, e13230 (2023).
78. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).
79. Redcay, E., Velnoskey, K. R. & Rowe, M. L. Perceived communicative intent in gesture and language modulates the superior temporal sulcus. *Hum. Brain Mapp.* **37**, 3444–3461 (2016).
80. Weisberg, J., Hubbard, A. L. & Emmorey, K. Multimodal integration of spontaneously produced representational co-speech gestures: an fMRI study. *Lang. Cogn. Neurosci.* **32**, 158–174 (2017).
81. Hassabis, D. et al. Imagine all the people: how the brain creates and uses personality models to predict behavior. *Cereb. Cortex* **24**, 1979–1987 (2014).
82. Van Overwalle, F., Ma, N. & Baetens, K. Nice or nerdy? The neural representation of social and competence traits. *Soc. Neurosci.* **11**, 567–578 (2016).
83. Thornton, M. A. & Mitchell, J. P. Theories of person perception predict patterns of neural activity during mentalizing. *Cereb. Cortex* **28**, 3505–3520 (2018).
84. Martin, R. C., Ding, J., Hamilton, A. C. & Schnur, T. T. Working memory capacities neurally dissociate: evidence from acute stroke. *Cereb. Cortex Commun.* **2**, tgab005 (2021).
85. Yue, Q., Martin, R. C., Hamilton, A. C. & Rose, N. S. Non-perceptual regions in the left inferior parietal lobe support phonological short-term memory: evidence for a buffer account? *Cereb. Cortex* **29**, 1398–1413 (2019).
86. Yue, Q. & Martin, R. C. Maintaining verbal short-term memory representations in non-perceptual parietal regions. *Cortex* **138**, 72–89 (2021).
87. Meyer, M. L. & Collier, E. Theory of minds: managing mental state inferences in working memory is associated with the dorsomedial subsystem of the default network and social integration. *Soc. Cogn. Affect. Neurosci.* **15**, 63–73 (2020).
88. Bemis, D. K. & Pylkkänen, L. Simple composition: a magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *J. Neurosci.* **31**, 2801 (2011).
89. Bemis, D. K. & Pylkkänen, L. Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cereb. Cortex* **23**, 1859–1873 (2013).
90. Graves, W. W. et al. Correspondence between cognitive and neural representations for phonology, orthography, and semantics in supramarginal compared to angular gyrus. *Brain Struct. Funct.* **228**, 255–271 (2023).

91. Hung, J., Wang, X., Wang, X. & Bi, Y. Functional subdivisions in the anterior temporal lobes: a large scale meta-analytic investigation. *Neurosci. Biobehav. Rev.* **115**, 134–145 (2020).
92. Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E. & Gallant, J. L. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* **532**, 453–458 (2016).
93. Seghier, M. L. The angular gyrus: multiple functions and multiple subdivisions. *Neuroscientist* **19**, 43–61 (2013).
94. Price, A. R., Bonner, M. F., Peelle, J. E. & Grossman, M. Converging evidence for the neuroanatomic basis of combinatorial semantics in the angular gyrus. *J. Neurosci.* **35**, 3276–3284 (2015).
95. Schell, M., Zaccarella, E. & Friederici, A. D. Differential cortical contribution of syntax and semantics: an fMRI study on two-word phrasal processing. *Cortex* **96**, 105–120 (2017).
96. *Chinese Linguistic Data Consortium* (Tsinghua University, State Key Laboratory of Intelligent Technology and Systems, and Chinese Academy of Sciences, Institute of Automation, 2003).
97. Lin, N., Yu, X., Zhao, Y. & Zhang, M. Functional anatomy of recognition of Chinese multi-character words: convergent evidence from effects of transposable nonwords, lexicality, and word frequency. *PLoS ONE* **11**, e0149583 (2016).
98. Vernon, R. J. W., Sutherland, C. A. M., Young, A. W. & Hartley, T. Modeling first impressions from highly variable facial images. *Proc. Natl Acad. Sci. USA* **111**, E3353–E3361 (2014).
99. Gao, W. et al. The CAS-PEAL large-scale Chinese face database and baseline evaluations. *IEEE Trans. Syst. Man Cybern. A Syst. Hum.* **38**, 149–161 (2008).
100. *The CAS-PEAL Face Database* (ICT-ISVISION Joint Research & Development Laboratory for Face Recognition, 2008); http://www.jdl.link/peal/home.htm
101. *Free Stock Video Footage* (Videvo Team, 2021); https://www.videvo.net/free-stock-footage/
102. Yan, C. & Zang, Y. DPARSF: a MATLAB toolbox for 'pipeline' data analysis of resting-state fMRI. *Front. Syst. Neurosci.* **4**, 13 (2010).
103. Yan, C.-G., Wang, X.-D., Zuo, X.-N. & Zang, Y.-F. DPABI: Data Processing & Analysis for (Resting-State) Brain Imaging. *Neuroinformatics* **14**, 339–351 (2016).
104. Ashburner, J. & Friston, K. J. Unified segmentation. *NeuroImage* **26**, 839–851 (2005).
105. Ashburner, J. A fast diffeomorphic image registration algorithm. *NeuroImage* **38**, 95–113 (2007).
106. Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. J. & Turner, R. Movement-related effects in fMRI time-series. *Magn. Reson. Med.* **35**, 346–355 (1996).
107. Yan, C.-G. et al. A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *NeuroImage* **76**, 183–201 (2013).
108. Blank, I. A. & Fedorenko, E. No evidence for differences among language regions in their temporal receptive windows. *NeuroImage* **219**, 116925 (2020).
109. Bates, D., Maechler, M., Bolker, B. & Walker, S. lme4: Linear Mixed-Effects Models using 'Eigen' and S4 (version 1.1-30) R package (2007).
110. R Core Team. R: A Language and Environment for Statistical Computing (version 4.2.1) R Foundation for Statistical Computing (2020).
111. Morey, R. & Rouder, J. N. BayesFactor. R package v.0.9.12-4.3 (2015).
112. Rouder, J. N., Morey, R. D., Speckman, P. L. & Province, J. M. Default Bayes factors for ANOVA designs. *J. Math. Psychol.* **56**, 356–374 (2012).
113. Friston, K. J., Harrison, L. & Penny, W. Dynamic causal modelling. *NeuroImage* **19**, 1273–1302 (2003).
114. Meyer, D. et al. E1071: Misc Functions of the Department of Statistics, TU Wien (version 1.7-11) R package (2022).
115. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995).

