

Experientially-grounded and distributional semantic vectors uncover dissociable representations of conceptual categories

Francesca Carota, Hamed Nili, Nikolaus Kriegeskorte & Friedemann Pulvermüller

To cite this article: Francesca Carota, Hamed Nili, Nikolaus Kriegeskorte & Friedemann Pulvermüller (2023): Experientially-grounded and distributional semantic vectors uncover dissociable representations of conceptual categories, *Language, Cognition and Neuroscience*, DOI: [10.1080/23273798.2023.2232481](https://doi.org/10.1080/23273798.2023.2232481)

To link to this article: <https://doi.org/10.1080/23273798.2023.2232481>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 12 Jul 2023.



Submit your article to this journal [↗](#)



Article views: 189



View related articles [↗](#)



View Crossmark data [↗](#)

Experientially-grounded and distributional semantic vectors uncover dissociable representations of conceptual categories

Francesca Carota ^{a,b,c}, Hamed Nili^{c,d}, Nikolaus Kriegeskorte^{c,e,f,g,h} and Friedemann Pulvermüller^{c,j,k,l,i}

^aMax Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; ^bDonders Institute for Cognitive Neuroscience, Radboud University, Nijmegen, The Netherlands; ^cCognition and Brain Sciences Unit, Medical Research Council, Cambridge, UK; ^dDepartment of Experimental Psychology, University of Oxford, Oxford, UK; ^eMRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK; ^fDepartment of Psychology, Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA; ^gDepartment of Neuroscience, Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA; ^hDepartment of Electrical Engineering, Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA; ⁱBrain Language Laboratory, Department of Philosophy and Humanities, WE4, Freie Universität Berlin, Berlin, Germany; ^jSchool of Mind and Brain, Humboldt Universität zu Berlin, Berlin, Germany; ^kCluster of Excellence “Matters of Activity”, Humboldt Universität zu Berlin, Berlin, Germany; ^lEinstein Center for Neuroscience Berlin, Berlin, Germany

ABSTRACT

Neuronal populations code similar concepts by similar activity patterns across the human brain’s semantic networks. However, it is unclear to what extent such meaning-to-symbol mapping reflects distributional statistics, or experiential information grounded in sensorimotor and emotional knowledge. We asked whether integrating distributional and experiential data better distinguished conceptual categories than each method taken separately. We examined the similarity structure of fMRI patterns elicited by visually presented action- and object-related words using representational similarity analysis (RSA). We found that the distributional and experiential/integrative models respectively mapped the high-dimensional semantic space in left inferior frontal, anterior temporal, and in left precentral, posterior inferior/middle temporal cortex. Furthermore, results from model comparisons uncovered category-specific similarity patterns, as both distributional and experiential models matched the similarity patterns for action concepts in left fronto-temporal cortex, whilst the experiential/integrative (but not distributional) models matched the similarity patterns for object concepts in left fusiform and angular gyrus.

ARTICLE HISTORY

Received 11 June 2022
Accepted 26 June 2023

KEYWORDS

Experiential semantics;
symbol grounding;
distributional statistics;
language network; RSA; fMRI

Introduction

Recent brain imaging research taking advantage of the information carried by population codes (e.g. neurones or voxels) in the brain has demonstrated that similar concepts and word meanings are represented by similar activity patterns (e.g. Carlson et al., 2014; Carota et al., 2017, 2021; Devereux et al., 2013; Fernandino et al., 2022; Kriegeskorte et al., 2008; Liuzzi et al., 2017) and that the statistical distributions of words, as captured by co-occurrence frequency vectors (Landauer & Dumais, 1997; Mikolov et al., 2013), relate to such representational patterns in the language networks (e.g. Carlson et al., 2014; Carota et al., 2017, 2021; Fu et al., 2022; Pereira et al., 2018; Wang et al., 2020; Xu et al., 2018). Distributional statistics, capturing the meanings of words and concepts by the company they take, have been proposed to provide a basis for language comprehension (e.g. Landauer, 1998; Landauer & Dumais, 1997).

However, distributional data in isolation fail to provide a cognitively realistic model of semantic knowledge (e.g. Harnad, 1990), because they do not clarify the links between word symbols and “the world” – the domain of referential semantics. Symbols become meaningful if anchored, or *grounded*, in the experiences of perceptions and actions that link them up with their extralinguistic referents (see Barsalou, 1999, 2008, 2017; Gallese & Lakoff, 2005; Glenberg & Robertson, 2000; Harnad, 2012; Marino et al., 2012; Pulvermüller, 1999, 2018) – for example, the object-related word *peach* with grounded semantic features such as [+ ROUND], [+ YELLOW] and the action word *hike* with [+ LEGS], [+ MOVEMENT]. Such experiential attributes can be used to efficiently code conceptual and semantic structure (Binder et al., 2016). They are reflected in reaction times during semantic judgments and lexical decision tasks (Andrews et al., 2014; Barsalou, 1999; Barsalou et al., 2008; Kousta et al., 2011; Paivio, 1971;

CONTACT Francesca Carota  francesca.carota@mpi.nl  Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands
This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Vigliocco et al., 2014; Zwaan, 2014), and likewise in the spatial distributions of brain activity patterns elicited by different semantic categories and indeed individual symbols (Fernandino et al., 2016, 2022; Hauk et al., 2004; Martin, 2007; Pulvermüller, 2018). Psycholinguistic theory holds that such experiential knowledge is linguistically encoded as a function of language use and task at hand (Symbol Interdependency hypothesis: Louwerse, 2008; 2018; 2021), so that an increase in perceptual simulation allows for richer and more detailed conceptual representations. Corroborating this view, recent behavioural results suggest that the sensorimotor properties of concepts enable a more precise specification of semantic categories, resulting into a best mapping of semantic similarities among concepts, as compared to distributional statistics *per se* (e.g.: Binder et al., 2016). As suggested by Carota et al. (2021), experiential properties are intercorrelated to a different degree with distributional data and with the hierarchical/taxonomic structure. The latter define categorical similarity between words, e.g. *peach* and *plum*, on their distance from their shared superordinate node, in this case “fruit”, and their shared overarching node, here “object”. Likewise, semantic dissimilarity, e.g. between *swan* and *hike*, depends on distance from the superordinate nodes “animal”/“object”, and “action”, respectively (see Carota et al., 2021 for discussion). In turn, distributional models coding word meaning in terms of linguistic contexts, based on word co-occurrences in texts, tend to produce semantic clusters that align with linguistic (or visual) context (Baroni et al., 2014; Bonner & Epstein, 2021; Crutch et al., 2013; Popham et al., 2021).

The goal of the present study is to investigate the commonalities and differences between grounded experiential and distributional semantic similarity mappings, and their integration.

It is well known that concepts and word meanings belonging to different categories are processed differently in the human brain; for example, action-related words and object words activate different sets of brain areas (Carota et al. 2012; Damasio et al., 1996; Hauk et al., 2004). The differences in activation patterns crucially depend on the different sensory and motor attributes that are pertinent for defining a given concept (Carota et al., 2017; Fernandino et al., 2016, 2022; Pulvermüller et al., 2009).

Previous work indicated that experiential semantic attributes are better reflected in the representational patterns of the brain’s “semantic network” than distributional information not only in these unimodal regions, but also in regions in multimodal association cortex (Fernandino et al., 2022), which integrate unimodal sensory and motor features. Different accounts

have ascribed an integration role to anterior temporal cortex (aTL: Lambon Ralph et al., 2017; Patterson et al., 2007), middle temporal gyrus (MTG: Hickok, 2014; Turken & Dronkers, 2011), inferior temporal gyrus (ITG, Price, 2000), angular gyrus (AG: Binder et al., 2009; Fernandino et al., 2016; Geschwind, 1965), or a range of fronto-parieto-temporal “connector hub” areas (Garagnani & Pulvermüller, 2016). Some of these multimodal integration regions are also good candidates for encoding distributional information, as they are part of the thematic network (Mirman et al., 2017; Wang et al., 2020; Xu et al., 2018), as well as intercorrelated experiential semantic properties flexibly mediated in multiple cortical regions (e.g. Reilly et al., 2016). In particular, recent evidence (Carota et al., 2021) suggested that the angular gyrus is important to interface distributional representations of concepts and their action-related properties, whereas anterior aspects of left inferior frontal gyrus (LIFG, BA 45-47), known to enable semantic combinatorial operations unifying simpler meaningful units (e.g. words) into larger semantic structures (e.g. sentences) (e.g. Hagoort, 2005, 2015, 2019), may primarily support representational similarity based on the distributional statistics (Carota et al., 2021). The relevance of this region for semantic combinatorics has been explained as rooted in its executive – rather than representational – functions (e.g. Hoffman et al., 2018). Intriguingly though, some studies indicate that also category-preferential visual and motor regions may take a role in representing distributional similarities for specific semantic word categories (Bonner & Epstein, 2021; Carlson et al., 2014; Carota et al., 2017; Hauk et al., 2004; Liuzzi et al., 2020; Mitchell et al., 2008; Popham et al., 2021; Pulvermüller, 1999). For instance, distributional semantic similarities among words were indexed in posterior inferior temporal cortices (e.g. Carlson et al., 2014), for which previous studies demonstrated category selectivity for different types of objects and nouns (e.g. places, foods, tools, and animals) (Mitchell et al., 2008; Mitchell and Cusack, 2016), and in prefrontal and posterior middle temporal regions, in which earlier work demonstrated the presence of representational similarities linked to rated experiential properties (particularly action-relatedness), and hierarchical taxonomies for action-related concepts (Carota et al., 2021 for detailed discussion). Reflections of the distributional links among action-related words has been also found in the dorsal region of the LIFG (BA44-45) (Carota et al., 2017), where the similarity in action-relatedness of the stimuli was also highly correlated with the similarity structure of the brain activity patterns (Carota et al., 2021).

These results lead to the hypothesis that qualitatively distinct sources of semantic information may be redundantly represented both in higher-order association cortex and in lower-level category-preferential regions depending on category information.

Here, we asked to what extent

- 1) Experientially-grounded and distributional semantic representational similarities are differently mapped in modality-preferential and multimodal areas and a multidimensional vector space combining grounded semantic information (based on human ratings of semantic properties) and distributional statistics (modelled using GloVe: Pennington et al., 2014) captured the semantic similarity structure of word representations better, or more completely, than each of these models on its own;
- 2) higher-order regions (e.g. LIFG, aTL, MTG/ITG, AG) and lower-level unimodal regions best indexed the neurometabolic correlates of the semantic similarities between concepts and whether such representational similarity mapping differed spatially between different word types (action- and object-related words).

To elucidate the relative contribution of experiential and distributional information, as well as their combination, to the mapping of the brain's multidimensional semantic space, we reanalysed an existing dataset of recorded fMRI responses (Carota et al., 2021), which had previously been scrutinised using Representational Similarity Analysis (RSA) (Kriegeskorte et al., 2008).

Materials and methods

Participants

Twenty-three healthy volunteers participated in the study. All participants were healthy right-handed, monolingual English native speakers, aged on average 29 years ($SE = 2.8$). All participants gave their informed consent to take part in the study and were remunerated for their time. Ethical approval was obtained from the Cambridge Psychology Research Ethics Committee.

Experimental procedure

Stimuli. Ninety-six words, sixteen from each individual category of leg-, arm-, face-related actions and tool-, animal-, food-related objects, were selected based on established semantic ratings (Carota et al., 2012, 2017, 2021; Hauk et al., 2008; Pulvermüller, 1999). Stimulus word groups were matched for a range of

psycholinguistic properties, including word length (counted in number of letters), letter bigram and trigram frequency, logarithmic word frequency, number of orthographic neighbours, and standardised lexical frequency, while differing in imageability, concreteness, and action-relatedness (see Table 1).

Relevant values were obtained from the CELEX database (Baayen et al., 1993) and the WordSmyth Web site (www.wordsmyth.net/). 21% of the action words were lexically unambiguous verbs and the lexically ambiguous ones which could be used as nouns and verbs were in the average 14 times more frequently used as verbs than as nouns (according to the Celex database; $SE = 4.2$). 58% of the object words were lexically unambiguous nouns and the lexically ambiguous ones which could be used as nouns and verbs were in the average 6 times more frequently used as nouns than as verbs ($SE = 2$). The object categories exhibited a gradient in action-relatedness, as animals were not associated with action (note that no pets were included), but to visual properties, food afford the actions of mouth and hand related to eating and preparing food, and tools afforded a greater variety of hand, foot, and mouth actions. The tool concepts were still highly associated with visual properties. We included object words with differing in their action-relatedness in order to test the power of the test models used here to capture the similarity and differences between the classes of object concepts and action concepts. This was indeed a part of the experimental design we were interested in, which also differs from most of earlier studies focusing on either action or object concepts.

Strings of meaningless hash marks matched in length to the stimulus words were used as low-level baseline stimuli during 120 trials. Null events consisting of a fixation cross displayed at the centre of the screen were presented during additional 60 trials. 60 trials consisting of misspelled words to be detected by the participants throughout the experimental task were presented. These "typo" trials did not include words from any of the semantic categories from which the 96 target words were taken – so as to avoid a bias towards one of these categories – and were discarded from the analysis. Statistical comparisons were carried out between brain activation patterns elicited by matched word categories.

After the fMRI experiment, participants completed an unannounced word recognition test containing both novel distractor and experimental words. They performed above chance (average hit rate: 85% [$STD = 8.3\%$]), indicating their attention to the words and compliance with the task. Results were used to confirm that subjects had been attentive continuously during the silent reading task.

Table 1. Psycholinguistic properties and semantic ratings are shown for each word sub-category, as well as for the categories of action- and object-related words. Means and standard errors (in brackets) are reported for each word category, along with results of an ANOVA comparing ratings between word groups.

