Simulating semantic dementia in a brain-constrained model of action and object words learning

Aleksei Efremov[a*,✉], Anastasia Kuptsova[a*], Thomas Wennekers[b], Yury Shtyrov[a,c], Boris Gutkin[a,d], and Max Garagnani[e,f]

[a]Institute for Cognitive Neuroscience, National Research University Higher School of Economics, Moscow, Russian Federation
[b]School of Engineering, Computing and Mathematics, University of Plymouth, Plymouth, United Kingdom
[c]Center of Functionally Integrative Neuroscience (CFIN), Institute for Clinical Medicine, Aarhus University, Aarhus, Denmark
[d]Département d'Etudes Cognitives, École Normale Supérieure, Paris, France
[e]Department of Computing, Goldsmiths, University of London, London, United Kingdom
[f]Brain Language Laboratory, Department of Philosophy and Humanities, WE4, Freie Universitat Berlin, Berlin, Germany
[*]Equal contribution first authors

The nature of semantic knowledge – conceptual information stored in the brain – is highly debated in the field of cognitive science. Experimental and clinical data specify various cortical regions involved in the processing of meaning. Those include semantic hubs that take part in semantic processing in general as well as sensorimotor areas that process specific conceptual categories according to their modality. Biologically inspired neurocomputational models can help adjudicate between different theories about the exact roles of those regions in the functioning of the semantic system.

Here, we used an existing neuroanatomically constrained model of frontotemporal brain areas implicated in language acquisition and grounding. We adapted it to replicate and explain the effects of semantic dementia on word processing abilities. Semantic dementia is a disease characterized by semantic knowledge deterioration that correlates with neural damage in the anterior temporal lobe. The behavior of our model is in accordance with clinical data – namely, word recognition performance decreases as SD lesions progress, whereas word repetition abilities remain preserved, or are less affected. Furthermore, our model makes novel predictions about category-specific effects of SD – namely, our simulation results indicate that word processing should be more impaired for object- than for action-related words, and that white matter degradation should lead to more severe consequences than the same proportion of grey matter degradation.

The present results provide a mechanistic, cortical-level explanatory account of a range of language impairments as observed in clinical populations during the onset and progress of semantic dementia.

semantic dementia | semantic knowledge | neurocomputational model | Hebbian learning | semantic grounding

Correspondence: ***efremov.aleksei.2016@post.bio.msu.ru***

Introduction

Semantic knowledge is information about the meaning of concepts and words accumulated by an individual throughout their lifetime (1, 2). We constantly use it to behave properly in different situations, express our thoughts, understand others, and for many more purposes. Difficulties experienced by people with semantic memory impairments can be of a diverse nature — they can fail to recognize or produce speech (3, 4); they can forget how to use ordinary things, for example, an umbrella (5); or they can fail to understand and explain the source of the pain that they feel (4). This shows the importance of understanding the nature of such impairments and the nature of semantic knowledge itself.

Although conceptual information processing in the brain has been one of the central research topics in recent decades, its neural substrate, processing mechanisms, or even the substance of conceptual information remain an open debate (6). One line of reasoning posits that the sensorimotor system plays only a secondary role in semantic representations (7, 8) and that it is important to have one higher-order brain area where all concepts are stored in an amodal way (9, 10). Other groups claim that when we process a concept, our brain simulates the perceptual experience of the interaction with this concept (11). Finally, multiple scholars propose hybrid theories that suggest both the necessity of symbol grounding in perceptual and motor experiences and the existence of 'higher order' convergence areas that link together experiences from different modalities (1, 12–14).

A huge amount of evidence on the functioning of the semantic system comes from clinical studies. Semantic dementia (SD) is an exemplar semantic disorder that gave birth to numerous insights into conceptual information processing mechanisms (5, 15, 16). It is characterized by deterioration of semantic knowledge, the degree of which correlates with the severity of anterior temporal lobe (ATL) atrophy (3, 4, 17). Furthermore, according to some computational models of SD (15, 18) and transcranial magnetic stimulation (TMS) studies (19), ATL atrophy is considered as the cause of impairments of semantic memory. Several researchers have suggested that this evidence supports the existence of an amodal core of concept representations and diminishes the importance of concepts embodiment (9, 20, 21). This line of thinking stems from the observation that semantic memory decline in SD patients is not concept selective (i.e. different semantic categories are affected) and ATL is not part of the sensory and motor systems, therefore ATL may be the core place where concepts from different categories are stored in an amodal way.

On the other hand, amodal representations cannot fully ex-

plain the data on the correlation between activation of different cortical sites and comprehension of specific conceptual categories. Patients with lesions in the motor and premotor cortex often suffer from impaired comprehension of action-related words compared with object-related ones, while patients with lesions in the visual cortex have problems with comprehension of object-related words (for a review, see (1, 22). Furthermore, it was shown that TMS of the hand motor areas could enhance comprehension of the hand-related words, while stimulation of the leg areas could have a similar effect on the leg-related words (23). Other TMS studies show that comprehension and retrieval of different categories of concepts (e.g., animals versus tools) correlate with different cortical stimulation sites (in this example, occipital and inferior temporal areas versus premotor areas, accordingly; for reviews see (24, 25). This line of reasoning posits that the neuronal networks integrating sensory and motor patterns that emerge as a result of concomitant action and perception experiences during word meaning and concept acquisition could be the basis for semantic representations. More specifically, Kiefer and Pulvermüller (1) suggest that concept representations are distributed over modality-specific sensorimotor areas and non-specific (or modality general) "connector hubs", which provide the neuroanatomical "bridge" between the different modality-specific cortices. Brain connectivity data show that relevant primary motor and sensory areas are not directly linked, but can only communicate by virtue of mediating areas (connector hub). Therefore, conceptual representations must necessarily be distributed over modality-specific sensorimotor areas and general connector hubs. They argue that this hypothesis follows directly from the perceptual origin of semantic knowledge acquisition and the neurobiological properties of the brain (14, 25). When we acquire the meaning of a new concrete word or concept, this usually occurs through different perceptual modalities simultaneously (for example, a mother teaches her child 'This is an apple' by showing an apple, so in this case, auditory and visual perception occurs at the same time). This co-occurrence could establish robust links between neural networks subserving different perceptual and motor modalities through Hebbian learning mechanisms (26).