Acknowledgements

This research was supported by grants from the National Natural Science Foundation of China (grant no. 31871105 to N.L.); the Scientific Foundation of the Institute of Psychology, Chinese Academy of Sciences (grant no. E2CX3625CX to X.W. and N.L.); the Scientific Foundation of the Institute of Psychology, Chinese Academy of Sciences (grant no. E1CX4725CX to X.W.); the National Science and Technology Innovation 2030 Major Program (grant no. 2021ZD0204104 to Y.B.); and the National Natural Science Foundation of China (grant no. 31925020 to Y.B.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the paper. Experiment 4 in this paper used the CAS-PEAL-R1 face database collected under the sponsor of the Chinese National Hi-Tech Program and ISVISION Tech. Co. Ltd. We thank X. Wang, H. Yang, H. Wen and W. Zhou for assistance in performing the MVPA and DCM analysis.

Author contributions

G.Z. and N.L. conceived the study. G.Z., N.L. and W.S. developed the methods. G.Z. performed the investigation and the data analysis. N.L. supervised the work. G.Z. and N.L. wrote the initial draft. All authors reviewed and edited the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-023-01704-8.

Correspondence and requests for materials should be addressed to Nan Lin.

Peer review information *Nature Human Behaviour* thanks Moritz Wurm, David Kemmerer and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

nature portfolio

Corresponding author(s): Nan Lin

Last updated by author(s): Aug 10, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used
Data analysis	The fMRI data were preprocessed using the Statistical Parametric Mapping software (SPM12; http://www.fil.ion.ucl.ac.uk/spm/) and the advanced edition of DPARSF V4.3 (Yan & Zang, 2010) implemented in DPABI V3.0 (Yan et al., 2016). After preprocessing, all statistical analyses were conducted using R (version 4.2.0). Bayesian analysis was conducted using BayesFactor package in R (version 0.9.12-4.3). All codes for analysis are available from Psychological Science Bank (https://doi.org/10.57760/sciencedb.psych.00138).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data for analysis are available from Psychological Science Bank (https://doi.org/10.57760/sciencedb.psych.00138).