	<i>Arm</i>	<i>Leg</i>	<i>Face</i>	<i>Animal nouns</i>	<i>Food nouns</i>	<i>Tool nouns</i>	Main effect of word-type (F)	<i>Action verbs</i>	<i>Object nouns</i>	Action x Object
Length	4.81 (.16)	4.63 (.16)	4.50 (.13)	4.69 (.15)	4.56 (.13)	4.63 (.16)	.533 $p=.751$	4.65 (.09)	4.63 (.082)	.030 $p=.862$
Bigram freq.	31248.64 (3138.66)	32472.14 (4035.53)	29029.40 (2887.87)	31538.82 (2985.33)	32699.44 (3939.17)	30550.20 (3422.15)	.155 $p=.978$	30916.73 (1926.37)	31596.15 (1964.98)	.061 $p=.806$
Trigram freq.	2475.48 (283.17)	2673.85 (367.88)	2535.13 (301.53)	2601.56 (399.53)	2825.27 (299.75)	2209.20 (300.13)	.401 $p<.847$	2561.49 (180.96)	2545.34 (193.65)	.004 $p=.952$
No. of neighbours	5.19 (0.91)	5.06 (0.60)	5.44 (0.84)	5.25 (0.76)	4.81 (0.79)	5.25(0.84)	.071 $p=.996$	5.23 (.45)	5.10 (.45)	.039 $p=.845$
No. of meanings	1.06 (0.06)	1.13 (0.09)	1.13 (0.09)	1.13 (0.09)	1.06 (0.06)	1.19 (0.10)	.333 $p=.892$	1.10 (.05)	1.13 (.05)	.101 $p=.752$
Log. word freq.	0.66 (0.11)	0.57 (0.13)	0.56 (0.10)	0.61 (0.11)	0.61 (0.14)	0.72 (0.12)	.250 $p=.939$.60 (.06)	.65 (.70)	.266 $p=.607$
Imageability	4.47 (0.17)	4.45 (0.27)	3.97 (0.26)	6.32 (0.09)	5.48 (0.27)	5.35 (0.35)	11.969 $p<.001$	4.30 (.14)	5.72 (.16)	44.722 $p<.0001$
Concreteness	4.14 (0.16)	3.59 (0.19)	3.62 (0.19)	6.60 (0.08)	6.21 (0.19)	5.73 (0.21)	62.306 $p<.001$	3.78 (.11)	6.18 (.11)	250.451 $p<.001$
Action-relatedness	4.83 (0.23)	4.91 (0.24)	5.31 (0.25)	1.60 (0.10)	2.02 (0.30)	3.22 (0.41)	34.967 $p<.001$	5.02 (.14)	2.28 (.19)	129.953 $p<.001$
Face-relatedness	1.56 (0.10)	1.40 (0.08)	5.75 (0.23)	1.20 (0.07)	2.06 (0.26)	1.27 (0.10)	123.410 $p<.001$	2.90 (.31)	1.51 (.11)	18.344 $p<.001$
Arm-relatedness	5.68 (0.13)	1.93 (0.17)	1.33 (0.09)	1.11 (0.06)	1.37 (0.12)	2.85 (0.36)	89.026 $p<.001$	2.98 (.29)	1.78 (.17)	12.816 $p=.001$
Body Sensation	3.74 (0.29)	3.55 (0.22)	3.92 (0.31)	1.16 (0.07)	1.35 (0.11)	1.40 (0.16)	40.353 $p<.001$	3.74 (.16)	1.30 (.07)	28.050 $p<.001$
Valence	3.45 (0.27)	4.05 (0.20)	3.66 (0.31)	3.52 (0.09)	4.08 (0.13)	3.85 (0.14)	1.626 $p=.161$	3.71 (.15)	3.81 (.08)	.377 $p=.541$
Arousal	3.25 (0.27)	2.98 (0.20)	2.60 (0.29)	1.30 (0.21)	1.44 (0.16)	1.65 (0.19)	14.464 $p<.001$	2.94 (.15)	1.46 (.11)	65.591 $p<.001$
Colour	2.10 (0.27)	1.56 (0.11)	1.43 (0.18)	2.55 (0.19)	3.39 (0.37)	1.41 (0.15)	11.707 $p<.001$	1.70 (.12)	2.45 (.19)	11.543 $p=.001$
Shape	2.96 (0.30)	2.46 (0.21)	1.73 (0.21)	3.67 (0.20)	2.90 (0.26)	3.21 (0.21)	8.140 $p<.001$	2.38 (.16)	3.26 (.13)	18.231 $p<.001$
Sound				1.39 (0.90)	1.03 (0.02)	1.85 (0.34)	57.105 $p<.001$		1.42 (.12)	
Taste/Smell				1.93 (0.24)	4.46 (0.30)	1.11 (0.08)	4.114 $p=.023$		2.49 (.25)	
Motion	3.11 (0.21)	4.37 (0.10)	2.63 (0.15)	1.63 (0.07)	1.41 (0.15)	2.23 (0.30)	4.716 $p=.014$		1.75 (.12)	

Experimental Design. We adopted a rapid, periodic single trial, event-related paradigm. Fixation cross was presented at the centre of the screen whenever no stimulus was shown. Word and hash string stimuli were presented tachistoscopically, to discourage subjects from eye movements, for 100 ms. The stimulus onset asynchrony (SOA) was randomly varied between 3.5 and 4s. This design yielded overlapping, yet detectable haemodynamic responses (Kriegeskorte et al., 2008; Nili et al., 2014). The 96 stimulus words were presented in a different pseudo-random order in each of the 6 runs. Each stimulus occurred once per run (6 repetitions of each word in total). Stimuli were visually presented by means of E-Prime software (Psychology software Tools, Inc., Sharpsburg, USA, 2001) through a back-projection screen positioned in front of the scanner and viewed on a mirror placed on the head coil.

Task. Participants were engaged in a typo-detection task. They were given the instruction to attend to all the stimuli, to silently read the words and to understand their meaning. In addition, they were instructed to press a button with their left index finger if a misspelled word appeared at the centre of the screen. To avoid confounding by body movements and responses, all catch trials with or without manual responses to typos were excluded from any analyses.

As the study made use of data from previously published work (Carota et al., 2017, 2021), we will summarise the principal method points, and refer the reader to the original article for the details (Carota et al., 2021).

Imaging methods

Subjects were scanned in a Siemens 3 T Tim Trio using a head coil. Echo-planar imaging (EPI) sequence parameters were TR = 2000 msec, TE = 30 msec, and flip angle = 78°. The functional images consisted of 32 slices covering the whole brain (slice thickness 3 mm, in-plane resolution 3 × 3 mm, inter-slice distance 0.75 mm).

Differences with earlier work

Previous work reporting on the present data set focused on related, yet different questions. A first study investigated the brain correlates of distributional semantics by evaluating brain responses averaged across semantic categories (Carota et al., 2017), whilst a second study focused on item-specific analyses of brain responses reflecting distributional semantics (using LSA methods) and hierarchical semantic structure (based on data from WordNet) (Carota et al., 2021). However, none of the previous studies explored the relative contribution

of different sensorimotor attributes and distributional semantics. Here, we now ask how distributional and grounded experiential semantics are reflected in the similarity structure of word-elicited brain activation and whether a combination of both semantic approaches leads to a closer relationship between semantic similarities and similarities of the brain responses of the words. Furthermore, in the present study (but not in our previous ones), the analyses were conducted for the entire group of words, irrespective of their category, as well as for specific categories of action and object words, in order to detect representational differences between category-general groups of concepts, and specific categories of concepts. Importantly, to do this, we directly compared our three target models in a whole set of regions of interest, as well as in each region of interest individually.

Data analysis: general linear model for representational similarity analysis (RSA)

For multivariate RSA (Kriegeskorte et al., 2008; Nili et al., 2014), the analysis was carried out in participant native space, using realigned, unsmoothed and non-normalised functional data, which were co-registered with MPRAGE of each subject. Data were analysed using the general linear model. Response-amplitude was estimated for each voxel and for each of the 96 stimuli by performing single univariate linear model fit. Runs were concatenated along the temporal dimension. A separate haemodynamic predictor was included for each of the 96 stimulus words. As each word stimulus occurred once in each run, each of the 96 stimulus words had one distinct haemodynamic response per run, extended across all 6 runs. The response-amplitude (beta) estimate map associated with each stimulus trial across all (6) runs was converted into a t map by contrasting them against the implicit baseline in order to compute the representational dissimilarity matrices for RSA (Carota et al., 2021; Nili et al., 2014).

Model specification

In order to test for the contribution of grounded experiential and distributional semantics on the representation of different word categories, we built up a semantic vector combining these qualitatively different properties (see Figure 1).

Experiential Model. The grounded experiential semantic model was constructed based on the following rated semantic properties relevant for the test words (for discussion, see Pulvermüller, 1999; and Binder et al., 2016): (1) action-relatedness (for discussion regarding

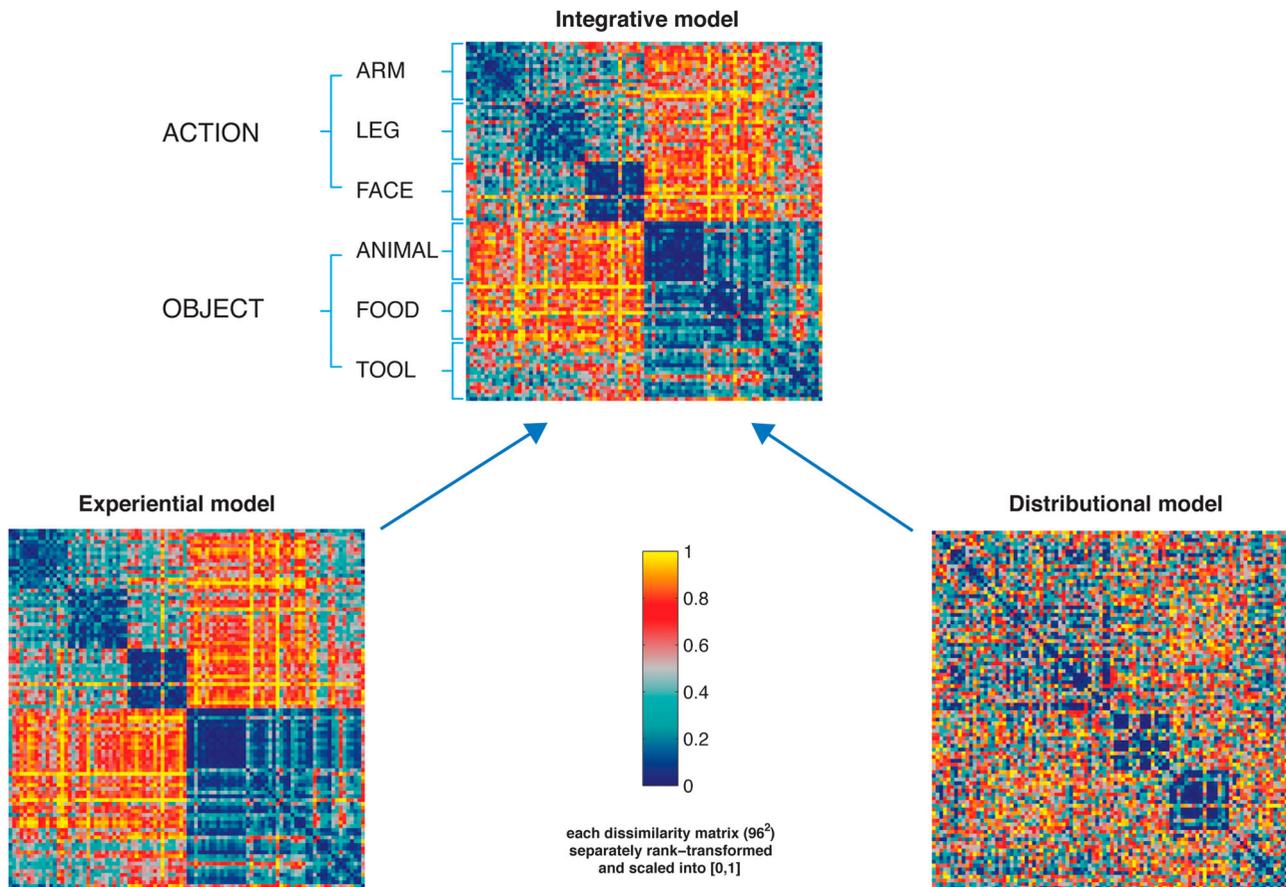


Figure 1. Representational dissimilarity matrices (RDM) displaying the semantic distances among the 96 stimulus words based on (a) the grounded “experiential” model reflecting human ratings (bottom left); (b) the distributional model based on GloVe (bottom right) and (c) the integrative model (top panel). The dissimilarity values from the two matrices were normalised before integration by averaging. Dark blue (= 0) indicates maximal similarity between identity pairs and yellow (= 1) indexes maximal dissimilarity. Each square in the model RDM shows (dis)similarities between the individual word items, which are plotted from top to bottom and from left to right – with dark blue indicating great semantic similarity or identity and red/yellow indexing dissimilarity. Therefore, the first diagonal in the matrix indicates the semantic identity of each single item with itself. Note that the integrated model at the top (see Table 2 for statistical tests) reflects a the categorial semantic structure of the stimulus vocabularies; please note the series of six blue squares from top left to bottom right representing semantic similarities within each of the six semantic subtypes. Also, the lexical categories of verbs and nouns are clearly revealed by this integrated model (big red/yellow “dissimilarity squares” at lower left and upper right). The only exception from the between-lexical category-dissimilarity is seen for tool words, which are indeed semantically associated with action verbs (for example, “knife” and “carve”/ “peel”). The sensorimotor model shows the similarities differences in lexicosemantic categories comparably. The distributional model also reflects the categorial semantic structure of the previous models to a lesser extent.

action-relatedness for object concepts, see Carota et al., 2012), (2) arm-relatedness, (3) leg-relatedness, (4) face-relatedness, (5) colour, (6) shape, (7) visual-relatedness, (8) imageability, (9) body-sensation, (10) arousal, and (11) valence. The ratings were acquired in a separate study using Lime Survey (<http://www.limesurvey.org> LimeSurvey Project Team/Carsten Schmitz, 2013). Fifteen native speakers of English were recruited to provide ratings for each word for a number of semantic variables, covering (1) sensorimotor meaning features – including imageability, concreteness, visual-relatedness (colour, shape), body-relatedness, and action-relatedness – and (2) affective-emotional features – including

arousal and valence (Bradley & Lang, 1994; Osgood et al., 1975). We administered explicit semantic ratings asking the following questions:

1. Is this word easy to imagine?
2. Is this word typically used to speak about something you can concretely experience (touch, see, hear)?
3. Is this word typically used to speak about visually perceived scenes or objects?
4. Is this word typically used to speak about colour?
5. Is this word typically used to speak about shape?
6. Is this word typically used to speak about activities you perform yourself (e.g. to move)?

7. Is this word typically used to speak about activities you perform with the face?
8. Is this word typically used to speak about activities you perform with the arm?
9. Is this word typically used to speak about activities you perform with the legs?
10. Is this word typically used to speak about something exciting (arousal)?
11. Is this word typically used to speak about negative or positive emotions (valence)?

Participants used a 1-7 Likert scale to provide their judgements. Please see Pulvermüller (1999) for further discussion.

Therefore, in the experiential model, each word was represented as $n(11)$ -dimensional vector, where dimensions correspond to the semantic properties (see list above) and the value for each dimension was the average rating of a given property. For each word pair, we calculated the correlation distance ($1-r$) between the values of each of these vectors, and entered the resulting values in each cell of a Representational Dissimilarity Matrix (RDM) (Figure 1, left bottom panel). We reasoned that 1-correlation (Pearson correlation across space) could be preferred for the experiential models and the neural RDMs (Kriegeskorte et al., 2008), because it is the most commonly used measure for RDMs. To construct the distributional vector model, we adopted cosine similarity to follow the classical metric for the calculation of semantic similarity based on co-occurrence frequencies in the tradition of Latent Semantic Analysis, as described in the next paragraph. Please note that Pearson correlation and cosine similarity are closely related, differing only in the fact that Pearson correlation subtractively normalises the mean to each pattern (to 0), while both methods divisively normalise the pattern variance (to 1).