In this study, we extend a neuroanatomically constrained computational model of word acquisition and grounding (14) to reflect neural mechanisms associated with semantic dementia. Using this model, we attempt to explain existing data about SD syndrome, namely, why word recognition abilities decline with an increase in the disease severity while word repetition abilities remain relatively preserved. In the extended model, we implement two types of neural anomalies reported to be associated with SD —damage to the grey matter and white matter of the ATL (3, 4, 17) — and compare the model dynamics resulting from these two alterations. We also compare the degradation dynamics of object-related words with the degradation dynamics of action-related words during the simulated SD progression.

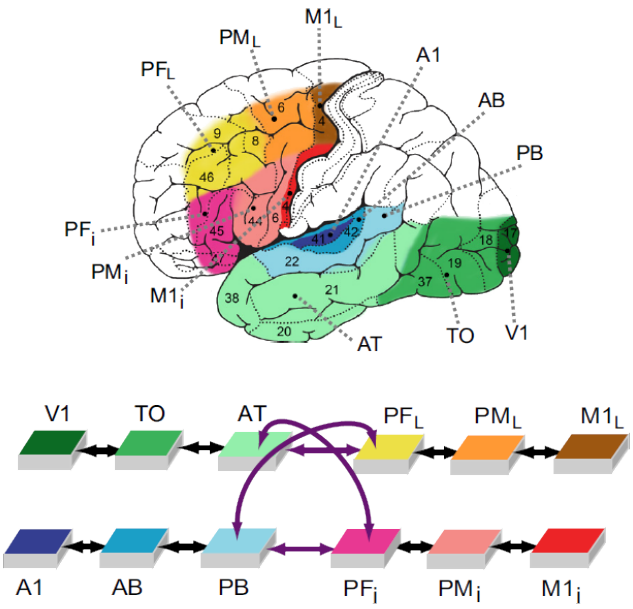

Fig. 1. Macrostructure of the model. Depicted here are 12 areas of the network and corresponding brain areas (coded by color). There are four 'zones' of three different modalities, with three areas in each zone. The 'auditory' zone comprises the superior and lateral auditory areas: primary auditory area (A1), auditory belt (AB) and auditory parabelt (PB). The 'visual' zone comprises the inferior temporo-occipital areas: primary visual area (V1), temporo-occipital area (TO) and anterior temporal area (AT). The motor cortex is represented by two zones – one corresponds to the articulatory movements, and the other (which we refer to as just the 'motor' zone) to nonarticulatory movements. The 'articulatory' zone comprises the inferior frontal areas: inferior primary motor area (M1i), inferior premotor area (PMi) and inferior prefrontal area (PFi). The motor zone comprises the superior-lateral frontal areas: lateral primary motor area (M1L), lateral premotor area (PML), and dorsolateral prefrontal cortex (PFL). We differentiate between the perisylvian cortex (auditory and articulatory areas) and the extrasylvian cortex (visual and motor areas). Black arrows indicate connections between adjacent areas, while the purple arrows indicate long-distance connections.

Methods

Overview of the model. To address the question of ATL involvement in the process of semantic knowledge deterioration during SD, we implemented a neurobiologically constrained model that replicates selected cortical areas and their connectivity. The model simulates 12 areas of the left hemisphere cortex (Fig.1). Six of them are perisylvian areas that are involved in language production and comprehension. The other six areas are believed to be involved in transferring and processing semantically relevant information during word meaning acquisition and comprehension / recognition – in the remainder of this article, they are referred to as extrasylvian areas (9, 11, 27, 28). The model is further divided into four modality-specific 'zones', with each zone containing three cortical areas: a primary sensorimotor cortex and adjacent 'higher' secondary and multimodal regions that have strong neuroanatomical links with this primary cortex (see subsection 'Connectivity of the Simulated Brain Areas' below).

Each area consists of two neuronal layers — an excitatory (e-cells) layer and an inhibitory (i-cells) one — with 625 (25x25) cells in each layer. Each i-cell corresponds to exactly one e-cell; a combination of an e-cell and an i-cell reflects approximately one cortical column that consists of pyrami-

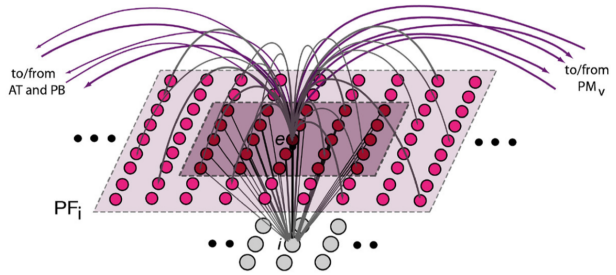

Fig. 2. Schematics of microconnectivity of one of the 7,500 excitatory neural elements (labelled 'e'). Within-area excitatory links (in grey) to and from 'cell' e are random and sparse, and limited to a local (19x19) neighborhood (light-pink shaded area).Lateral inhibition between e and neighboring excitatory elements is realized as follows: the underlying cell 'i' inhibits e, while its activity depends on the total excitatory input it receives from the 5x5 neighborhood around e (dark-purple shaded area); by means of analogous connections (not depicted), e can inhibit its neighbors.

dal excitatory neurons and inhibitory interneurons. Cells are modeled as graded-response neurons (see Appendix (Detailed Model Implementation) for specifications).

Functional and structural features of the model reflect well-documented properties of the human cortex such as:

1) known structure of the neuroanatomical links between the modeled brain systems;

2) sparse, patchy, and topographic between- and within-area connections, with the probability of the existence of a synaptic link between two cells falling off with their distance (29, 30) (see Appendix (Detailed Model Implementation) for specifications);

3) local lateral inhibition (Fig.2) and area-specific global regulation mechanisms (31);

4) Hebbian learning mechanisms that represent phenomena of long-term potentiation and depression (32);

5) neurophysiological dynamics of single cells including temporal summation of inputs, sigmoid transformation of membrane potentials into neuronal outputs, and adaptation (33);

6) presence of uniform white noise (simulating spontaneous baseline neuronal firing) in all parts of the network at all time points (34).

Connectivity of the simulated brain areas. Neuroanatomical studies show that adjacent cortical areas tend to be connected with each other (35, 36). We implemented such connections in all of the four zones of our model, between (1) inferior frontal areas PFi–PMi–M1i; (2) dorsolateral frontal areas PFL–PML–M1L (see also (37–40); (3) superior and lateral auditory areas A1–AB–PB (41–43); and (4) inferior temporo-occipital areas V1–TO–AT (44, 45).