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Our findings apply to both sexes and genders. Sex and gender were not considered in our study design. We performed no sex- or gender-based analyses, because there was no sufficient evidence indicating differences in neural correlates of social semantic working memory between sexes or genders.
Population characteristics	Participants were all right-handed and native Chinese speakers. None of them had experienced psychiatric or neurological disorders or had sustained a head injury. The sample sizes of Experiments 1, 2, 3, 4, 5 and 6 were 20 (16 women, M age = 22.3 years, SD age =2.3 years), 20 (13 women, M age = 23.5 years, SD age = 1.9 years), 20 (11 women, M age = 21.8 years, SD age = 2.4 years), 16 (9 women, M age = 24.0 years, SD age = 2.4 years), 20 (14 women, M age = 22.8 years, SD age = 2.7 years), and 39 (28 women, M age = 22.9 years, SD age = 2.2 years), respectively. There were 82 participants in total (56 women, M age = 22.7 years, SD age = 2.4 years).
Recruitment	All participants were recruited online from colleage students in Beijing. Participant should be right-handed and native Chinese speaker. None of them had experienced psychiatric or neurological disorders or had sustained a head injury. Each participant read and signed the informed consent form before taking part in the experiments. Due to the college student participants, the research results may not generalize to other populations (e.g., children).
Ethics oversight	Institutional Review Board of the Institute of Psychology of the Chinese Academy of Sciences

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This is a quantitative basic research involves human subjects
Research sample	Participants were Native Chinese college students in Beijing. There were 82 participants in total (56 women, M age = 22.7 years, SD age = 2.4 years). The sample sizes of Experiments 1, 2, 3, 4, 5 and 6 were 20 (16 women, M age = 22.3 years, SD age =2.3 years), 20 (13 women, M age = 23.5 years, SD age = 1.9 years), 20 (11 women, M age = 21.8 years, SD age = 2.4 years), 16 (9 women, M age = 24.0 years, SD age = 2.4 years), 20 (14 women, M age = 22.8 years, SD age = 2.7 years), and 39 (28 women, M age = 22.9 years, SD age = 2.2 years), respectively. College students, commonly used in psychology experiments, are representative of young, healthy populations, while they may not fully represent other groups (e.g., children).
Sampling strategy	Participants were all right-handed and native Chinese speakers. None of them had experienced psychiatric or neurological disorders or had sustained a head injury. Sample sizes were determined by referencing those of previous fMRI studies on social-semantic and sentence effects, which have been summarized in two meta-analyses, conducted by Zhang et al.(2021) and Zaccarella et al.(2017) respectively.
Data collection	In all experiments, participants' responses were recorded with a computer, while the ongoing brain activity during the task was recorded using a MRI scanner. No one was present in the room together with the participants during the experiments. The researcher was aware of the experimental conditions and the study hypothesis during data collection.
Timing	The data collection started July 2019 and ended in September 2021.
Data exclusions	No data were excluded from the analyses.
Non-participation	No participants declined participation or dropped out.
Randomization	Participants were not allocated into experimental groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study	
☒	☐	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology and archaeology
☒	☐	Animals and other organisms
☒	☐	Clinical data
☒	☐	Dual use research of concern

Methods

n/a	Involved in the study	
☒	☐	ChIP-seq
☒	☐	Flow cytometry
☐	☒	MRI-based neuroimaging

Magnetic resonance imaging

Experimental design

Design type

Experiment 1, 2, and 5 were task state with block design. Experiment 3 and 4 were task state with event-related design. Experiment 6 was resting state.

Design specifications

Experiment 1: there were 6 sessions for each participant; each session contained 24 blocks; each blocks contained 4 trials, lasting 20s in total; the interblock invertal was 10s.
Experiment 2: there were 4 sessions for each participant; each session contained 15 blocks; each blocks contained 4 trials, lasting 24s in total; the interblock invertal was 10s.
Experiment 3: there were 4 sessions for each participant; each session contained 32 trials; each trial lasted 19s in average; the intertrial invertal was 1.5-4.5s (M = 3s).
Experiment 4: there were 8 sessions for each participant; each session contained 32 trials; each trial lasted 10.2s in average; the intertrial invertal was 1.3-4.3s (M = 2.8s).
Experiment 5: there was a signle sessions for each participant, containing 18 blocks; each blocks contained 5 trials, lasting 27s in total; the interblock invertal was 10s.
Experiment 6: there was a single session for each participant to collect resting-state fMRI data, lasting 8 mins.

Behavioral performance measures

Experiment 1, 2, 3: We recorded button press (Yes/No) and reaction time. For both accuracy and reaction time, mean and standard deviation across participants in different experiment conditions were used to establish that the participants were performed the task as expected.
Experiment 4: We recorded button press to assess task performance. Because there was no theoretically correct response, inter-subject correlation were used to establish that the participants were performed the task as expected.
Experiment 5: We recorded button press of the pleasantness rating task. This task was set only to ensure that participants pay attention to the video stimuli. The task of interest was video watching, which does not require any behavioral response.
Experiment 6: Only resting-state fMRI data were collected and there was no behavioral task.