Distributional Model. We applied a state-of-art extension of Latent Semantic Analysis (LSA, Landauer & Dumais, 1997), GloVe (Pennington et al., 2014). Similar to LSA, the model assumes that words have similar meaning if they tend to occur, beyond the *same* textual span, in *similar* contexts. Therefore, our distributional model indexes the semantic relationships between both words which co-occur in the same texts and paragraphs (first-order co-occurrence: for example, words linked to a common event or function, e.g. *cake* and *spoon*, *peach* and *to peel*, *harp* and *to play*), and, most importantly, words which do not appear in the same text, but can co-occur in similar contexts (second order co-occurrence). For instance, although *play* and *music* may not appear in the same text, they may separately co-occur with words like *symphony*, *orchestra* or

recording. Thus, the distributional model captures abstract, second-order semantic information about word meanings, reflecting statistical knowledge about their actual usage in language. The distributional model was generated from a lexical corpus, according to established methods (see Bruffaerts et al., 2013 for a discussion on the use of different types of semantic models in neuroimaging). We used the pretrained word vectors that the GloVe authors (Pennington et al., 2014) reported their highest model performance with on the word analogy and word similarity tasks. These vectors have a length of 300 dimensions and were obtained by training on 42 billion tokens of text. The GloVe method is a “global log-bilinear regression model that combines the advantages of the two major model families in the literature: global matrix factorization and local context window methods” (Pennington et al., 2014). In natural language processing, global matrix factorisation is based on matrix factorisation methods from linear algebra to reduce large term frequency matrices representing the occurrence or absence of words in a document. The model is based on weighted least squares, and trains on global word-word co-occurrence counts exploiting only the nonzero elements in a word-word co-occurrence matrix, rather than on the entire sparse matrix or on individual context windows in a large corpus. The embeddings are then optimised directly, so that the dot product of two-word vectors corresponds to the log of the number of times the two words will occur near each other. For example, if the two words “apple” and “pear” occur in the context of each other, say 20 times in a 10-word window in the document corpus, then: $\text{Vector}(\text{apple}) \cdot \text{Vector}(\text{pear}) = \log(10)$. This forces the model to encode the frequency distribution of words that occur near them in a more global context.

The model comprised 300 dimensions, the number yielding the highest model performance in the semantic task (word analogy task) evaluated by the authors (Pennington et al., 2014). Semantic similarity between words was then measured as the cosine between two-word vectors: the smaller the cosine, the greater the similarity between word stimulus pairs. These values were expressed as a dissimilarity matrix (Figure 1, right bottom panel).

Integrative (Experiential and Distributional) Model. The third semantic model integrated rated semantic properties and distributional properties of words, and was obtained by pointwise averaging the values contained in the two model RDMs. The similarity values from the two matrices were normalised before integration by averaging. The resulting RDM is displayed in the right panel of Figure 1 (top panel). Such integrative semantic

model expressed the contribution of the distributional and sensorimotor information about the test words in terms of pairwise similarity and dissimilarity relations, positing a mutual relationship between the experiential properties of two words and their co-occurrence likelihood. In particular, the smaller the dissimilarity between the set of experiential properties for two items (e.g. for the pair peach/plum: similarity in shape, “round”; action-related, “edible/manipulable”, taste, odour, etc.), the higher their likelihood to co-occur in a similar and coherent semantic context (e.g. groceries at the market, or cooking recipe). Reversely, the bigger the experiential property dissimilarity between two items (e.g. peach and jive), the smaller their likelihood to fit similar contexts.

Whole-brain RSA searchlights

Data were extracted for each participant individually using a “sphere of information” searchlight approach (roaming sphere with 10 mm radius: Kriegeskorte et al., 2008; Nili et al., 2014; this included about 121 voxels). See Carota et al. (2021) for method details. The correlation distances (1-correlation) between the response patterns for each word paired with every other word were expressed as representational dissimilarity matrices (RDMs), which are symmetric about a diagonal of zeros (Kriegeskorte et al., 2008). These brain data RDMs were then correlated with our model RDMs (using Spearman’s rank correlation) at each brain location. FDR correction at 0.05 for multiple comparisons across voxels and number of models was applied. 10,000 permutations were used in the analysis. In order to ensure that the searchlight maps we reported did not suffer from distortion due to either the searchlight size or the detection of fewer informative voxels, we performed additional analyses testing for the relatedness of the models and the brain activity patterns in selected regions of interest, as described below.

Analyses per regions of interest

Definition of ROIs. An exploratory analysis was run in hypothesis-based, pre-selected Regions of Interest (ROIs) in the left hemisphere with well-studied semantic roles in language comprehension (for reviews: Binder & Desai, 2011; Pulvermüller, 2013), as described in the introduction. These included the (1) LIFG (BA 44, BA 45 and BA47); (2) superior, middle, and inferior temporal gyrus (STG, MTG, ITG), further divided into anterior (aSTG, aMTG, aITG) and posterior (pSTG, pMTG, pITG) sections; (3) temporal pole, TP; (4) fusiform gyrus (FuG: Mion et al., 2010); (5) precentral gyrus (PCG); and (6)

inferior parietal cortex (both supramarginal gyrus: SMG, and angular gyrus: AG) (see Figure 3(A)). We automatically defined these ROIs using the standard Wake Forest University (WFU) Pickatlas toolbox, which generates ROI masks in standard MNI space based on the Automated Anatomical Labelling (AAL) parcellation. In order to carry out multivariate analysis within individual-subject native space, all ROI-masks were transformed to subject native space by inverting the spatial normalisation applied during GLM analysis. See Carota et al. (2021) for method details.

Relatedness between fMRI patterns and models as assessed by RSA

RDMs were computed for each participant for the abovementioned ROIs and related to the model RDMs. In order to test for the sensitivity of each of our three semantic models (i.e.: experiential, distributional, and the integrative model resulting from their combination) to semantic word categories, a first analysis in the whole brain (searchlights) and in selected ROIs was performed using the 96×96 matrix containing all words.

Results from statistical model comparisons

Analyses were run to statistically compare the model-fMRI pattern fit in all the abovementioned ROIs (results are reported only for the ROIs exhibiting significant semantic similarity effects). A first analysis directly compared the performance of the distributional, experientials and integrative models focusing on the 96×96 RDMs containing all words in order to assess the respective successes of the models in mapping of semantic similarity in different cortical areas. A second analysis compared the performance of the distributional, experientials and integrative models based on the 48×48 RDMs specific to action and object words in mapping category-specific semantic similarity in different cortical areas. A third analysis focused on the dominant attributes of action and object words, known to be reflected in the corresponding action and perception systems (e.g. Hauk et al., 2008), in order to test for the replicability of more specific category-specific semantic similarity effects in our data. To this aim, for the semantic sub-spaces of action and object words, two 48×48 model RDMs were constructed, a first one coding for the rated action-related properties (obtained from the ratings of action-relatedness, leg-, arm-, face-relatedness), and a second one coding for the rated visual properties (colour, shape, imageability). The model-fMRI pattern comparisons were performed by computing, for each subject, non-parametric Spearman’s rank correlations between model and brain activity RDMs. The two models were compared by subtracting the r-value of the

correlation between the second model and the brain activity patterns (fMRI RDM) from the r -value of the correlation between the first model and the brain activity patterns (fMRI RDM). The difference in r -value across all subjects was then tested against the null hypothesis of the value 0, to test for a difference in correlation, using a 1-sided Wilcoxon signed-rank test for all comparisons (both 96×96 RDM and 48×48 RDM). P -values surviving FDR correction for multiple comparisons are reported (Benjamini & Hochberg, 1995).

Discriminability of semantic categories. We further tested the hypothesis that action and object words are respectively encoded in fronto-central and temporal regions which are engaged by the processing of grounded action-related and visual properties of words. To this aim, an ANOVA contrasted the correlation values (as obtained from RSA) between the experiential model RDMs (48×48) specific to action and object words and the brain activity RDMs in a hypothesis-driven sub-set of frontal and temporal regions associated with the processing of action and object word (Hauk et al., 2004). These ROIs covered (1) the pars triangularis (BA 45) and (2) and opercularis (BA 44) of the LIFG, (3) a premotor region which was non adjacent to inferior premotor cortex and BA 44 (4) anterior temporal cortex (Patterson et al., 2007), (5) fusiform gyrus, and (6) angular gyrus (Binder & Desai, 2011). We additionally performed a 2×2 ANOVA (design: Region (6 levels, with 3 frontocentral and 3 temporo-parietal ROIs) \times Word Category (2 levels: action and object words)) to test for a double dissociation for the representations of action- and object-related words, contrasting r values between the corresponding experiential sub-models and the brain activity patterns in the abovementioned ROIs between word groups.

Results

Behavioural results

To evaluate the relationship between experiential semantic properties and the six semantic word categories, results of the behavioural semantic rating experiment were evaluated. Mean values and standard errors for both semantic properties, for which word groups differed, and psycholinguistic variables, for which they were matched, are summarised in Table 1.

Semantic similarities revealed by three linguistic models

The discriminability of semantic word categories in each model RDM was analysed by comparing (t-test) the

within-category dissimilarity values for each semantic type with the between-category similarity values. The resulting t -values were larger for the combination of experiential ratings and distributional semantics ($t = 6.30$; $p < 0.0001$) than for each individual model of experiential information ($t = 5.40$; $p < 0.0001$) and distributional semantics ($t = 4.50$; $p < 0.0001$). This suggests that the combined model of integrative experiential and distributional information produces better semantic category discrimination than each of the two measures alone on the level of purely cognitive-linguistic classification.

For both distributional and experiential models, the within-category dissimilarities were lower ($M = 0.75$ for Distributional; $M = 0.26$ for Experiential) than the between-category dissimilarities ($M = 0.80$ for Distributional; $M = 0.62$ for Experiential), as revealed by a Wilcoxon rank sum test ($Z = -3.08$, $p = .002$ for Distributional; $Z = -3.23$, $p = .001$ for Experiential). The integrative model combining distributional and sensorimotor dissimilarities showed a similar pattern ($M = 0.51$ within-category; $M = 0.66$ between-category; $Z = -6.15$, $p < .001$), i.e. the difference between the within/between category dissimilarities ($D = 0.16$) was larger than that for the distributional model ($D = 0.05$) but smaller than that of the experiential model ($D = 0.26$).

Results from whole-brain RSA searchlights

Whole-brain RSA searchlights revealed that the integrative model correlated with the similarity structure of the brain activity patterns in fronto-temporal and parieto-occipital cortex (see Table 2 and the top panel of Figure 2). A major cluster of activity was identified in left pITG and FuG, followed by a second one in left occipital lobe, and AG. A third cluster was seen in the pars orbitalis (BA 47), pars triangularis (BA 45) and pars opercularis of the LIFG (BA 44), extending to premotor cortex.

Additional significant correlations were found in the right MTG, and, to a lesser extent, in right parieto-occipital regions. The grounded experiential related to the brain activity patterns in the same regions as the integrative model, except the LIFG BA 47 (see Table 3 and Figure 2, middle panel).

Semantic similarity mapping of the distributional model led to pronounced correlations with the neural pattern in more anterior fronto-temporal regions, including bilateral IFG (BA 47), bilateral dorsal pre-supplementary motor area (pre-SMA), and AG, which showed pronounced correlations with the model in the right hemisphere (see Table 4 and bottom panel of Figure 2).

When the activity correlated with the experiential model was subtracted from the activity correlated with

Table 2. Results from searchlight RSA for the integrative model. Table of coordinates and significance voxel-level peak values (p) in each activation cluster that was correlated with the integrated semantic model integrating distributional and sensorimotor information.

Regions	Cluster Extent	Voxel-level P	Pseudo T	Coordinates		
				x	y	z
Left fusiform	2850	0.017	4.2	-39	-46	-24
Left Inferior Temporal		0.017	4.45	-51	-46	-12
Left Mid Occipital	1021	0.017	4.31	-24	-64	-36
Left Mid Frontal		0.017	4.68	-33	23	48
Left Insula		0.018	4.14	-30	23	-1
Left Inferior Frontal Opercularis		0.026	3.77	-51	11	29
Left Precentral Gyrus	120	0.030	3.60	-45	11	32
Right Middle Temporal	1514	0.017	4.52	51	-64	10
Right Middle Temporal		0.018	4.23	45	-73	21
Right Angular	1604	0.018	4.20	36	-70	44
Right Inferior Frontal Orbitalis		0.019	4.27	33	32	-12
Right Inferior Frontal Opercularis		0.019	4.02	51	11	25
Right Inferior Frontal Triangularis		0.019	3.74	48	26	2

the integrative model (Integrative minus Experiential: Table 5, Figure 2; right top panel), we found correlation differences in left posterior ITG and MTG.

In turn, when the activity correlated with the distributional model was subtracted from the activity correlated with the integrative model (Integrative minus distributional: Table 6; Figure 2, right bottom panel), we found correlation differences in dorsal pre-supplementary motor area, and right AG.

RSA results per ROI for all words

The three models produced significant correlations between semantic and brain activity differences in 3 (distributional), 7 (experiential) and 9 (integrative model) areas (see bar diagrams in Figure 3). Therefore, the integrative model returned significant correlations in a greater number of ROIs than the other two models, producing significant semantic similarity effects in pITG-MTG and LIFG (BA 44-45) with additional significant effects in left angular gyrus, precentral, anterior temporal cortex and fusiform gyrus (see Table 7, Figure 3, right top).

Likewise, the grounded experiential revealed significant effects across subjects in pITG-MTG, LIFG BA44-45, and angular gyrus (Table 7, Figure 3, right middle) and thus in more areas than the distributional model. The distributional model led to significant effects only in LIFG, and anterior ITG (Table 7, Figure 3, right bottom).

Results from the distributional statistical model comparisons respectively reveal distributional- and experiential-specific similarity in LIFG (BA 47) and pITG

As shown in Figure 4, the brain activation similarities correlated differently with the semantic similarities obtained from the 3 models. In left pITG, we found significantly higher correlations for integrative and experiential models, as compared with distributional ones, whereas the opposite result was achieved for LIFG. Only significant differences are reported.

Double dissociation of action and object words in motor and temporo-parietal regions

Contrasting the experiential and the distributional sub-models for action and object words separately revealed category-specific differences between the two semantic types. For action concepts (see Table 8), both experiential and the distributional models correlated with the patterns of activity in LIFG (BA 44-45-47), and left posterior middle temporal cortex.

The integrative model performed comparably to both models, and also revealed a significant correlation with the brain activity patterns in PCG (Figure 5(A)). In contrast, for object concepts (see Table 9), the experiential model performed better than the distributional model in a network of regions comprising the LIFG (BA 44-45), left pITG, left FuG, and left AG.

The integrative model performed similarly well as the experiential, and was significantly superior to the distributional model in the left pMTG, AG, and FuG (Figure 5 (B)). This indicates that mapping of semantic similarities of action-related words is relatively robust across distributional and experiential methods, whereas the meaning of object word is relatively better mapped by the experiential method.