Implementation of long-distance cortico-cortical links (purple arrows in Fig.1) that connect distant areas with each other is abundantly supported by evidence. Arcuate and uncinate fascicles provide the connections between anterior, inferior, and posterior-superior parts of the temporal cortex (areas AT and PB) and the inferior prefrontal cortex (PFi; (46–55). The extreme capsule connects the dorsolateral prefrontal cortex (PFL) to anterior and inferior temporal regions (AT; (56–58) and to the superior temporal cortex (PB; (54, 56, 59–61).

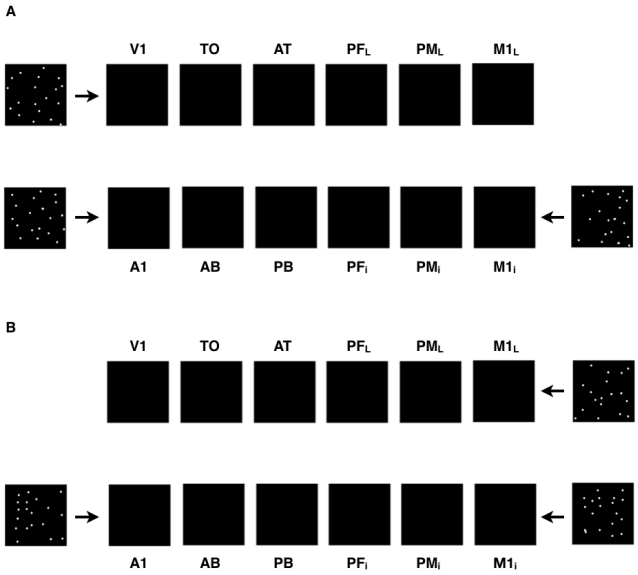

Fig. 3. Example of activity patterns presented to the network's primary areas to simulate word learning. A – The acquisition of object-related words was simulated by providing input patterns simultaneously to the model correlates of the auditory (A1), articulatory (M1i) and visual (V1) areas; B – Similarly, the simulation of action-related words acquisition involved presenting concomitant inputs to the auditory (A1), articulatory (M1i) and motor (M1L) areas. See text for details.

Learning words of different semantic categories. As in (14), the model was taught to differentiate between two semantic categories: action- and object-related words. The model simulates the acquisition of a word having object-related semantics by means of co-experiencing and associating auditory, articulatory, and visual patterns, as inputs to A1, M1i, and V1 are presented simultaneously during the learning phase (Fig.3A). To teach action-related semantics, input patterns are provided to auditory (A1), articulatory (M1i), and motor (M1L) primary areas (Fig.3B). Presenting a "pattern" as input involves activating 19 pre-specified cells in each of the aforementioned areas. The set of cells to be activated was randomly assigned for each word. Six word patterns for each of the two semantic categories were generated and taught; therefore, the model learned 12 different word patterns.

During the learning procedure, each word pattern is presented for 3000 trials. One trial lasts 16 time-steps. The next trial starts as soon as the activity in the network falls below the threshold but not earlier than 30 time-steps from the previous trial end. Trials are randomly shuffled. To reflect larger variability of activity in noninvolved primary areas (V1 for action-related words and M1L for object-related words), a randomly generated 19-cell noise pattern is provided to them in each trial. Finally, in addition to the uniform white noise constantly present in all areas of the network (simulating spontaneous neuronal activity), four primary areas are presented with additional "environmental" white noise both during training and testing.

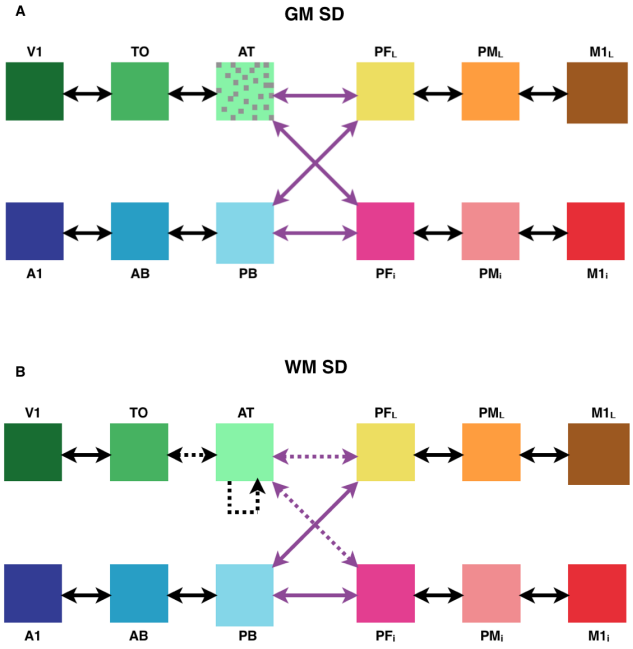

Fig. 4. Simulated semantic dementia lesions. Schematic representation indicating the type and location of the model lesions applied to simulate grey matter semantic dementia (GM SD, A) and white matter semantic dementia (WM SD, B). Dotted area represents inactivation of some portion of e-cells in AT area. Dashed lines represent inactivation of some portion of corresponding links.

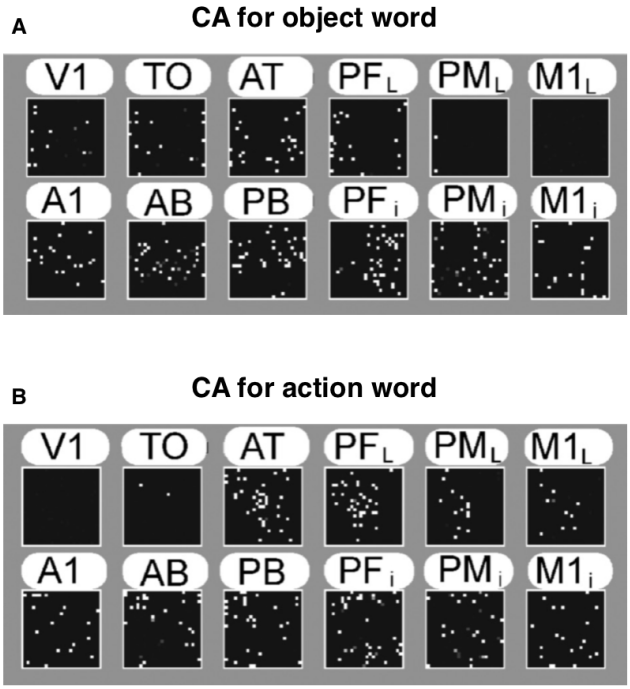

Fig. 5. Formation of cell assembly circuits in the network as a result of simulated word learning and semantic grounding. A – example of an object-related word CA circuit. B – example of an action-related word CA circuit. Note the differential cortical distribution of the two semantic types of word circuits in the extrasylvian (but not perisylvian) areas of the model.