Acquisition

Imaging type(s)

Functional

Field strength

3 T

Sequence & imaging parameters

From Experiments 1 to 5, functional BOLD data were collected using a gradient-echo echo-planar imaging sequence in 42 near-axial slices (repetition time = 2 seconds; echo time = 30 milliseconds; flip angle = 70°; matrix size = 64 × 64; voxel size = 3.0 mm × 3.0 mm × 3.0 mm; image type = EPI). In Experiment 6, functional BOLD data were collected using a gradient-echo echo-planar imaging sequence in 33 axial slices (repetition time = 2 seconds, echo time = 30 milliseconds, flip angle = 90°, matrix size = 64 × 64; voxel size = 3.5 mm × 3.5 mm × 4.2 mm; image type = EPI).

Area of acquisition

A whole brain scan.

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software

The fMRI data were preprocessed using the Statistical Parametric Mapping software (SPM12; http://www.fil.ion.ucl.ac.uk/spm/) and the advanced edition of DPARSF V4.3 (Yan & Zang, 2010) implemented in DPABI V3.0 (Yan et al., 2016).

Normalization

For each participant, structural image was segmented using a unified segmentation module (Ashburner & Friston 2005). Next, a custom, study-specific template was generated by applying diffeomorphic anatomical registration through exponentiated lie algebra (DARTEL; Ashburner, 2007). The parameters obtained during segmentation were used to normalize the functional images of each participant

Normalization template

MNI305

Noise and artifact removal

For the preprocessing of the task fMRI data, the first five volumes of each functional run were discarded to reach signal equilibrium. Slice timing and 3-D head motion correction were performed. After normalization, the functional images were spatially smoothed using a 6-mm full-width-half-maximum Gaussian kernel for univariate analysis but not for multivariate pattern analysis. For the preprocessing of the resting-state fMRI data, after the same procedure for univariate analysis, linear trends were removed to reduce the effects of low-frequency drifts. The effects of nuisance variables, including 24 rigid head motion parameters (Friston, Williams, Howard, Frackowiak, & Turner, 1996; Yan et al., 2013), white matter signal, and cerebrospinal fluid signal, were removed by linear regression from each voxel's time course. Temporal bandpass filtering (0.01–0.1 Hz) was performed to reduce the effects of high-frequency noises.

Volume censoring

None

Statistical modeling & inference

Model type and settings

Experiment 1, 2, 3, 5: mass univariate; at the first level, general-linear-model (GLM) analyses were performed to explore the fixed effect of each regressor for each participant; in addition to regressors of interest, for each GLM, six head motion parameters obtained by head motion correction were included as nuisance regressors and a high-pass filter (128 seconds) was used to remove low-frequency signal drift for each run; the estimated beta-maps for each regressor obtained from the first-level analysis were entered into second-level (between-subject) random-effect analysis.
Experiment 4: multivariate pattern analysis (MVPA); GLM analysis was first performed to obtain results of each regressors; during GLM analysis, six head motion parameters were included as nuisance regressors, and a high-pass filter (128 seconds) was used to remove low-frequency signal drift for each run; then MVPA were performed; the results of MVPA were entered into second-level (between-subject) random-effect analysis.

Effect(s) tested

Experiment 1 and 2: 1) social semantic effect, (HSS+HSWL) - (NSS+NSWL); 2) sentence effects, (HSS+NSS) - (HSWL+NSWL); 3) interactions, (HSS-HSWL)-(NSS-NSWL); 4) sentence effects in high-socialness conditions, HSS-HSWL; 5) sentence effects in non-socialness conditions, NSS-NSWL.
Experiment 3: 1) social semantic effect, (HSHML+HSLML) - (NSHML+NSLML; 2) cognitive demanding effects, (HSHML +NSHML) - (HSLML+NSLML); 3) interaction of socialness and cognitive demand, (HSHML - NSHML) - (HSLML - NSLML).
Experiment 4: classification accuracy, (HD or LD) > 0.5, (HT or LT) > 0.5, (BF or SF) > 0.5, (LE or SE) > 0.5.
Experiment 5: social semantic effect, HS - NS, HS - SP, SP - NS.
Experiment 6: RSFC strength, (RSFC to social-semantic-processing ROIs) - (RSFC to sentence processing ROIs).
For experiment 1, 2, and 3, flexible factorial models were applied to accommodate their multifactor designs.
For all experiments, all contrasts were performed using one-sample t tests.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☒ Both

Anatomical location(s)