Overall, then, this pattern of results confirmed category-specific semantic mapping differences across distributional, experiential, and integrative models in left frontotemporal and inferior parietal cortex. Semantic mappings were generally more robust for models including experiential semantic information about the items. Indeed, the results also suggest that the distributional model captured the similarity for action concepts in frontotemporal and inferior parietal cortex well, but less so those of object concepts. Semantic mappings for object words were present in left pMTG, FuG and AG reflected experiential (and integrative) information.

In turn, the direct statistical comparisons between the models of the experiential visual and action-related semantic properties for action and object words

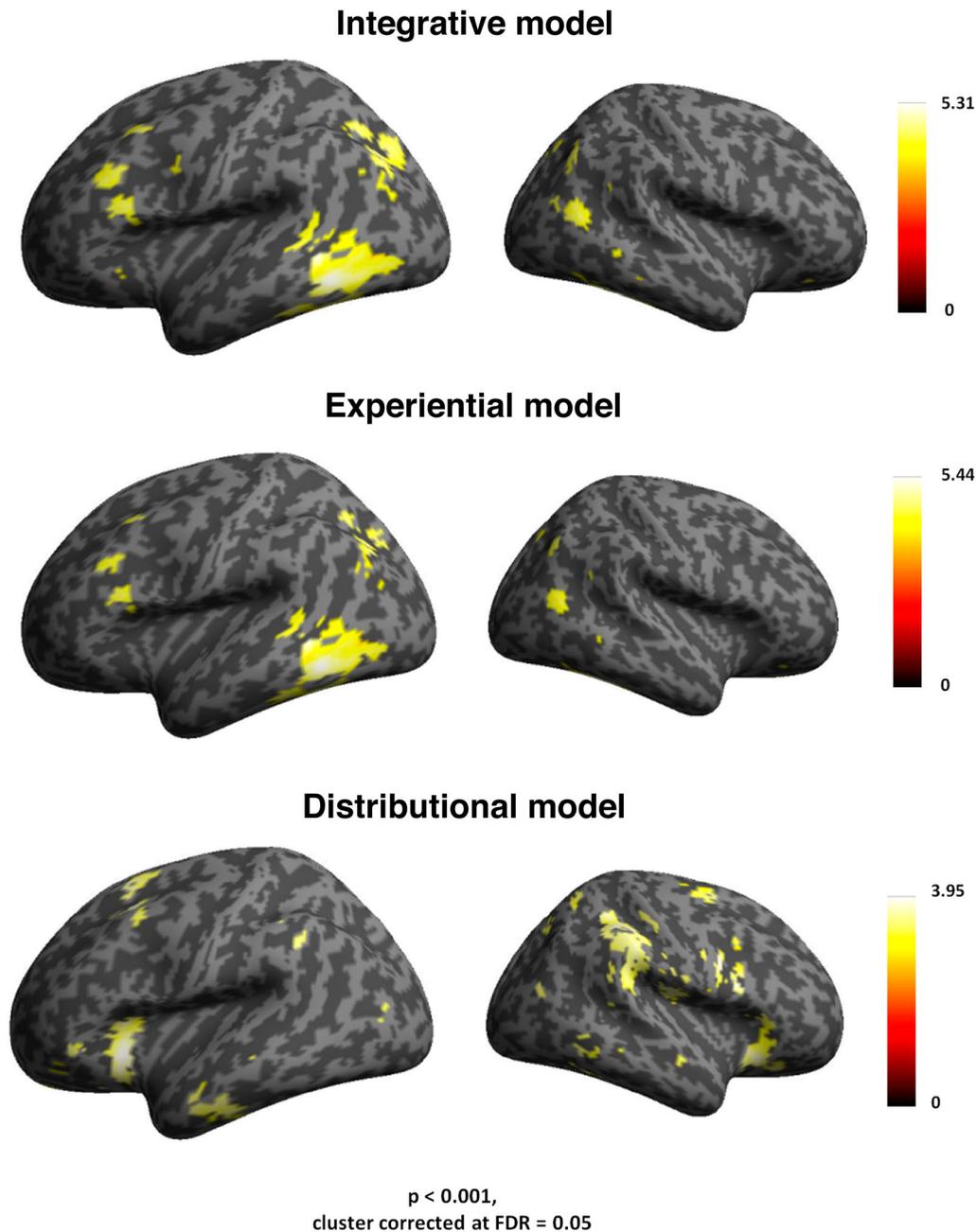


Figure 2. Results from whole-brain searchlight RSA (on the left). Left top panel: widespread activity triggered by the integrative model RDM in distributed prefrontal, premotor, inferior temporal and parieto-occipital cortex. Left middle panel: comparable correlational effects of the experiential model in the pars opercularis of the inferior frontal cortex (BA 44), premotor, inferior temporal and parietal cortex. Left bottom panel: activity triggered by the distributional model RDM in higher-order regions in left inferior frontal cortex (BA 47) and anterior temporal cortex. Results from model comparisons for the whole-brain searchlight RSA (on the right). Right top panel: correlation differences between the brain activity RDMs correlating with integrative model RDM minus the brain activity RDMs correlating with experiential model RDM in left posterior ITG and MTG. Right bottom panel: correlation differences between brain activity RDMs correlating with the integrative model RDM minus the brain activity RDMs correlating with distributional model RDM were seen in dorsal pre-supplementary motor area, and right AG.

revealed a clear differentiation between the underlying representations respectively in the LIFG (BA 44), PCG, and SMG, and in the FuG and AG (see Figure 6(A)). These significantly distinct patterns (FDR = 0.05) indicated a double dissociation between the representation of action and object words based on their respective

experiential attributes. These results confirm category-specificity in model-performance differences in motor and higher-order visual regions supporting the comprehension of action and object concepts.

Additional ROI analyses were performed to compare semantic similarities between the large lexicosemantic

Table 3. Results from searchlight RSA for the experiential. Table of coordinates and significance voxel-level peak values (p) in each activation cluster that was correlated with the sensorimotor semantic model.

Regions	Cluster Extent	Voxel-level P	Pseudo T	Coordinates		
				x	y	z
Left Inferior Temporal	694	0.016	5.44	-51	-46	-12
Left fusiform		0.016	4.45	-39	-46	-24
Left Mid Temporal		0.016	3.60	-48	-40	4
Left Occipital	312	0.017	4.27	-33	-70	40
Left Angular		0.026	2.97	-30	-55	-31
Left Inferior Frontal Triangularis	153	0.016	4.13	-48	26	21
Left Inferior Frontal Opercularis		0.017	4.03	-48	10	29
Left Precentral		0.017	4.52	-48	14	31
Left Mid Frontal	45	0.027	4.05	-33	23	48
Right Middle Temporal	196	0.018	4.23	51	-64	10
Right Middle Temporal		0.018	4.20	45	-73	21

categories of action-related verbs and object related nouns, as a further sanity check of the data. A statistical analysis of pooled values from category-preferential frontal and posterior areas (see Methods) indicated a differential mapping of lexicosemantic categories (ANOVA design: Word Category (action verbs vs. object nouns) \times Region (frontal vs. temporo-occipital)). Please refer to Figure 6(B).

Discussion

In the present study, we asked how experiential semantic information about the actions and perceptions in which word meaning is grounded and distributional data about word co-occurrences in texts are

Table 4. Results from searchlight RSA for the distributional model. Table of coordinates and significance voxel-level peak values (p) in each activation cluster that was correlated with the distributional semantic model.

Regions	Cluster Extent	Voxel-level P	Pseudo T	Coordinates		
				x	y	z
Left Inferior Frontal (BA 47)	581	0.0051	6.38	-21	23	-20
Right Middle Frontal Gyrus		0.0051	6.23	30	50	25
Left Inferior Parietal (BA 40)	629	0.0051	4.73	-48	-49	48
Left Angular (BA 39)	190	0.0051	4.44	-39	-70	40
Right Precuneus	90	0.0051	5.12	15	-73	44
Left Middle Temporal (BA 22)	518	0.0051	3.52	-57	-64	-24
Left Precentral		0.0051	3.32	-54	5	29
Right Inferior Frontal (BA 47)	581	0.0051	5.27	51	38	-5

Table 5. RSA results from model comparisons in the searchlight framework. Table of coordinates and significance voxel-level peak values (p) in each activation cluster that was correlated with the integrative model contrasted against the distributional semantic model.

Regions	Cluster Extent	Voxel-level P	Pseudo T	Coordinates		
				x	y	z
Left Inferior Temporal	320	0.019	2.98	-51	-46	-12
Left Mid Temporal		0.019	2.30	-39	-46	-24

mapped to distinct neurometabolic activity patterns of the human brain and whether integrating these approaches with each other may lead to improved semantic mappings, including any category discriminability of semantically related action and object concepts. We found that the experiential and integrative models produced semantic mappings in a range of areas, including lateral prefrontal, parietal and middle temporal association cortices. In contrast, the distributional model best correlated with a more focused region in left inferior frontal cortex (BA 47). The integrative model performed better than the distributional model alone in left posterior inferior and middle temporal cortex, and it performed better than the experiential in left supplementary motor area, and right AG.

A further major result concerns the differences in our models' performance when separately analysing their relatedness to brain activity patterns for the large categories of action and object related concepts and words. Both the experiential and the distributional models, as well as the integrative one, explained the neural similarity structure for action words/concepts (although to a different extent) in LIFG and pMTG, whilst only the experiential and integrative models produced significant correlations with semantic similarities of object words in the pITG, FuG, and AG. Both combined and grounding models showed a double dissociation with selected frontal lobe areas better mapping action semantic features and posterior areas

Table 6. RSA results from model comparisons in the searchlight framework. Table of coordinates and significance voxel-level peak values (p) in each activation cluster that was correlated with the integrative model contrasted against the experiential model.

Regions	Cluster Extent	Voxel-level P	Pseudo T	Coordinates		
				x	y	z
Left Mid Frontal	138	0.024	2.98	-33	23	48
Right Angular	200	0.024	2.60	36	-70	44

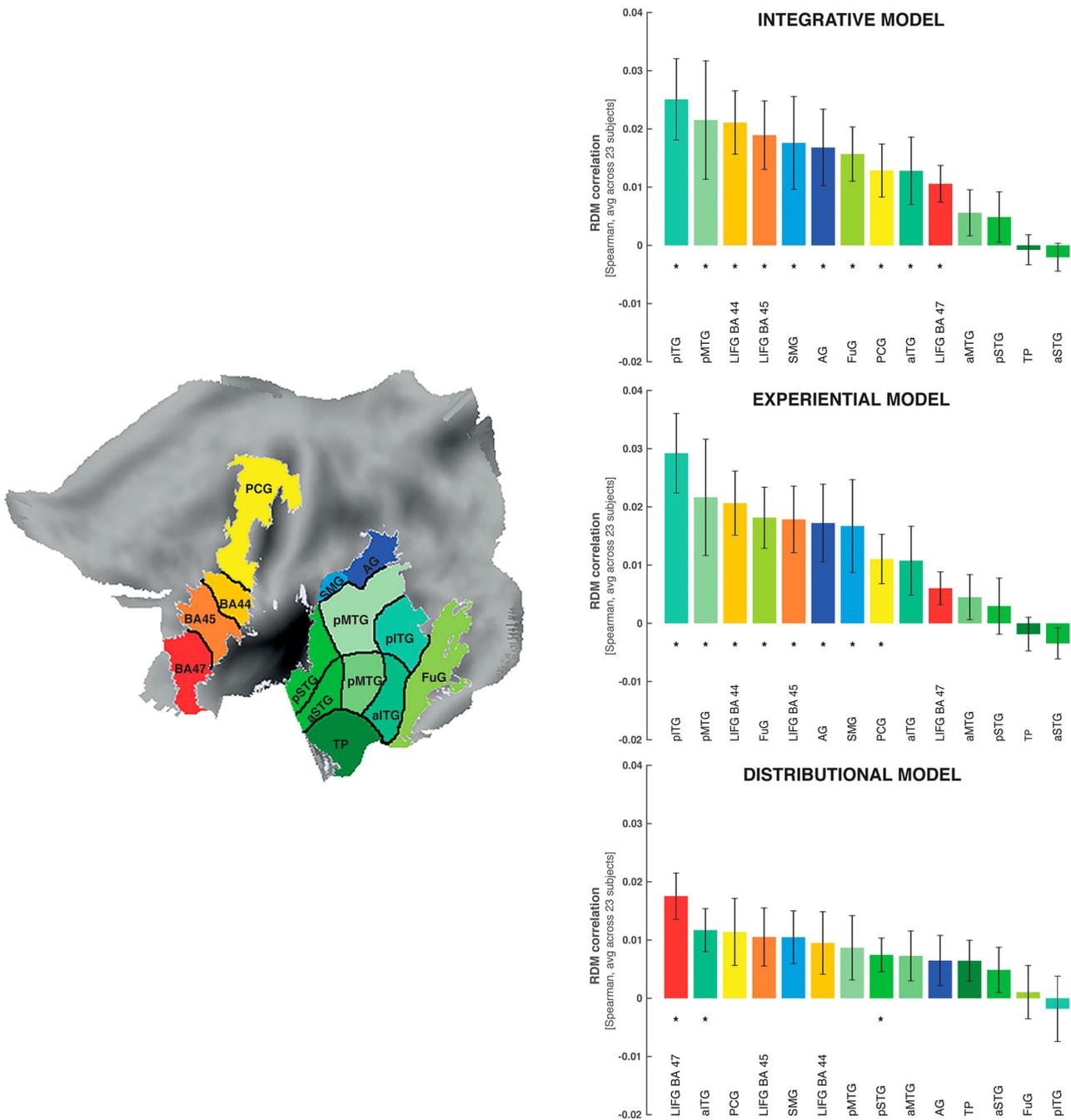


Figure 3. RSA results in the selected ROIs projected onto a flat map of the left hemisphere (the left panel). On the right: The bar graph depicts the averaged model-fMRI pattern correlations for each of the ROI under examination (the corresponding bar is indicated with the same colour as the ROI) across subjects (23). Spearman's rank correlations were calculated to assess the relatedness between brain activity and model RDMs and statistical inference was applied on the single subject correlations using a one-sided signed-rank test across subjects, testing whether the resulting correlation coefficients were significantly greater than zero. Below each bar, the significance value for the test is reported, corrected for multiple testing across brain regions by applying the FDR procedure (marked with an asterisk); the expected FDR was less than 5% (Benjamini & Hochberg, 1995). The horizontal bars in black indicate significant differences from model comparisons after FDR correction across ROIs (FDR = 0.05). Left panel: Effects specific to the LIFG (BA 44), where the distributional model differed significantly from both the sensorimotor and integrative models. Left panel: Effects specific to the left pITG. There was a significant difference (FDR = 0.05) between the effects of the sensorimotor and integrative models and the ones triggered by the distributional model.

indexing object related perceptual ones. These results suggest that both distributional statistics and grounded experiential data contribute to a reliable

decoding of symbol meaning of different semantic types, thus supporting the view that both grounding and distributional information are relevant for the

Table 7. RSA results from ROI analyses testing for relatedness between models and brain activity patterns for all words. Table of correlation values (r) and significance values (P) between the brain activity patterns in ROIs and the three models for all words per ROI. Correlations values which survive FDR correction for multiple comparisons and the model are indicated by asterisk (*).