Implementation of semantic dementia. Neuronal, anatomical and functional changes during SD can be described as a bilateral and asymmetric pattern of progressive atrophy and hypometabolism in white and grey matter (3, 4, 17). Grey matter changes are detected in the anterior temporal lobes — superior, middle, inferior temporal gyri, fusiform gyrus, temporal pole, parahippocampal gyrus — and progress to the basal ganglia and the medial orbitofrontal cortex during the course of the disease (for a review, see (17, 62). Changes in white matter are detected in regions that are connected to or are adjacent to the temporal lobes: left inferior fronto-occipital fasciculus; uncinate fasciculus and inferior longitudinal fasciculus bilaterally (for a review, see (17, 62). The most severe and the most robust atrophy in SD patients is detected in the anterior temporal pole and ventral parts of the ATL (20).

Following that evidence, we apply two types of degradation to simulate SD — grey matter degradation (GM SD) and white matter degradation (WM SD). To simulate GM SD we inactivate e-cells in the AT area (Fig.4A) and to simulate WM damage we remove connections to and from those cells (both within-area and between-areas; Fig.4B). We also use three severity levels for each degradation type — 30%, 60%, and 90% loss of matter — to simulate effects of the progressive nature of this disorder (3, 4, 17).

First, we trained 13 nets using the learning procedure described above, then we copied the resulting trained nets, and then the copies underwent one of the two degradation types (GM SD or WM SD) of the particular severity level (30%, 60%, or 90% loss of matter). Thus, we replicated the situation with SD patients who acquired semantic knowledge before the onset of the disease.

Cell assemblies and word recognition procedure. Previous results obtained with analogous architectures (14, 63, 64) showed that the learning procedure described above leads to the formation of distributed associative circuits or 'cell assemblies' (CA, Fig.5). CAs are sets of cells that are "...more strongly connected to each other than to other" cells (65). Once developed, they behave as discrete functional units with two quasi-stable states: 'on' and 'off' (63, 64, 66). Thus, CAs are stable stimulus-specific distributed memory circuits that emerge in the network as a result of learning and exhibit complete reactivation (or 'ignition') in response to the presentation of the stimulus, or part of it (or even spontaneously, due to noise-driven activity accumulation (67). After the training phase (see section 'Learning words of different semantic categories'), we applied a standard procedure to identify and quantify the cells that make up the CA circuits corresponding to each word stimulus (14, 63, 68). In the present model, spontaneously formed CAs linked word forms to aspects of those words meaning acquired through (simulated) sensory-motor areas (14).

Crucially, before and after lesions we assessed the ability of the network to "recognize" words. To do that, the auditory component of each word pattern was presented once (for 2 time-steps) to the A1 area only while no other input was provided. The activity of each e-cell was recorded during stimulus presentation and for the following 15 time-steps. An e-cell was considered as a responsive CA cell if its time-averaged activity during this period reached a given thresh-