We used two methods to define the left vTPJ and lATL areas sensitive to sentences. The first way was to define the ROIs based on a previously published meta-analysis (Zaccarella et al., 2017). Zaccarella et al. (2017) reported 11 peak MNI coordinates where sentences induced reliably stronger activity than word lists, among which we selected the coordinates of -44 -56 18 and -54 -4 -22 to represent the left vTPJ and lATL, respectively. We chose these coordinates because they are most consistent with the anatomical positions "vTPJ" and "lATL". For each coordinate, the ROI was defined as a 6-mm radius sphere centered on it. The second way to define the ROIs was based on individual data. This method was applicable to Experiments 1 and 2 because the task used in these experiments can serve as a localizer for the brain areas sensitive to sentences. The localizing method was modified from the method proposed by Fedorenko et al. (2010). Fedorenko et al. (2010) provided a set of group-constrained masks for the areas involved in language processing (http://web.mit.edu/evlab/funcloc/). Because the masks covered broader regions than the left vTPJ and lATL, we overlapped the original group-constrained masks with a social-constrained map to obtain the neural overlaps between language and social cognition in the left vTPJ and lATL. The social-constrained map was defined by the Neurosynth meta-analysis (neurosynth.org; Yarkoni et al., 2011) using the term 'social' as the key word (association test; false discovery rate criterion of 0.01).

Statistic type for inference
(See Eklund et al. 2016)

All effects in whole-brain level were tested by one sample t-tests and cluster-wise FWE correction as implemented in SPM12. In Experiment 4, ROI-level classification accuracy against chance-level (50% for each classification) was tested by non-parametric bootstrapping test. In the other experiments, effects at ROI level were tested by null-hypothesis one sample t-tests (two-tailed) and Bayesian one sample t-tests, simultaneously.

Correction

For whole-brain analysis, multiple comparison corrections were conducted using cluster-level FWE correction ($p < .05$) as implemented in SPM12 (voxel-wise $p < .001$). For ROI analysis, multiple comparison corrections were conducted using Bonferroni correction.

Models & analysis

n/a	Involved in the study
☐	☒ Functional and/or effective connectivity
☒	☐ Graph analysis
☐	☒ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Resting-state functional connectivity (RSFC) using pearson correlation.
There were 15 seed ROIs in total, among which 2 key ROIs (i.e. the left vTPJ and lATL) and 9 sentence-processing ROIs were defined based on the meta-analysis results of Zaccarella et al. (2017) and 4 social-semantic-processing ROIs were defined based on the meta-analysis results of Zhang et al. (2021). For each

pair of seed ROIs, each participant's mean time series of each seed ROI was calculated and correlated with each other. The correlation coefficients were then Fisher-transformed to represent the RSFC. We conducted two analyses to examine whether the left vTPJ and lATL have stronger RSFC to the social-semantic or sentence-processing areas. In the first analysis, for each key ROI, we compared its mean RSFC to the social-semantic-processing ROIs with that to the sentence-processing ROIs across participants using both Bayesian and classical parametric t-test. In the second analysis, the mean RSFC matrix of these 15 seed ROIs was transformed back to correlation coefficients and then applied with k-means clustering to group them into 2 to 10 clusters. The ideal number of clusters was selected on the basis of the highest silhouette score.

Multivariate modeling and predictive analysis

Features in MVPA were voxel-based beta value of regressors. We conducted both whole-brain searchlight MVPA and ROI-based MVPA. Whole-brain searchlight analysis was conducted within a group-based gray mask. To obtain the mask, the normalized structural image was segmented into different tissues for each participant. The resulting gray matter probabilistic images were resliced to the same spatial resolution as that of the functional image, averaged across participants, and thresholded at 0.25 to generate a binary mask for searchlight mapping. For each voxel within the gray matter mask, support vector machine (SVM) decoding was conducted within a 5 x 5 x 5 voxels cube centered at that voxel using the leave-one-run-out cross-validation approach (Cortes & Vapnik, 1995). For the encoding stages of both sentences, we trained 4 classifiers to discriminate the poles of the 4 dimensions (HD or LD, HT or LT, BF or SF, and LE or SE) described in the current sentence. For the encoding stages of the second sentence, we additionally trained 4 classifiers to discriminate the poles of the 4 dimensions described in the context sentence (the first sentence). Before the SVM decoding was conducted, beta values within a cube were normalized to remove the common response pattern by subtracting the mean across the conditions to be discriminated. The resulting accuracy images were smoothed using a 6 mm FWHM Gaussian kernel for subsequent second-level statistical analyses.

ROI-based MVPA was conducted within the ROIs. After fitting the GLM, for each regressor of the encoding stage of the first and second sentences in each run, the estimated beta-values of all voxels within a given ROI mask were normalized and concatenated to form a fMRI pattern vector. Based on these fMRI pattern vectors, SVM decoding was conducted to discriminate the poles of the 4 dimensions described in the current or last sentences, just as in the whole-brain searchlight cubes.