ROI	Integrative model			Experiential			Distributional model		
	mean r-val	p-val	sig	mean r-val	p-value		mean r-val	p-value	sig
LIFG BA44	0.021	0.000319	**	0.021	0.000175	**	0.009	0.067067	
LIFG BA45	0.019	0.000127	**	0.018	0.001363	**	0.011	0.020763	
LIFG BA47	0.011	0.001363	**	0.006	0.037435		0.018	7.63E-05	**
PCG	0.013	0.003726	*	0.011	0.006764	*	0.011	0.052271	
TP	-0.001	0.600114		-0.0002	0.699477		0.006	0.032518	
aITG	0.013	0.032518	*	0.011	0.059323		0.012	0.003355	*
aMTG	0.006	0.157277		0.004	0.196536		0.007	0.084762	
aSTG	-0.002	0.7774		-0.003	0.811739		0.005	0.055713	
pITG	0.025	9.07E-05	**	0.029	5.33E-05	**	-0.002	0.117004	
pMTG	0.022	0.006764	*	0.022	0.001726	**	0.009	0.123375	
pSTG	0.005	0.172342		0.003	0.365694		0.007	0.006764	*
FuG	0.016	0.001208	**	0.018	0.001726	**	0.001	0.446606	
SMG	0.018	0.016335	*	0.017	0.020763	*	0.01	0.024221	
AG	0.017	0.012711	*	0.017	0.006147	*	0.006	0.061209	

discrimination and representation of concepts (see, for example, Barsalou et al., 2008; Louwse, 2008). Consistent with earlier work (Fernandino et al., 2022), our results also suggest that grounding information about the actions and perceptions words are used to speak about are generally better reflected in brain response similarities than distributional information about word co-occurrences.

The task applied in the present experiment involved attentive silent reading with occasional typo detection. This task was chosen to ascertain alertness and focused attention to the stimuli. As we were interested in semantic processes elicited by attended-to written stimuli, we avoided a task forcing subjects towards a specific type of semantic processing. Note that a semantic task, such as semantic judgements under a pre-defined aspect or matching to a target concept or category, would have elicited neurocognitive processes of comparison, which would have overlaid, modulated and thus potentially confounded the process of interest,

namely word meaning access. One may still argue that our methodological choice leaves the theoretical possibility that subjects attended to the words but did not understand or semantically process them. However, it is widely believed that skilled readers and users of a language cannot avoid understanding written words they see and attend to. This position is supported by a range of previous neurocognitive studies revealing consistent brain correlates of word and sentence meaning in passive and non-semantic tasks (see, for example, Carota et al., 2012, 2017; Hauk et al., 2004; Pulvermüller et al., 2005; for reviews see Binder & Desai, 2011; Pulvermüller, 2018). Consistent with this claim, the present dataset showed semantic signatures of the word materials that mapped semantic similarities obtained from different linguistic methods. Therefore, the correlation of brain activation similarities with semantic similarities provides clear evidence that the words presented were also understood and their meaning processed. Consistent with earlier work, these results thus show semantic

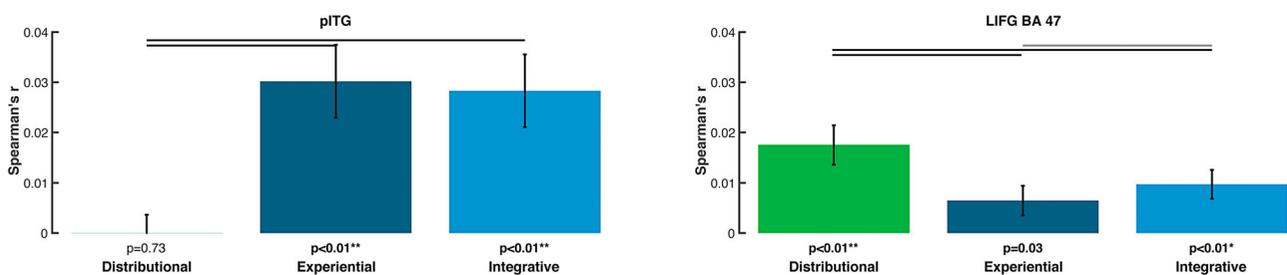


Figure 4. RSA results from model comparisons across all words for selected ROIs and for each of the three models: distributional (green), experiential (dark blue), and integrative model (light blue) (lines above bars indicating significance, $p < 0.05$ FDR corrected). Left panel: Semantic similarity effects specific to the left pITG, where the sensorimotor and integrative models matched the brain activity patterns significantly better than the distributional model. Right panel: Semantic similarity effects specific to the LIFG (BA 47), where the distributional model fits the brain activity patterns significantly better than the sensorimotor and the integrative models. The diagrams are based on the Spearman's rank correlations calculated to assess the relatedness between brain activity and model RDMs.

Table 8. RSA results from ROI analyses testing for relatedness between models and brain activity patterns for action words. Table of correlation values (r) and significance values (P) between the brain activity patterns in ROIs and the three models for object words per ROI. Correlations values which survive FDR correction for multiple comparisons and the model are indicated by asterisk (*).

ROI	Integrative model			sig	Experiential			sig	Distributional model		sig
	mn(r)	p -value			mn(r)	p -value			mn(r)	p -value	
LIFG BA 44	0.0297	0.0008	**	0.0242	0.0107	*	0.0185	0.0374			
LIFG BA 45	0.0255	0.0027	**	0.0212	0.0107	*	0.0168	0.0459			
LIFG BA 47	0.0205	0.0027	**	0.0154	0.0374		0.0140	0.0303			
PCG	0.0163	0.0281		0.0127	0.0755		0.0106	0.0557			
TP	0.0081	0.1234		0.0042	0.1113		0.0080	0.1501			
aITG	0.0099	0.1001		0.0076	0.3110		0.0017	0.3110			
pITG	0.0310	0.0019	**	0.0272	0.0037	*	0.0151	0.0401			
aMTG	0.0067	0.1363		0.0027	0.3434		0.0032	0.1723			
pMTG	0.0222	0.0034	**	0.0149	0.0242		0.0202	0.0208			
aSTG	0.0042	0.2410		-0.0003	0.4822		0.0099	0.1001			
pSTG	0.0046	0.1965		0.0022	0.2800		0.0042	0.6675			
TP	0.0081	0.1234		0.0042	0.1113		0.0080	0.1501			
FuG	0.0137	0.0401		0.0089	0.0801		0.0120	0.0490			
SMG	0.0194	0.0082	*	0.0163	0.0046	*	0.0148	0.1723			
AG	0.0164	0.0593		0.0149	0.8353		0.0036	0.3545			

processing occurring independently of the application of a semantic task.

The importance of experiential information for the representations of concepts

Our present results demonstrate that the grounded experiential and the distributional models predict the representational patterns across different regions considered in the present study. This confirms earlier evidence suggesting

that experiential information is representationally predominant over distributional data in multimodal semantic regions (Fernandino et al., 2022). However, the present data further reveal striking differences in model performance when the models are compared within separate regions, as the distributional and experiential models respectively mapped the high-dimensional semantic space in left inferior frontal gyrus (LIFG, BA 47) and in left inferior posterior temporal cortex, in which also the integrative model performed best.

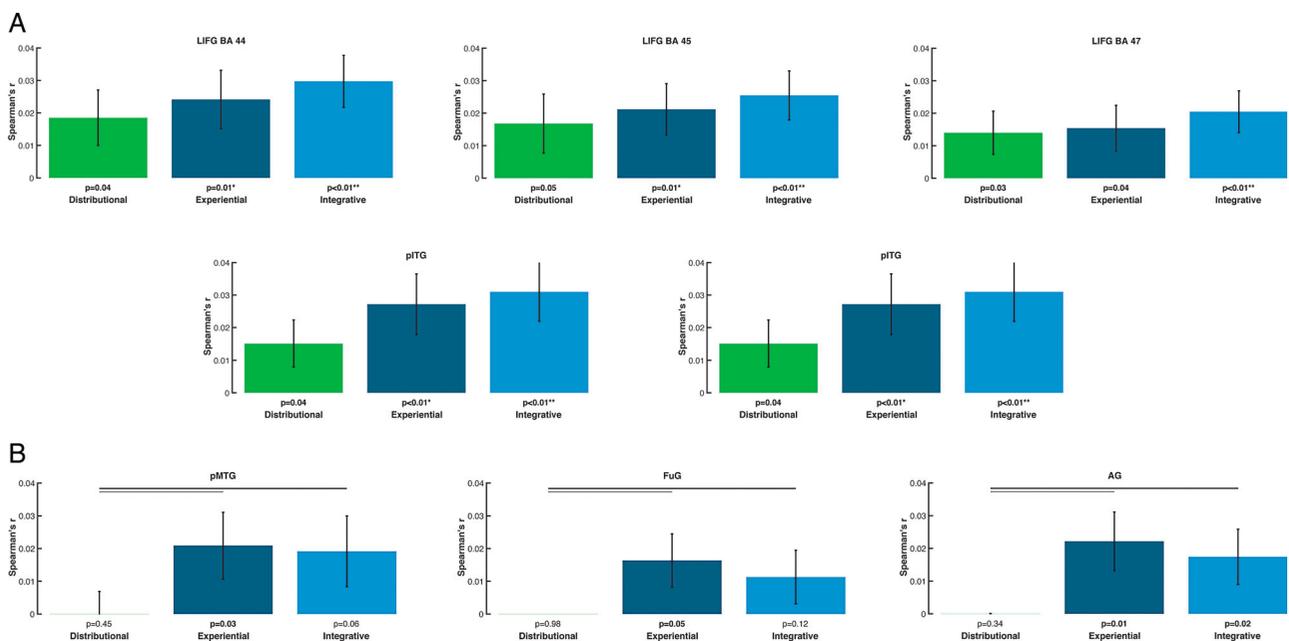


Figure 5. Category-specific results for object- and action-related concepts from model comparisons. The bar graphs depict the averaged correlations between the brain activity RDMs in selected ROIs and the (48×48) RDMs for the distributional (green), the experiential (dark blue), and the integrative (light blue) models for the sub-spaces of action and object words. Significant differences are indicated by the horizontal bars in black after FDR correction across models (FDR = 0.05). (A) For action words, there was no significant difference in model performance, as the three models performed equally for LIFG (BA 44-45-47); similar results were seen in pITG and pMTG. (B) For object words, the integrative and experiential models were significantly superior to the distributional model in the pMTG, FuG, and AG.

Table 9. RSA results from ROI analyses testing for relatedness between models and brain activity patterns for object words. Table of correlation values (r) and significance values (P) between the brain activity patterns in ROIs and the three models for object words per ROI. Correlations values which survive FDR correction for multiple comparisons and the model are indicated by asterisk (*).

ROI	Integrative model		sig	Experiential		sig	Distributional model		sig
	mn(r)	p -value		mn(r)	p -value		mn(r)	p -value	
LIFG BA 44	0.0117	0.1113		0.0162	0.0801		-0.0107	0.8828	
LIFG BA 45	0.0122	0.1234		0.0180	0.0848		-0.0163	0.9776	
LIFG BA 47	0.0114	0.1363		0.0091	0.2226		0.0052	0.1802	
PCG	0.0133	0.0261		0.0151	0.0303		-0.0030	0.6675	
TP	0.0007	0.7495		0.0003	0.6675		-0.0125	0.9926	
aITG	0.0060	0.2410		0.0051	0.3657		0.0019	0.4584	
pITG	0.0167	0.0557		0.0169	0.0948		-0.0004	0.2317	
aMTG	0.0114	0.0755		0.0099	0.1723		-0.0005	0.5297	
pMTG	0.0192	0.0631		0.0209	0.0303		-0.0031	0.4466	
aSTG	0.0015	0.6890		0.0010	0.5885		-0.0133	0.9949	
pSTG	0.0029	0.3545		0.0005	0.5769		-0.0014	0.3884	
TP	0.0007	0.7495		0.0003	0.6675		-0.0125	0.9926	
FuG	0.0113	0.1234		0.0164	0.0459		-0.0150	0.9776	
SMG	0.0065	0.2050		0.0070	0.1363		-0.0031	0.6890	
AG	0.0174	0.0163		0.0222	0.0061		-0.0078	0.3434	

Furthermore, based on both experiential and distributional data, the semantic spaces of action-related verbs and object related nouns could be respectively mapped in specific category-preferential and category-general frontal (motor) and temporo-parietal (visual) areas. In turn, the distributional model complemented these effects, by capturing the brain activity patterns in left frontotemporal regions for action words, but did not match the brain activity patterns specific to object words. Furthermore, these category-specific representational profiles corresponded to the effects seen if the pertinent set of action-related and visually-related properties was selected for model comparison, when a clear double dissociation became manifest in the spatial distribution of their respective representations in category-preferential premotor, and higher-level regions in the ventral and dorsal visual streams in posterior temporal and parietal cortex.