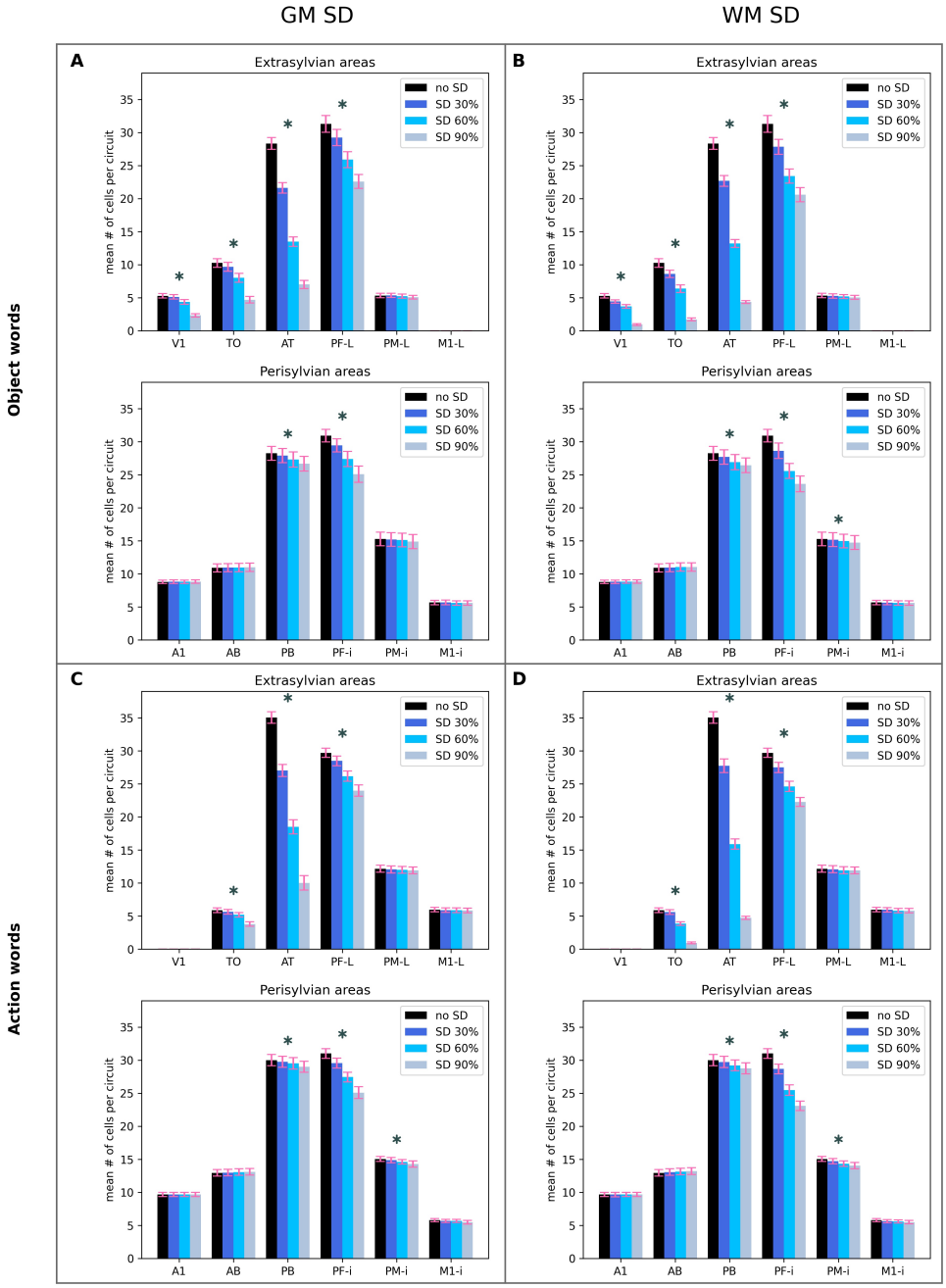

Fig. 6. Network response during simulated auditory word recognition. The average number (in 13 simulations) of input-specific cells exhibiting above threshold activity in response to presentation of a learned auditory stimulus ("word sound") to the model correlate of auditory cortex is plotted as a function of area, for intact networks (no SD) and three different degrees of simulated SD lesions. A, B: responses to object words. C, D: responses to action words. A, C: simulated grey-matter (GM) SD. B, D: simulated white-matter (WM) SD. See main text for details. Error bars indicate standard error of the mean. * – difference between no SD and 90% SD cases, p<0.001, t-test.

old (ϑ); this threshold is specified for each word (w) and each area (a) separately in the following way:

$$\theta(\omega, a) = \gamma \max_{e \in a} \overline{O(e,t)_\omega} \quad \textbf{(1)}$$

meaning that the time-averaged output is calculated for each e-cell (e) in each area (a) per each word (w). The threshold for the area for the word is equal to the fraction (γ) of the time-averaged output of the most active cell in this area for this word.

Further details of model implementation are given in the Appendix (Detailed Model Implementation).

Results

Main Hypothesis Testing. Figure 6 shows the number of CA cells in the extrasylvian and perisylvian areas of the network exhibiting a significant response (as defined by the recognition procedure, see 'Methods') as a function of SD severity. As can be seen, these results suggest that the extrasylvian portions of the CA circuits degrade more rapidly than the perisylvian portions. This difference was expected and is a direct consequence of the WM and GM lesions being applied to the AT area, which is an important component of the extrasylvian system. The qualitatively observed differential severity of degradation in the extrasylvian CA circuit was confirmed by a 2-way RM ANOVA with factors ExtraPeri (two levels — extrasylvian and perisylvian areas) and Severity (four levels — no SD, 30% SD, 60% SD, and 90%

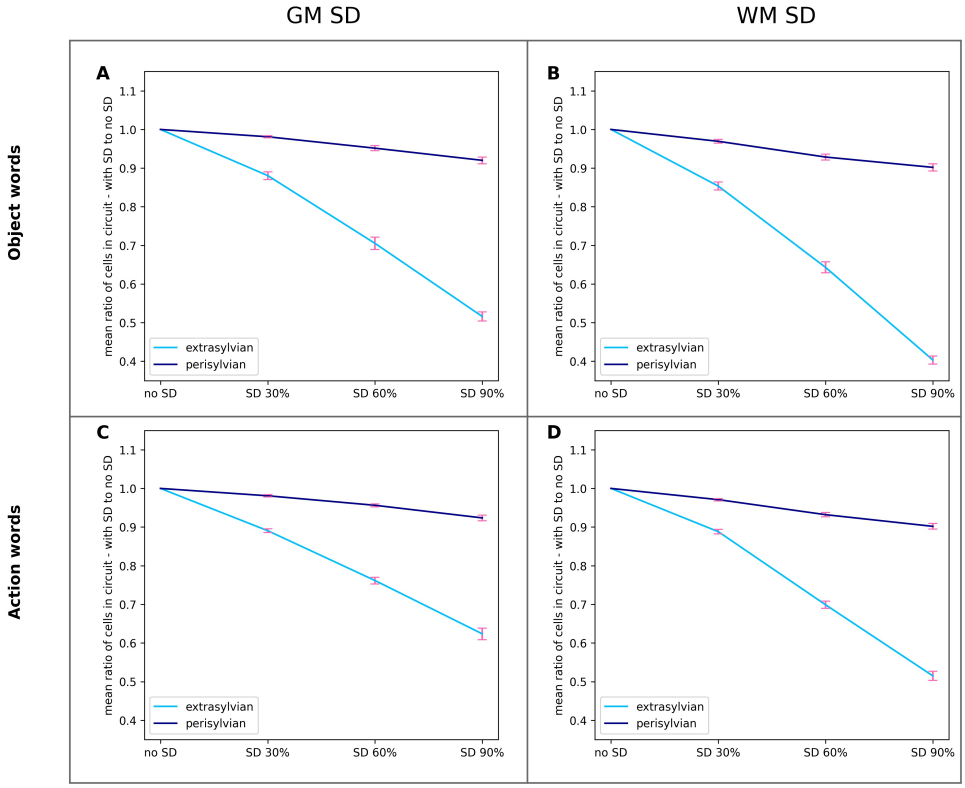

Fig. 7. Per-system network response during simulated auditory word recognition. Same data as in Figure 6 but grouped by extra- (semantic) or peri- (phonological) - sylvian system. The ratios plotted are the percentage of word-circuit-specific reactivated cells relative to the total number of CA cells (across the system) reactivated in the intact-network (no-SD) condition. See the text for details. Error bars indicate the standard error of the mean.

SD), revealing a significant interaction between these factors ($F_{3,36} = 1166.29$, $p < 0.001$).

To further explore the differences between extrasylvian and perisylvian areas, we performed an analysis of the total numbers of CA cells under those areas in four conditions: object versus action words and GM SD versus WM SD. Figure 7 shows how the number of responsive cells in the CA circuits decreases with SD progression — from no SD through 30% and 60% SD to 90% SD — in four conditions separately. Note that the data here are normalized by the number of CA cells before SD. The results show that in all four conditions the CA circuits deteriorate far more in extrasylvian areas than in perisylvian areas. For each of the four conditions, a 2-way RM ANOVA with factors ExtraPeri and Severity shows a significant interaction between factors (for four conditions: $F_{3,36} > 307.37$, $p < 0.001$).

Dependent samples t-tests confirm a significant difference in the number of responsive CA cells between no SD and 90% SD cases in each of the eight conditions (extra/peri x object/action x GM/WM; $t_{12} > 10.71$, $p < 0.001$ for all conditions). We also performed post hoc dependent samples t-tests to ensure that the number of CA cells in extrasylvian areas significantly differs from the number of CA cells in perisylvian areas in the most severe SD case (90% SD) for each of the four conditions (object/action x GM/WM; $t_{12} > 17.88$, $p < 0.001$). We have made a Bonferroni correction for all 12 comparisons (p-value threshold is 0.0042).