Overall, then, our results from direct measures of metabolic signals during word comprehension support the assumption that the combinatorial knowledge of statistical regularities with which word symbols co-occur in linguistic contexts (e.g. texts or speeches), are not sufficient to form semantic representations of symbol meanings (Andrews et al., 2014; Cangelosi & Harnad, 2001; Harnad, 1990; Searle, 1980; Vigliocco et al., 2004). Rather, a condition sine qua non for a complete picture of the semantic similarity space is the grounding of these symbols in the motor, perceptual, and affective properties of their referents in the “world”, as mediated by dedicated action and perception systems (Barsalou, 2008, 2017; Harnad, 1990; Pulvermüller, 2013, 2018; Searle, 1980). On the basis of earlier work (Cangelosi & Harnad, 2001; Harnad, 1990; Vincent-Lamarre et al., 2016), an integrative model was expected to lead to a better approximation of human

semantic similarity judgments as compared to the use of each estimate taken separately. Here though, we found no significant difference in performance on the entire word set as compared to the experiential model, although the numerically larger numbers of areas with significant semantic similarity mappings were greater for the integrative model compared with the other two. The extent to which the explanatory power of integrative models combining distributional and experiential semantics has proven successful for describing the multidimensional semantic space is an empirical issue, and may depend on the linguistic meaning and context under examination. For example, distributional models alone are performing well in decoding sentential meaning, as well as in computing the similarity structure between the semantic content of sentences (e.g. Pereira et al., 2018). Furthermore, successful decoding in sentential contexts could not be achieved based on the experiential semantics of single words alone but requires additional information not explained by compositional semantics (e.g. about the meaning of idioms) for which combinatorial distributional information is essential. A recent study showed that the distributional GloVe model could predict fMRI representational similarities well when applied in isolation, model integration with a model including semantic attributes significantly improved fMRI prediction of sentence meaning (over GloVe alone and attributes alone) in left lateral temporal lobe and left inferior frontal gyrus (also see Anderson et al., 2019, 2020, 2021). Therefore, language-based computational models may come to the fore for predicting multi-faceted concepts/memories (as described by sentences), whereas they may be less valuable for isolated word-related concepts (current study and especially Ferdinandino et al., 2022). Although the extent to which distributional and experiential models are successful for

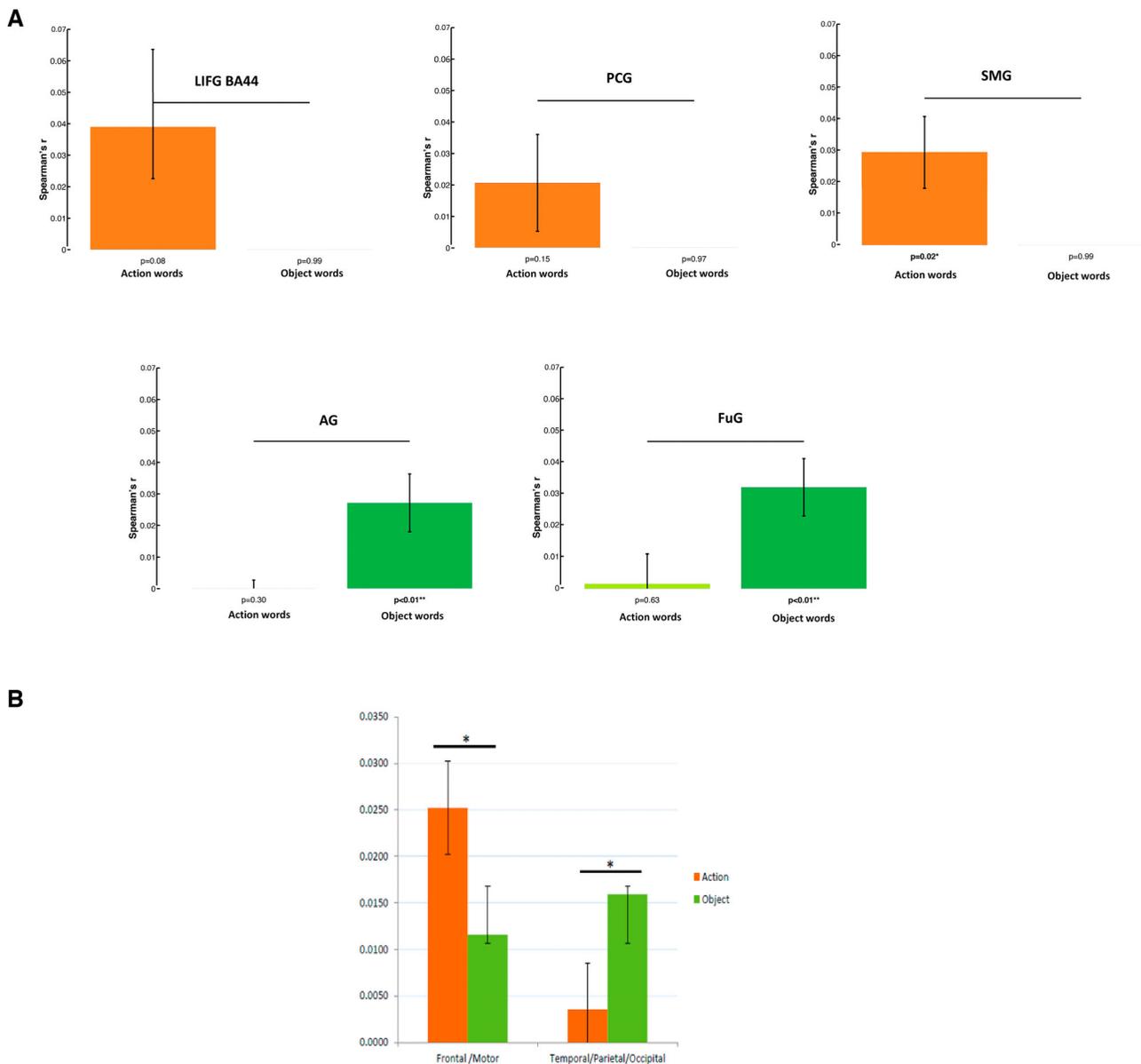


Figure 6. (A) Category-specific results for object- and action-related concepts from model comparisons. The bar graphs depict the averaged correlations between the 48×48 brain activity RDMs in specific cortical areas and the model RDMs based on action-related semantic properties (top panels) and visual properties (bottom panel) for the sub-spaces of action (orange) and object (green) words. Significant differences are indicated by the horizontal bars in black after FDR correction across models (FDR = 0.05). Top panel: in LIFG (BA 44), PCG and SMG, the model based on action-related semantic properties was correlated more strongly with the similarity patterns specific to action words as compared to the ones specific to object words. Bottom panel: in the FuG and AG, the opposite was found, i.e. greater correlations between the similarity structure based on the visual properties and the similarity structure of the brain activity patterns for object words than for action words. (B) Results from statistical analyses showing differential mapping of lexicosemantic categories obtained (ANOVA design: Word Category (action verbs vs. object nouns) \times Region (frontal vs. temporo-occipital regions of interest)), with a significant cross-over interaction between the factors Region and Word Category ($F[1, 22] = 7.45, p = 0.01$). This effect obtained by employing the integrated semantic vector model revealed a double dissociation in the semantic similarity mappings of action and object words onto focal category-preferential (BA 44, PCG) and category general (SMA, AG, FUG) regions. A comparable interaction effect was also found when the experiential model alone was applied ($F[1, 22] = 7.85, p = 0.012$). The distributional model did not reveal any semantic mapping differences across word categories.

describing the multidimensional semantic space is an inherently empirical issue, their explanatory power may depend on the type of linguistic meaning and the semantic context under examination, and may increase

for semantic contexts beyond single words. Integrating experiential attributes with distributional data is also beneficial for predicting autobiographical event imagination fMRI data (using GloVe models constructed from

participants offline descriptions of what they imagined) (Anderson et al., 2020). One reason is that sentential contexts may constrain distributional statistics to the computation of the words that they actually contain (first-order co-occurrence) for estimating semantic similarity, thus reducing the need to exploit second-order co-occurrence in *similar* contexts, as is the case for single words (as performed here exploiting a large corpus of sentences, paragraphs, etc.).

As mentioned, the present study is not the first to apply the distributional semantic model in RSA of brain signals to words (e.g. Anderson et al., 2019; Carota et al., 2017; Fernandino et al., 2022; Liuzzi et al., 2020; Mirman et al., 2017; Pereira et al., 2018; Wang et al., 2020; Xu et al., 2018). In our earlier studies, we used the classical approach of latent semantic analysis (Landauer & Dumais, 1997; see Carota et al., 2017), and a more state-of-art word embedding model (Word2Vec: Mikolov et al., 2013; see Carota et al., 2021), whereas the more recent GloVe method (Pennington et al., 2014) was applied here. When comparing the results, there is surprising consistency regarding anterior inferior frontal cortex, where the highest correlations between semantic and activation similarities were obtained. In addition, consistent similarity mappings were found in parietal cortex (AG). The GloVe method also led to additional significant similarity mapping in dorsal premotor and anterior temporal areas (see Figure 2). In contrast to our first study (Carota et al., 2017), where a range of inferior frontal areas (BA 44, 45, 47) revealed the clearest distributional semantic mapping correlates, the more posterior areas (44, 45) just missed the significance threshold in the present analysis.

A clear performance difference between distributional and grounded models was seen in the analyses of the action vs. object word analyses, where both models performed equally well in frontal areas, but only the grounded model succeeded in similarity mapping in posterior temporal cortex. The reason why, for our data set, the GloVe distributional model explained the similarity structure of the brain activity patterns for action words better than the ones for object words (where sometimes negative RDM correlation values were observed) might be found in the richer argument structure of action words (verbs) as compared to object nouns, that is, the lexical representation of the lexical arguments they take, such as other verbs, nouns and nominalizations, adjectives, and even prepositions, which provides specific information not only about the syntactic expression of the items that can be realised but also about their semantics (e.g. the verb “drink” requiring two noun phrases, one

related to a living being and the other to a liquid). However, this suggestion should be addressed in future experiments. Still, these considerations lead to an important caveat concerning our results on category-specificity from the comparison of action and object concepts based on GloVe. Because the performance of different types of distributional language models varies due to the particular type of techniques their metrics rely on (Anderson et al., 2021; Pereira et al., 2018), it is possible that even if different distributional models still differentiate between semantic categories, they may capture category-specific patterns to different extent depending on how much contextual information they take into account, and how. It is also possible that other models are better optimised for explaining object concept similarity. For instance, recent neuroimaging evidence suggested that topological network models may outperform both co-occurrence and experientials in language-specific regions when object-related words are considered (Fu et al., 2022). In particular, Fu et al. (2022) evaluated the performance of different models to assess task-invariant neural representations. The authors employed brain imaging fMRI data from two different tasks, involving picture naming and word familiarity judgements of written words. They constructed one model based on complex network (graph) path relations, with words as nodes and their simple co-occurrences as edges, and proposed that when language is represented as a graph, rich topological properties (node and edge layout patterns) can be computed, capturing information that affects semantic knowledge and language learning. They assessed the performance of the network graph topological models (graph-common-neighbours and graph-shortest-path), of a standard word embedding model such as Word2Vec (Mikolov et al., 2013), and of a model based on simple (edge) co-occurrence. They found that the distributional (Word2Vec) model correlated with activity in left occipital and posterior MTG when the brain imaging data from the picture naming task were examined. The same model correlated with activity in a greater number of cortical regions, also including bilateral anterior temporal lobe when the data from the familiarity judgement were included. However, a conjunction analysis testing for task-invariant effects by assessing the overlapping regions that showed significant positive correlations between neural activity and models showed effects (no surviving voxels). The study of Fu et al. (2022) employed a partial correlation approach regressing out of each language model potential confounds such as visual co-occurrence features, and sensorimotor information. The results bring novel insights

on the potential relevance of the (so far little tested) network graph topological models (describing edges and nodes arrangements in a network) to the semantic similarity structure of object-related words. However, important methodological differences regarding the type of tasks, the stimulus set (object nouns), the language (Chinese), the models themselves, none of which was based on the GloVe model employed here, make it difficult to establish a comparison with our present data. Also, these interesting results require further validation in the context of other word semantic classes, such as action verbs. Concerning the results from the RSA analyses, in our present study, our main semantic similarity effect in posterior middle/inferior temporal cortex was linked to the experiential model. However, the corresponding sensorimotor model was only used as a control model for partial correlation analyses, and not compared with the brain activity patterns on its own. Furthermore, no direct comparison of the models was performed, as we did in our present study. However, it is interesting to note that Fu et al. (2022) report task-invariant effects of the network graph topological model just in the two regions highlighted in our present results: the LIFG and the left posterior ITG/MTG. Therefore, despite the abovementioned methodological differences, our present results match the ones by Fu and colleagues in the loci of semantic similarity effects, but also diverge from those results because the semantic similarity mapping in the LIFG and in posterior inferior/middle temporal regions emphasises distributional and experiential information, respectively.

Hierarchical architecture of the representational semantic systems: category-general semantic similarity mapping in higher-level association cortex

The present results highlight a hierarchical architecture of the representational systems encoding lexical meaning, according to which distributional and grounded experiential semantic properties of word symbols are represented in partly dissociable sets of category-general and category-specific regions in motor and sensory cortex. Our data are consistent with models positing the co-presence of multimodal semantic properties in several higher-order convergence zones holding conjunctive representations derived from multiple low-level sensory and motor representations (Binder & Desai, 2011; Damasio, 1989a, 1989b; Meyer & Damasio, 2009; Pulvermüller 2013, 2021; Pulvermüller et al., 2009; Reilly et al., 2016).

Indeed, we found that the LIFG and the pITG respectively best reflected similarity in distributional and experiential properties of the stimulus words. These areas had been found active to written words from a range of different semantic categories, as shown by earlier univariate fMRI studies and likewise by applying k-means cluster analysis (Pulvermüller et al., 2009). Interestingly, these observations are also in line with results from connectivity analyses showing that a cortical network comprising left inferior frontal, and posterior inferior temporal, and inferior parietal cortex supports semantic processing functions specifically (e.g. Xiang et al., 2010). In particular, posterior inferior temporal cortex is known to store lexical information about conceptual features in memory (Carlson et al., 2014; Coutanche et al., 2016; Devereux et al., 2013; Ghio et al., 2016; Hagoort, 2019; Mitchell and Cusack, 2016; Mitchell et al., 2008; Tyler et al., 2013; Clarke and Tyler, 2014), thus supporting a semantic route to reading (e.g. Price, 2000).

Neuronal populations coding for representations of lexical-semantic similarities in the posterior temporal memory circuit may communicate via long-distance connections with population codes in inferior frontal regions, especially in the LIFG (see Fuster, 1997; Pulvermüller, 1999, 2018), as recently confirmed across imaging modalities (e.g. Lau et al., 2008; Pulvermüller, 2013; Xiang et al., 2010). Strengthened activity in this network may arise as a result of the spread of activation from the LIFG.

In turn, the more anterior/ventral aspect of the left inferior frontal cortex is critical for unification of semantic information into linguistic structures (e.g. Hagoort, 2015), a process which operates on combinatorial semantic information and which may require access to the representation of statistical knowledge, such as the one captured by our distributional model. The present results fit this picture well, stressing the importance of the LIFG (BA 47) in semantic processing highlighted in previous literature (e.g. Poldrack et al., 1999; Thompson-Schill et al., 1997), and particularly in the encoding of distributional representations, a finding confirming earlier results (Carota et al., 2017, 2021).

Also, anterior temporal regions showed a relatively stronger correlation of activity patterns with the distributional semantic model, which however did not reach a significant difference compared with the other two models, and did not stand out from the other language regions. In conclusion, the present study confirms a role of the anterior temporal regions in the encoding of semantic similarity, but failed to confirm this region as a possible apex for all forms of semantic knowledge to be brought together, as expected on the basis of the

“Hub-and-spokes” model (Hoffman et al., 2018; Lambon Ralph et al., 2017; Patterson et al., 2007), highlighting its contribution within a widespread network of distributed language representations.

Category-specific representations of action-related and visually-related words

A major finding of the present study was that a sub-network of inferior parietal (AG) and fusiform (FuG) regions reflected object-specific representations of visually-related semantic similarity. Previous fMRI studies have shown that the preferential response of different portions of the FuG to different object categories, for example animals and tools as employed here (Chao et al., 1999; Forseth et al., 2018; Martin, 2007; Tyler et al., 2013; Warrington & Shallice, 1984). Such categorical differentiation has been proposed to arise from higher and lower levels of shared features across the members of different categories of objects, giving origin to category structure (Tyler et al., 2013). It is interesting that, in the present data, the AG may become co-activated with an object-specific FuG region in the ventral visual pathway to code for an intermediate and more specialised level of object-specific representation

The AG encodes cross-modally integrated representations of semantic knowledge (e.g. Fernandino et al., 2015, 2016), as well as category-general representations of statistical knowledge about object words, as our present results show. This area may thus be a multimodal association region binding together different formats and modalities of semantic representations to enable the comprehension of complex conceptual information about object concepts. For instance, earlier evidence suggests that the AG supports the representation of complex object concepts (e.g. plaid jacket), based on simple conceptual constituents (e.g. jacket and plaid) (Price et al., 2015). The semantically coherent representation of these conceptual combinations presupposes the multimodal integration of cross-modal attributes in this well-established semantic “hub” area (Binder et al., 2009; Binder & Desai, 2011; Fernandino et al., 2015, 2022).

In turn, we here found that the posterior ITG/MTG reflected experiential and integrative similarities, but not distributional codes for the set of object words, whilst all models predicted the brain activity patterns for action words in posterior MTG. These differences are consistent with a role taken by this region in storing experiential information, and particularly action-relatedness (Carota et al., 2017, 2021), and action semantics (Hauk et al., 2008). The greater

importance of the pMTG in coding for distributional similarities among action verbs as compared to object concepts is also consistent with the it responds specifically to verbs (Elli et al., 2019).

Other studies also found sensitivity of the posterior ITG/MTG to the semantic similarity for a variety of object categories, including tools, when using distributional metrics (Word2Vec) and visual features, including shape (Liuzzi et al., 2020). Discrepancies between these results and our present research outcomes can depend not only on the distributional model employed (see discussion above), but also, more obviously, on the definition of the ROIs in terms of size and locus. For instance, the MTG is more sensitive to action-related tool nouns than nouns referring to animals (Carota et al., 2012, 2017; Desai et al., 2009; Fernandino et al., 2015; Hauk et al., 2008; Martin et al., 1996; Pillon & d’Honinchtun, 2011; Vannuscorps & Pillon, 2011). Furthermore, responses to object nouns are more posterior and inferior relative to the ones specific to verbs (Bedny & Caramazza, 2011; Perini et al., 2014), and this can explain differences in the semantic similarity effects reported for object words across recent studies in these regions. Finally, also the use of different distributional methods for generating semantic vectors (e.g.: Word2Vec vs. GloVe) could underlie these differences.

In the current data, it was noteworthy that specialised sub-networks encode distinctive information supporting the disambiguation and identification of specific categories of words and concepts. The additional results from the comparisons of the models coding for specific visual and motor properties clearly further demonstrate the neurocognitive distinctiveness, and distinction, of action and object words. Our present findings thus point to a genuinely experiential semantic origin of category-specific representations: classifying an object and the word used to name it as a member of a category may require simple conjunction of experiential attributes, which confines activity to category-specific temporo-parietal regions.

The semantic sub-space in left motor, inferior frontal (LIFG BA 44), motor, and the SMG was, in turn, well mapped by the model coding for rated action-related properties of the words, consistent with earlier findings on the role of these regions in action semantics (e.g. Damasio & Tranel, 1993; Davey et al., 2015; Dreyer et al., 2020; Grossman et al., 2008; Hauk et al., 2004; Hillis et al., 2004, 2006; Kemmerer et al., 2012; Kemmerer & Tranel, 2003; Pulvermüller et al., 2005; Rueschmeyer et al., 2010; Vukovic & Shtyrov, 2019). Interestingly, these same regions reflected both distributional and experiential data as well.

The need for integration of distributional data with grounded experiential semantics

On the basis of earlier work (Cangelosi & Harnad, 2001; Harnad, 1990; Vincent-Lamarre et al., 2016), an integrative model was expected to lead to a better approximation of human semantic similarity judgments as compared to the use of each estimate taken separately. Here though, we found no significant difference in performance as compared to the experiential model.

Our results support the theoretical assumption that the distributional information resulting from the computational manipulation of symbols alone is insufficient for making those symbols interpretable (Andrews et al., 2014; Cangelosi & Harnad, 2001; Harnad, 1990; Searle, 1980; Vigliocco et al., 2004). As postulated by the well-known Searlean Chinese room example, a computer could manipulate strings of Chinese symbols so well to perform as convincingly as a Chinese speaker would, even passing the Turing test, yet it would not understand their meaning. As this would require knowledge of what these symbols are used to communicate about (Harnad, 1990; Searle, 1980). Rather, the grounding of arbitrary symbolic word forms in action and perception knowledge and corresponding brain systems holding this information, at least for some of the symbols (“grounding kernel”: Cangelosi & Harnad, 2001), is essential for building the referential links between the arbitrary and symbolic word forms and the actions, objects and entities in the external world they are used to talk about (Barsalou, 2008, 2017; Harnad, 1990, 2012; Pulvermüller, 2013, 2018; Searle, 1980). The present data are consistent with the view that both linguistic context and the action and object context, in which concepts are experienced since the early stages of language acquisition, critically contribute to coherent word meaning representations (Wittgenstein, 1953/2009).

Conclusions

We showed that distributional and experiential similarities between concrete written words is reflected by the similarities between word-elicited graded activation patterns in partly distinct networks of frontal and temporo-parietal regions. The resulting neurocognitive picture informs semantic theories suggesting that a representationally complete characterisation of the complexity of lexical meaning cannot be based on distributional information alone, but requires experiential knowledge about the grounding of symbol meaning in action and perception. In the present data, the conjunction of a limited amount of specific experiential properties, such as the dominant action-related and

visual attributes of action and object words, was sufficient for mapping the semantic spaces of the corresponding categories in category-specific motor and visual cortex, as well as multimodal connector hub areas, thus being consistent with the relevance of category-specific as well as category-general semantic grounding information (Binder et al., 2016; Fernandino et al., 2016, 2022; Pulvermüller et al., 2009). Still, distributional information contributed to the differentiation between categories as well, drawing upon slightly different and sometimes overlapping association areas (e.g. left inferior frontal cortex). Semantic similarity-mappings of the categories of action and object words were equally successful in frontal cortex with both grounded experiential and distributional models, whereas only the former model succeeded in similarity mapping of object words in temporal cortex. Further work is needed for determining to what extent the different sources of semantic information (e.g. distributional and experiential properties of words, and their interplay, as employed here) contribute to the mapping of the representational content of language in relevant brain networks for different types of concepts, and beyond single words (e.g. sentence, discourse, naturalistic contexts).

Acknowledgement

We would like to thank Olaf Hauk, and three anonymous reviewers for their valuable comments on an earlier version of our paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Francesca Carota  <http://orcid.org/0000-0002-6315-9858>

References

- Anderson, A. J., Binder, J. R., Fernandino, L., Humphries, C. J., Conant, L. L., Raizada, R. D., Lin, F., & Lalor, E. C. (2019). An integrated neural decoder of linguistic and experiential meaning. *Journal of Neuroscience*, 39(45), 8969–8987. <https://doi.org/10.1523/JNEUROSCI.2575-18.2019>
- Anderson, A. J., Kiela, D., Binder, J. R., Fernandino, L., Humphries, C. J., Conant, L. L., Raizada, R. D. S., Grimm, S., & Lalor, E. C. (2021). Deep artificial neural networks reveal a distributed cortical network encoding propositional sentence-level meaning. *Journal of Neuroscience*, 41(18), 4100–4119. <https://doi.org/10.1523/JNEUROSCI.1152-20.2021>
- Anderson, A. J., McDermott, K., Rooks, B., Heffner, K. L., Dodel-Feder, D., & Lin, F. V. (2020). Decoding individual identity from brain activity elicited in imagining common

- experiences. *Nature Communications*, 11(1), 5916. <https://doi.org/10.1038/s41467-020-19630-y>
- Andrews, M., Frank, S., & Vigliocco, G. (2014). Reconciling embodied and distributional accounts of meaning in language. *Topics in Cognitive Science*, 6(3), 359–370. <https://doi.org/10.1111/tops.12096>
- Baayen, H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Linguistic Data Consortium.
- Baroni, M., Dinu, G., & Kruszewski, G. (2014). Don't count, predict! A systematic comparison of context-counting vs. Context-predicting semantic vectors. In *Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 238–247). Association for Computational Linguistics. <https://aclanthology.org/P14-1023/>
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Science*, 22(4), 577–609. <https://doi.org/10.1017/S0140525X99002149>.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59(1), 617–645. <https://doi.org/10.1146/annurev.psych.59.103006.093639>
- Barsalou, L. W. (2017). What does semantic tiling of the cortex tell us about semantics? *Neuropsychologia*, 105, 18–38. <https://doi.org/10.1016/j.neuropsychologia.2017.04.011>
- Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg, & A. C. A. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 245–283). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199217274.003.0013>
- Bedny, M., & Caramazza, A. (2011). Perception, action, and word meanings in the human brain: The case from action verbs. *Annals of the New York Academy of Sciences*, 1224(1), 81–95. <https://doi.org/10.1111/j.1749-6632.2011.06013.x>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Ser. B*, 57(1), 289–300.
- Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M., & Desai, R. H. (2016). Toward a brain-based componential semantic representation. *Cognitive Neuropsychology*, 33(3–4), 130–174. <https://doi.org/10.1080/02643294.2016.1147426>
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Science*, 15(11), 527–536. <https://doi.org/10.1016/j.tics.2011.10.001>
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767–2796. <https://doi.org/10.1093/cercor/bhp055>
- Bonner, M. F., & Epstein, R. A. (2021). Object representations in the human brain reflect the co-occurrence statistics of vision and language. *Nature Communications*, 12(1), 4081. <https://doi.org/10.1038/s41467-021-24368-2>
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59. [http://dx.doi.org/10.1016/0005-7916\(94\)90063-9](http://dx.doi.org/10.1016/0005-7916(94)90063-9)
- Bruffaerts, R., Dupont, P., Peeters, R., De Deyne, S., Storms, G., & Vandenberghe, R. (2013). Similarity of fMRI activity patterns in left perirhinal cortex reflects semantic similarity between words. *Journal of Neuroscience*, 33, 18597–18607.
- Cangelosi, A., & Harnad, S. (2001). The adaptive advantage of symbolic theft over sensorimotor toil: Grounding language in perceptual categories. *Evolution of Communication*, 4(1), 117–142. <https://doi.org/10.1075/eoc.4.1.07can>
- Carlson, T. A., Simmons, R. A., Kriegeskorte, N., & Slevc, L. R. (2014). The emergence of semantic meaning in the ventral temporal pathway. *Journal of Cognitive Neuroscience*, 26(1), 120–131. https://doi.org/10.1162/jocn_a_00458
- Carota, F., Kriegeskorte, N., Nili, H., & Pulvermüller, F. (2017). Representational similarity mapping of distributional semantics in left inferior frontal, middle temporal and motor cortex. *Cerebral Cortex*, 27(1), 294–309. <https://doi.org/10.1093/cercor/bhw379>
- Carota, F., Moseley, R., & Pulvermüller, F. (2012). Body-part-specific representations of semantic noun categories. *Journal of Cognitive Neuroscience*, 24(6), 1492–1509. http://dx.doi.org/10.1162/jocn_a_00219.
- Carota, F., Nili, H., Pulvermüller, F., & Kriegeskorte, N. (2021). Distinct fronto-temporal substrates of distributional and taxonomic similarity among words: Evidence from RSA of BOLD signals. *NeuroImage*, 1(224), 117408. <https://doi.org/10.1016/j.neuroimage.2020.117408>
- Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, 2(10), 913–919. <https://doi.org/10.1038/13217>
- Clarke, A., & Tyler, L. (2014). Object-specific semantic coding in human perirhinal cortex. *Journal of Neuroscience*, 34(14), 4766–4775. <https://doi.org/10.1523/JNEUROSCI.2828-13.2014>
- Coutanche, M. N., Solomon, S. H., & Thompson-Schill, S. L. (2016). A meta-analysis of fMRI decoding: Quantifying influences on human visual population codes. *Neuropsychologia*, 82, 134–141. <https://doi.org/10.1016/j.neuropsychologia.2016.01.018>
- Crutch, S. J., Troche, J., Reilly, J., & Ridgway, G. R. (2013). Abstract conceptual feature ratings: The role of emotion, magnitude, and other cognitive domains in the organization of abstract conceptual knowledge. *Frontiers in Human Neuroscience*, 7, 186. <http://dx.doi.org/10.3389/fnhum.2013.00186>.
- Damasio, A. R. (1989a). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation*, 1(1), 123–132.
- Damasio, A. R. (1989b). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, 33(1–2), 25–62.
- Damasio, A. R., & Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences, USA*, 90(11), 4957–4960. <https://doi.org/10.1073/pnas.90.11.4957>
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, 380(6574), 499–505. <http://dx.doi.org/10.1038/380499a0>.
- Davey, J., Rueschemeyer, S.-A., Costigan, A., Murphy, N., Krieger-Redwood, K., Hallam, G., & Jefferies, E. (2015). Shared neural processes support semantic control and action understanding. *Brain and Language*, 142, 24–35. <https://doi.org/10.1016/j.bandl.2015.01.002>