Combining this evidence, we infer that in all four conditions (object/action × GM/WM) the total number of CA cells in extrasylvian areas greatly decreases with the increase of SD severity – the average drop is 48.4% for the 90% SD case – while in perisylvian areas this number also declines, but much more slowly: the average drop is 8.7%.

To explore the SD impact on different areas, we performed post hoc dependent samples t-tests on the number of CA cells in no SD and 90% SD cases for each of the four conditions separately: object versus action and GM SD versus WM SD. All comparisons were Bonferroni corrected. Differences are presented as asterisks in Figure 6. In extrasylvian areas, SD leads to a strong decline in V1, TO, AT, and PFL areas, while PML and M1L remain intact. In perisylvian areas, SD leads to a strong decline in PFi and a mild but statistically significant decline in PB and PMi, while other areas are intact.

Additional Findings: model predictions. We asked what predictions our model would make about recognition of object versus action words when we implemented the structural alterations reflecting the pathology of SD. (Note that we assumed here that the number of responsive CA cells can be taken as an indicator of the network's ability to recognize words). Results shown in Fig.8 suggest that the network's ability to respond to the presentation of the auditory part of a learned "word" declines more dramatically for object- than action-related words. The statistical validity of this observation was confirmed by the result of 2-way RM ANOVAs, ran on the total number of CA cells in extrasylvian areas in GM and WM SD, separately with factors WordType and Severity. The ANOVA for GM SD showed a significant interaction of the two factors ($F_{3,36} = 7.04$, $p = 0.0012$), but only a trend toward significance ($F_{3,36} = 3.04$, $p = 0.067$) for WM SD. However, analysis on WM SD data revealed the presence of the main effects of both WordType ($F_{1,12} = 18.28$, $p = 0.001$)

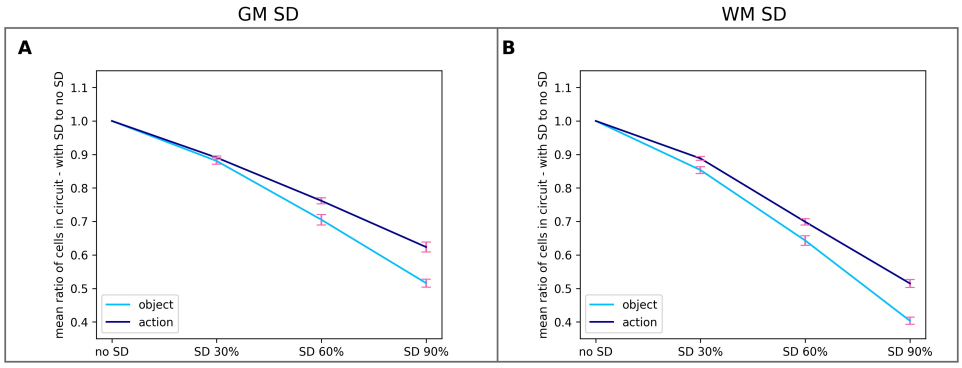

Fig. 8. Total network response during simulated auditory word recognition for different types of words. The ratios plotted are the percentage of (word-circuit)-specific reactivated cells relative to the total number of CA cells reactivated in the intact-network (no-SD) condition. See the text for details.

and Severity ($F_{3,36} = 1578.12$, $p < 0.001$). Therefore, both for GM and WM SD, the network's response to presentation of a known (simulated) word sound – i.e., its ability to recognize that item – decreased significantly more for object- than action-related words (48.3% decline for object words, 37.4% decline for action words, $t_{12} = 4.61$, $p = 0.001$ for GM SD; 59.5% decline for object words, 48.5% decline for action words, $t_{12} = 7.82$, $p < 0.001$ for WM SD).

We also noticed that recognition abilities appeared to decline more dramatically in WM SD than in GM SD (see Fig.9). To test this observation statistically, we ran 2-way RM ANOVAs for the total number of CA cells in extrasylvian areas (separately for object- and action-related words) with factors SD-Type and Severity. In both cases, the interaction between factors was significant ($F_{3,36} > 28.52$, $p < 0.001$), confirming that the network's decrease in recognition abilities for object- and action-related words is significantly greater for WM SD than for GM SD (for object words it is 59.5% decline in WM SD vs 48.3% decline in GM SD, $t_{12} = 8.49$, $p < 0.001$; for action words it is 48.5% decline in WM SD vs. 37.4% decline in GM SD, $t_{12} = 6.64$, $p < 0.001$).

Discussion

We used a neuroanatomically and biologically constrained neural model of the areas of the brain involved in language acquisition and semantic grounding to simulate semantic dementia (SD) lesions as progressive damage to the anterior temporal lobe – ATL –area. As a result of SD lesions, the "extrasylvian" parts of the word circuits (which include the ATL) degraded more rapidly than their perisylvian counterparts. A second striking result is the observed category-specific effects of SD lesions – namely, our simulations indicate that word processing abilities should be more impaired for object- than for action-related words; furthermore, the model shows that white matter degradations should lead to stronger deficits than grey matter lesions of analogous severity.

These simulation results are consistent with, and explain existing data about the decline of language function in clinical populations during onset and progress of semantic dementia. Specifically, the observed dissociation between word recognition deficits and relatively preserved word repetition abilities observed in SD patients is explained here by the cortically distributed character of word circuits, consisting of cell assemblies (CA) that include neurons from both extrasylvian (semantic) and perisylvian (language) areas. Previous modeling results (14) suggested that while extrasylvian areas may encode semantic knowledge and information pertaining to aspects of word meaning (and hence provide a neural substrate for word and concept recognition processes), perisylvian areas mediate the formation and storage of associations between acoustic and articulatory features of syllables and words (necessary, e.g., during word – and non-word – repetition tasks). The changes in activity patterns that we observed after introducing SD to our model speak in favor of our main hypothesis: while the CA circuits that are needed for successful word repetition (located in the perisylvian areas) remain relatively intact, word recognition abilities are progressively impaired as SD lesions increase due to a strong reduction in responsive CA cells in the extrasylvian areas. We believe that these simulated cortical changes can explain the dynamics of SD patients, who experience a strong decline of word recognition abilities but largely retain word repetition skills (10, 69, 70). We suggest that such a dissociation between perisylvian and extrasylvian areas can be further tested in neuroimaging experiments in SD patients.

We should note that the above interpretation of the model results is based on a key assumption: we are postulating that a decrease in the number of responsive extrasylvian CA cells corresponds to a de facto deficit in the word recognition ability of the network. A large body of experimental data suggests that (extrasylvian) sensorimotor areas convey information on aspects of word meaning (71–73); thus, even if parts of the CA circuits in extrasylvian areas are still activating in response to a presentation of an auditory stimulus, we submit that this partial activation contains less semantic information about the word, and therefore this information deficit translates directly into a corresponding observable word recognition deficit.

The results of the simulations also enable us to make predictions and adjudicate between competing theories on the question of whether SD patients have category-specific or category-general recognition problems. Several studies have suggested that SD leads to category-general deficits (9, 21, 74); however, our model shows that recognition of object-related words should decrease more sharply than recognition of action-related words as SD progresses. Some studies confirm that object naming is more affected in SD patients than action naming (75), and others show similar results for le-

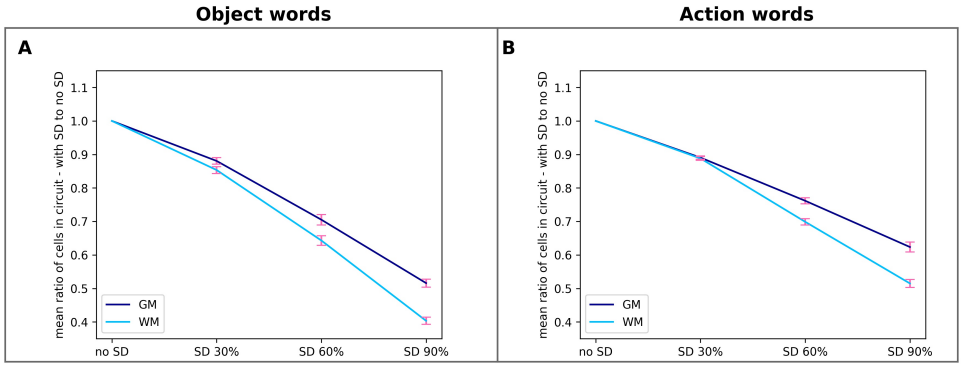

Fig. 9. Total network response during simulated auditory word recognition for different types of SD. The ratios plotted are the percentage of (word-circuit)-specific reactivated cells relative to the total number of CA cells reactivated in the intact-network (no-SD) condition. See the text for details.

sions in the ATL (76). Our model predicts (and explains) the emergence of this effect on the basis of the neuroanatomical links existing between brain areas relevant for object- and action-word processing, simulated here (see Fig.1), and specifically of the locus of the ATL itself. The ATL acts as the interface via which object-related information coming from the visual system links up with language circuits in perisylvian areas. If the grey matter within – or the white matter bundles that link to – this connector hub is lesioned, the word CA circuits that extend strongly into the visual system (i.e., object-related word circuits – see Fig.6) are severely affected, whereas CA circuits linking perisylvian and motor areas through the prefrontal hub area (PFL in the model – see Fig.1), i.e., action-related word circuits, will be only marginally affected. Unfortunately, most previous studies investigating this matter tend to group words based on grammatical categories – namely, nouns or verbs – rather than on the basis of semantics properties (77); hence, to validate the above predictions, further experimental testing is desirable.

To the best of our knowledge, this is the first brain-constrained model of SD where both types of degradation – white and grey matter – were implemented; previous computational studies only modeled degradation of links (WM) (15, 18). Our simulation results show that white matter degradation should lead to a more severe recognition decline than the same loss of grey matter. This additional prediction can be tested experimentally by investigating possible correlations between the grey matter / white matter loss ratio in SD and the disease severity. The results we have obtained on the distinction between grey and white matter degradation address an important feature of the cell assembly concept in our model. Although the constituent parts of CAs are individual cells, the emergence and integrity of those CAs are based on the links between them. Disabling a portion of cells naturally leads to the inactivation of some part of CA, but the same fraction of connections is more sparsely distributed and affects more cells. We can see from the results that even if not all links of some particular cell are inactivated – which would be the same as disabling the cell itself – those that are could be crucial for the cell's functioning as a part of CA.

Another aim of this paper has been to argue computationally that hybrid theories of semantic system organization offer the most promising theoretical approach to explain a wide range of experimental data. The present model can be considered an example of such hybrid paradigms: the primary areas in the action and perception brain systems are particularly important to provide sensorimotor co-experiences during word meaning acquisition; the central (semantic hub) areas enable these experiences to converge and be associated together into a single (word, or conceptual) circuit; the emergent semantic representations are distributed between the modality-specific sensorimotor areas and convergence hubs. One of the arguments of proponents of amodal theories is that ATL must be the place where concepts are stored in an amodal way because its degradation leads to semantic system deterioration (9, 21). However, as we show here, if word circuits are distributed across multiple diverse areas (all of which together support word recognition and processing), ATL damage can still lead to semantic impairments as observed in SD without this hub necessarily being the locus of amodal conceptual representations.

To conclude, the present neurobiologically constrained model explains – in terms of cortical mechanisms and neuroanatomical characteristics – the existing data on language impairments as observed in SD clinical populations. Furthermore, for the first time, simulated white or grey matter lesions to the anterior-temporal lobe area of the model enable making novel, theory-driven predictions about (i) differential effects of semantic dementia on word processing deficits (specifically, on object- vs action-related word recognition) and (ii) different degrees of language impairment depending on the type of tissue damaged by the neurodegenerative disease.