- Desai, R. H., Binder, J. R., Conant, L. L., & Seidenberg, M. S. (2009). Activation of sensory-motor areas in sentence comprehension. *Cerebral Cortex*, 20(2), 468–478. <https://doi.org/10.1093/cercor/bhp115>
- Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *Journal of Neuroscience*, 33(48), 18906–18916. <https://doi.org/10.1523/JNEUROSCI.3809-13.2013>
- Dreyer, F., Picht, T., Frey, D., Vijkoczy, P., & Pulvermüller, F. (2020). The functional relevance of dorsal motor systems for processing tool nouns—evidence from patients with focal lesions. *Neuropsychologia*, 141, 107384. <https://doi.org/10.1016/j.neuropsychologia.2020.107384>
- Elli, G. V., Lane, C., & Bedny, M. (2019). A double dissociation in sensitivity to verb and noun semantics across cortical networks. *Cerebral Cortex*, 29(11), 4803–4817. <https://doi.org/10.1093/cercor/bhz014>
- Fernandino, L., Binder, J. R., Desai, R. H., Pendl, S. L., Humphries, C. J., Gross, W. L., Conant, L. L., & Seidenberg, M. S. (2016). Concept representation reflects multimodal abstraction: A framework for embodied semantics. *Cerebral Cortex*, 26(5), 2018–2034. <https://doi.org/10.1093/cercor/bhv020>
- Fernandino, L., Colin, J., Humphries, C. J., Seidenberg, M. S., Gross, W. L., Conant, L. L., & Binder, J. R. (2015). Predicting brain activation patterns associated with individual lexical concepts based on five sensory-motor attributes. *Neuropsychologia*, 76, 17–26. <https://doi.org/10.1016/j.neuropsychologia.2015.04.009>
- Fernandino, L., Tong, J. Q., Conant, L. L., Humphries, C. J., & Binder, J. R. (2022). Decoding the information structure underlying the neural representation of concepts. *Proceedings of the National Academy of Sciences*, 119(6), e2108091119. <https://doi.org/10.1073/pnas.2108091119>
- Forseth, K. J., Kadipasaoglu, C. M., Conner, C. R., Hickok, G., Knight, R. T., & Tandon, N. (2018). A lexical semantic hub for heteromodal naming in middle fusiform gyrus. *Brain*, 141(7), 2112–2126. <https://doi.org/10.1093/brain/awy120>
- Fu, Z., Wang, X., Wang, X., Yang, H., Wang, J., Wei, T., Liao, X., Liu, Z., Chen, H., & Bi, Y. (2022). Different computational relations in language are captured by distinct brain systems. *Cerebral Cortex*, 33(4), 997–1013. <https://doi.org/10.1093/cercor/bhac117>
- Fuster, J. M. (1997). Network memory. *Trends in Neurosciences*, 20(10), 451–459. [https://doi.org/10.1016/S0166-2236\(97\)01128-4](https://doi.org/10.1016/S0166-2236(97)01128-4)
- Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22(3-4), 455–479. <https://doi.org/10.1080/02643290442000310>
- Garagnani, M., & Pulvermüller, F. (2016). Conceptual grounding of language in action and perception: A neurocomputational model of the emergence of category specificity and semantic hubs. *European Journal of Neuroscience*, 43(6), 721–737. <https://doi.org/10.1111/ejn.13145>
- Geschwind, N. (1965). Disconnection syndromes in animals and man. *Brain*, 88(2), 237–294; 585–644. <https://doi.org/10.1093/brain/88.2.237>
- Ghio, M., Vaghi, M. M. S., Perani, D., & Tettamanti, M. (2016). Decoding the neural representation of fine-grained conceptual categories. *NeuroImage*, 132, 93–103. <https://doi.org/10.1016/j.neuroimage.2016.02.009>
- Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*, 43(3), 379–401. <https://doi.org/10.1006/jmla.2000.2714>
- Grossman, M., Anderson, C., Khan, A., Avants, B., Elman, L., & McCluskey, L. (2008). Impaired action knowledge in amyotrophic lateral sclerosis. *Neurology*, 71(18), 1396–1401. <https://doi.org/10.1212/01.wnl.0000319701.50168.8c>
- Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Science*, 9(9), 416–423. <https://doi.org/10.1016/j.tics.2005.07.004>
- Hagoort, P. (2015). MUC (memory, unification, control): A model on the neurobiology of language beyond single word processing. *Neurobiology of Language*, 2016, 339–347. <https://doi.org/10.1016/B978-0-12-407794-2.00028-6>
- Hagoort, P. (2019). The meaning making mechanism(s) behind the eyes and between the ears. *Philosophical Transactions of the Royal Society B*, 375(1791), 20190301. <https://doi.org/10.1098/rstb.2019.0301>
- Harnad, S. (1990). The symbol grounding problem. *Physica*, D42, 335–346.
- Harnad, S. (2012). From sensorimotor categories and pantomime to grounded symbols and propositions. In K. R. Gibson, & M. Tallerman (Eds.), *The Oxford handbook of language evolution* (pp. 387–392). Oxford University Press.
- Hauk, O., Davis, M., & Pulvermüller, F. (2008). Modulation of brain activity by multiple lexical and word form variables in visual word recognition: A parametric fMRI study. *NeuroImage*, 42(3), 1185–1195. <https://doi.org/10.1016/j.neuroimage.2008.05.054>
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in the motor and premotor cortex. *Neuron*, 41(2), 301–307. [https://doi.org/10.1016/S0896-6273\(03\)00838-9](https://doi.org/10.1016/S0896-6273(03)00838-9)
- Hickok, G. (2014). *The myth of mirror neurons. The real neuroscience of communication and cognition*. Norton & Company.
- Hillis, A. E., Heidler-Gray, J., Newhart, M., Chang, S., Ken, L., & Bak, T. H. (2006). Naming and comprehension in primary progressive aphasia: The influence of grammatical word class. *Aphasiology*, 20(2-4), 246–256. <https://doi.org/10.1080/02687030500473262>
- Hillis, A. E., Oh, S., & Ken, L. (2004). Deterioration of naming nouns versus verbs in primary progressive aphasia. *Annals of Neurology*, 55(2), 268–275. <https://doi.org/10.1002/ana.10812>
- Hoffman, P., McClelland, J. L., & Lambon Ralph, M. A. (2018). Concepts, control, and context: A connectionist account of normal and disordered semantic cognition. *Psychological Review*, 125(3), 293–328. <https://doi.org/10.1037/rev0000094>
- Kemmerer, D., Rudrauf, D., Manzel, K., & Tranel, D. (2012). Behavioral patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions. *Cortex*, 48(7), 826–848. <https://doi.org/10.1016/j.cortex.2010.11.001>
- Kemmerer, D., & Tranel, D. (2003). A double dissociation between the meanings of action verbs and locative prepositions. *Neurocase*, 9(5), 421–435. <https://doi.org/10.1076/neur.9.5.421.16551>
- Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, 140(1), 14–34. <https://doi.org/10.1037/a0021446>

- Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2(1), 4–5. <https://doi.org/10.3389/neuro.01.016.2008>
- Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1), 42–55. <https://doi.org/10.1038/nrn.2016.150>
- Landauer, T. K. (1998). Learning and representing verbal meaning: The latent semantic analysis theory. *Current Directions in Psychological Science*, 7(5), 161–164. <https://doi.org/10.1111/1467-8721.ep10836862>
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211–240. <https://doi.org/10.1037/0033-295X.104.2.211>
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920–933. <https://doi.org/10.1038/nrn2532>
- LimeSurvey Project Team, Carsten Schmitz. (2013). *LimeSurvey: An Open Source survey tool*. LimeSurvey Project, Hamburg, Germany. <http://www.limesurvey.org>
- Liuzzi, A. G., Aglinskas, A., & Fairhall, S. L. (2020). General and feature-based semantic representations in the semantic network. *Scientific Reports*, 10(1), 1–12. <https://doi.org/10.1038/s41598-020-65906-0>
- Liuzzi, A. G., Bruffaerts, R., Peeters, R., Adamczuk, K., Keuleers, E., De Deyne, S., Storms, G., Dupont, P., & Vandenberghe, R. (2017). Cross-modal representation of spoken and written word meaning in left pars triangularis. *Neuroimage*, 150, 292–307. <https://doi.org/10.1016/j.neuroimage.2017.02.032>
- Louwerse, M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin and Review*, 15, 838–844.
- Louwerse, M. (2018). Knowing the meaning of a word by the linguistic and perceptual company it keeps. *Trends in Cognitive Science*, 10(3), 573–589.
- Louwerse, M. (2021). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science*, 3(2), 273–302. <https://doi.org/10.1111/j.1756-8765.2010.01106.x>
- Marino, B. F., Gallese, V., Buccino, G., & Riggio, L. (2012). Language sensorimotor specificity modulates the motor system. *Cortex*, 48(7), 849–856. <https://doi.org/10.1016/j.cortex.2010.12.003>
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58(1), 25–45. <https://doi.org/10.1146/annurev.psych.57.102904.190143>
- Martin, A., Wiggs, C., Ungerleider, L., & Haxby, L. V. (1996). Neural correlates of category-specific knowledge. *Nature*, 379(6566), 649–652. <https://doi.org/10.1038/379649a0>
- Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in Neuroscience*, 32(7), 376–382. <https://doi.org/10.1016/j.tins.2009.04.002>
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 26, <https://doi.org/10.48550/arXiv.1310.4546>
- Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y. T., Fryer, T. D., Williams, G. B., Hodges, J. R., & Nestor, P. J. (2010). What the left and right anterior fusiform gyri tell us about semantic memory. *Brain*, 133(11), 3256–3268. <https://doi.org/10.1093/brain/awq272>
- Mirman, D., Landrigan, J.-F., & Britt, A. E. (2017). Taxonomic and thematic semantic systems. *Psychological Bulletin*, 143(5), 499–520. <https://doi.org/10.1037/bul0000092>
- Mitchell, D. J., & Cusack, R. (2016). Semantic and emotional content of imagined representations in human occipitotemporal cortex. *Scientific Reports*, 6(20232). <https://doi.org/10.1038/srep20232>
- Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K. M., Malave, V. L., Mason, R. A., & Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *Science*, 320(5880), 1191–1195. <https://doi.org/10.1126/science.1152876>
- Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A toolbox for representational similarity analysis. *PLoS in Computational Biology*, 10(4), e1003553. <https://doi.org/10.1371/journal.pcbi.1003553>
- Osgood, C. E., May, W. H., & Miron, M. S. (1975). *Cross-cultural universals of affective meaning*. University of Illinois Press.
- Paivio, A. (1971). *Imagery and verbal processes*. Holt, Rinehart and Winston.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), 976–987. <https://doi.org/10.1038/nrn2277>
- Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543).
- Pereira, F., Lou, B., Pritchett, B., Ritter, S., Gershman, S. J., Kanwisher, N., Botvinick, M., & Fedorenko, E. (2018). Toward a universal decoder of linguistic meaning from brain activation. *Nature Neuroscience*, 9, 963. <https://doi.org/10.1038/s41467-018-03068-4>
- Perini, F., Caramazza, A., & Peelen, M. V. (2014). Left occipitotemporal cortex contributes to the discrimination of tool-associated hand actions: fMRI and TMS evidence. *Frontiers in Human Neuroscience*, 8, 591. <https://doi.org/10.3389/fnhum.2014.00591>
- Pillon, A., & d'Honincthun, P. (2011). A common processing system for the concepts of artifacts and actions? Evidence from a case of a disproportionate conceptual impairment for living things. *Cognitive Neuropsychology*, 28(1), 1–43. <https://doi.org/10.1080/02643294.2011.615828>
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*, 10(1), 15–35. <https://doi.org/10.1006/nimg.1999.0441>
- Popham, S. F., Huth, A. G., Bilenko, N. Y., Deniz, F., Gao, J. S., Nunez-Elizalde, A. O., & Gallant, J. L. (2021). Visual and linguistic semantic representations are aligned at the border of human visual cortex. *Nature Neuroscience*, 24(11), 1628–1636. <https://doi.org/10.1038/s41593-021-00921-6>
- Price, A. R., Bonner, M. F., Peelle, J. M., & Grossman, M. (2015). Converging evidence for the neuroanatomic basis of combinatorial semantics in the angular gyrus. *Journal of Neuroscience*, 35(7), 3276–3284.

- Price, C. J. (2000). The anatomy of language: Contributions from functional neuroimaging. *Journal of Anatomy*, 197(3), 335–359.
- Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, 22(2), 253–336. <https://doi.org/10.1017/S0140525X9900182X>
- Pulvermüller, F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Science*, 17(9), 458–470. <https://doi.org/10.1016/j.tics.2013.06.004>
- Pulvermüller, F. (2018). The case of CAUSE: Neurobiological mechanisms for grounding an abstract concept. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1752), 20170129. <https://doi.org/10.1098/rstb.2017.0129>
- Pulvermüller, F., Kherif, F., Hauk, O., Mohr, B., & Nimmo-Smith, I. (2009). Cortical cell assemblies for general lexical and category-specific semantic processing as revealed by fMRI cluster analysis. *Human Brain Mapping*, 30(12), 3837–3850. <https://doi.org/10.1002/hbm.20811>
- Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience*, 17(6), 884–892. <https://doi.org/10.1162/0898929054021111>
- Pulvermüller, F., Tomasello, R., Henningsen-Schomers, M. R., & Wennekers, T. (2021). Biological constraints on neural network models of cognitive function. *Nature Reviews Neuroscience*, 22(8), 488–502. <http://dx.doi.org/10.1038/s41583-021-00473-5>
- Reilly, J., Peelle, J. E., Garcia, A., & Crutch, S. J. (2016). Linking somatic and symbolic representation in semantic memory: The dynamic multilevel reactivation framework. *Psychonomic Bulletin and Review*, 23(4), 1002–1014. <https://doi.org/10.3758/s13423-015-0824-5>
- Rueschmeyer, S., van Rooji, D., Lindemann, O., Willems, R., & Bekkering, H. (2010). The function of words: Distinct neural correlates for words denoting differently manipulable objects. *Journal of Cognitive Neuroscience*, 22(8), 1844–1851. <https://doi.org/10.1162/jocn.2009.21310>
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–457. <https://doi.org/10.1017/S0140525X00005756>
- Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences, USA*, 94(26), 14792–14797. <https://doi.org/10.1073/pnas.94.26.14792>
- Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: Converging evidence from lesion and connectivity analyses. *Frontiers in System Neuroscience*, 201(5), 1. <https://doi.org/10.3389/fnsys.2011.00001>
- Tyler, L. K., Chiu, S., Zhuang, J., Randall, B., Devreux, B. J., Wright, P., Clarke, A., & Taylor, K. I. (2013). Objects and categories: Feature statistics and object processing in the ventral stream. *Journal of Cognitive Neuroscience*, 25(19), 1723–1735. https://doi.org/10.1162/jocn_a_00419
- Vannuscorps, G., & Pillon, A. (2011). A domain-specific system for representing knowledge of both man-made objects and human actions. Evidence from a case with an association of deficits. *Neuropsychologia*, 49(9), 2321–2341. <https://doi.org/10.1016/j.neuropsychologia.2011.04.006>
- Vigliocco, G., Kousta, S.-T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The neural representation of abstract words: The role of emotion. *Cerebral Cortex*, 24(7), 1767–1777. <http://dx.doi.org/10.1093/cercor/bht025>
- Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, 48(4), 422–488. <https://doi.org/10.1016/j.cogpsych.2003.09.001>
- Vincent-Lamarre, P., Massé, A. B., Lopes, M., Lord, M., Marcotte, O., & Harnad, S. (2016). The latent structure of dictionaries. *Topics in Cognitive Science*, 8(3), 625–659. <https://doi.org/10.1111/tops.12211>
- Vukovic, N., & Shtyrov, Y. (2019). Learning with the wave of the hand: Kinematic and TMS evidence of primary motor cortex role in category-specific encoding of word meaning. *Neuroimage*, 202, 116179. <https://doi.org/10.1016/j.neuroimage.2019.116179>
- Wang, X., Men, W., Gao, J., Caramazza, A., & Bi, Y. (2020). Two forms of knowledge representations in the human brain. *Neuron*, 107(2), 383–393. <https://doi.org/10.1016/j.neuron.2020.04.010>
- Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairment. *Brain*, 107(3), 829–854. <https://doi.org/10.1093/brain/107.3.829>
- Wittgenstein, L. (1953/2009). *Philosophical investigations*. Wiley-Blackwell.
- Xiang, H., Fonteijn, H. M., Norris, D. G., & Hagoort, P. (2010). Topographical functional connectivity pattern in the perisylvian language networks. *Cerebral Cortex*, 20(3), 549–560. <https://doi.org/10.1093/cercor/bhp119>
- Xu, Y., Wang, X., Wang, X., Men, W., Gao, J.-H., & Bi, Y. (2018). Doctor, teacher, and stethoscope: Neural representation of different types of semantic relations. *Journal of Neuroscience*, 38(13), 3303–3317. <https://doi.org/10.1523/JNEUROSCI.2562-17.2018>
- Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. *Trends in Cognitive Sciences*, 18(5), 229–234. <http://dx.doi.org/10.1016/j.tics.2014.02.008>