Bibliography

1. Markus Kiefer and Friedemann Pulvermüller. Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, 48(7):805–825, July 2012.
2. M R Quillan. Semantic memory. 1966.
3. John R Hodges and Karalyn Patterson. Semantic dementia: a unique clinicopathological syndrome, November 2007.
4. Maxime Montembeault, Simona M Brambati, Maria Luisa Gorno-Tempini, and Raffaella Migliaccio. Clinical, anatomical, and pathological features in the three variants of primary progressive aphasia: A review, August 2018.
5. Jaap M J Murre, Kim S Graham, and John R Hodges. Semantic dementia: Relevance to connectionist models of long-term memory, 2001.
6. Bradford Z Mahon and Gregory Hickok. Arguments about the nature of concepts: Symbols, embodiment, and beyond. *Psychon. Bull. Rev.*, 23(4):941–958, August 2016.
7. Bradford Z Mahon. What is embodied about cognition? *Language, Cognition and Neuroscience*, 30(4):420–429, April 2015.
8. Bradford Z Mahon and Alfonso Caramazza. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology-Paris*, 102(1):59–70, January 2008.
9. Karalyn Patterson, Peter J Nestor, and Timothy T Rogers. Where do you know what you

know? the representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.*, 8(12):976–987, December 2007.
10. Timothy T Rogers, Matthew A Lambon Ralph, Peter Garrard, Sasha Bozeat, James L McClelland, John R Hodges, and Karalyn Patterson. Structure and deterioration of semantic memory: a neuropsychological and computational investigation. *Psychol. Rev.*, 111(1): 205–235, January 2004.
11. Vittorio Gallese and George Lakoff. The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.*, 22(3):455–479, May 2005.
12. Lawrence W Barsalou. On staying grounded and avoiding quixotic dead ends, August 2016.
13. Jeffrey R Binder and Rutvik H Desai. The neurobiology of semantic memory. *Trends Cogn. Sci.*, 15(11):527–536, November 2011.
14. Max Garagnani and Friedemann Pulvermüller. Conceptual grounding of language in action and perception: a neurocomputational model of the emergence of category specificity and semantic hubs. *Eur. J. Neurosci.*, 43(6):721–737, March 2016.
15. Lang Chen, Matthew A Lambon Ralph, and Timothy T Rogers. A unified model of human semantic knowledge and its disorders. *Nat Hum Behav*, 1(3), March 2017.
16. John R Hodges, Naida Graham, and Karalyn Patterson. Charting the progression in semantic dementia: Implications for the organisation of semantic memory. *Memory*, 3(3-4): 463–495, September 1995.
17. Edoardo G Spinelli, Maria Luisa Mandelli, Zachary A Miller, Miguel A Santos-Santos, Stephen M Wilson, Federica Agosta, Lea T Grinberg, Eric J Huang, John Q Trojanowski, Marita Meyer, Maya L Henry, Giancarlo Comi, Gil Rabinovici, Howard J Rosen, Massimo Filippi, Bruce L Miller, William W Seeley, and Maria Luisa Gorno-Tempini. Typical and atypical pathology in primary progressive aphasia variants. *Ann. Neurol.*, 81(3):430–443, March 2017.
18. Taiji Ueno, Satoru Saito, Timothy T Rogers, and Matthew A Lambon Ralph. Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron*, 72(2):385–396, October 2011.
19. Gorana Pobric, Elizabeth Jefferies, and Matthew A Lambon Ralph. Category-specific versus category-general semantic impairment induced by transcranial magnetic stimulation. *Curr. Biol.*, 20(10):964–968, May 2010.
20. Matthew A Lambon Ralph, Elizabeth Jefferies, Karalyn Patterson, and Timothy T Rogers. The neural and computational bases of semantic cognition, December 2016.
21. Edouard Machery. The amodal brain and the offloading hypothesis. *Psychon. Bull. Rev.*, 23 (4):1090–1095, August 2016.
22. Lotte Meteyard, Sara Rodriguez Cuadrado, Bahador Bahrami, and Gabriella Vigliocco. Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex*, 48 (7):788–804, July 2012.
23. Friedemann Pulvermüller, Olaf Hauk, Vadim V Nikulin, and Risto J Ilmoniemi. Functional links between motor and language systems. *Eur. J. Neurosci.*, 21(3):793–797, February 2005.
24. L W Barsalou. Perceptual symbol systems. *Behav. Brain Sci.*, 22(4):577–609; discussion 610–60, August 1999.
25. Friedemann Pulvermüller. Brain reflections of words and their meaning. *Trends in Cognitive Sciences*, 5(12):517–524, December 2001.
26. Friedemann Pulvermüller. Words in the brain's language. *Behav. Brain Sci.*, 22(2):253–336, April 1999.
27. Joseph T Devlin, Paul M Matthews, and Matthew F S Rushworth. Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *J. Cogn. Neurosci.*, 15(1):71–84, January 2003.
28. E K Warrington and T Shallice. Category specific semantic impairments. *Brain*, 107 (Pt 3): 829–854, September 1984.
29. Valentino Braitenberg and Almut Schüz. *Cortex: Statistics and Geometry of Neuronal Connectivity*. Springer, Berlin, Heidelberg, 1998.
30. J H Kaas. Topographic maps are fundamental to sensory processing. *Brain Res. Bull.*, 44 (2):107–112, 1997.
31. Alan L Yuille and Davi Geiger. Winner-take-all mechanisms. In *The handbook of brain theory and neural networks*, pages 1056–1060. MIT Press, Cambridge, MA, USA, October 1998.
32. A Artola and W Singer. Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *Trends Neurosci.*, 16(11):480–487, November 1993.
33. G Matthews. Neurobiology: Molecules, cells and systems. *undefined*, 2001.
34. Gustavo Deco, Edmund T Rolls, and Ranulfo Romo. Stochastic dynamics as a principle of brain function. *Prog. Neurobiol.*, 88(1):1–16, May 2009.
35. Deepak N Pandya and Edward H Yeterian. Architecture and connections of cortical association areas. In Alan Peters and Edward G Jones, editors, *Association and Auditory Cortices*, pages 3–61. Springer US, Boston, MA, 1985.
36. M P Young, J W Scannell, G A Burns, and C Blakemore. Analysis of connectivity: neural systems in the cerebral cortex. *Rev. Neurosci.*, 5(3):227–250, July 1994.
37. T Arikuni, K Watanabe, and K Kubota. Connections of area 8 with area 6 in the brain of the macaque monkey. *J. Comp. Neurol.*, 277(1):21–40, November 1988.
38. Richard P Dum and Peter L Strick. Motor areas in the frontal lobe of the primate. *Physiol. Behav.*, 77(4):677–682, December 2002.
39. Richard P Dum and Peter L Strick. Frontal lobe inputs to the digit representations of the motor areas on the lateral surface of the hemisphere. *J. Neurosci.*, 25(6):1375–1386, February 2005.
40. M T Lu, J B Preston, and P L Strick. Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *J. Comp. Neurol.*, 341(3):375–392, March 1994.
41. Jon H Kaas and Troy A Hackett. Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. U. S. A.*, 97(22):11793–11799, October 2000.
42. D N Pandya. Anatomy of the auditory cortex. *Rev. Neurol.*, 151(8-9):486–494, August 1995.
43. Josef P Rauschecker and Biao Tian. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc. Natl. Acad. Sci. U. S. A.*, 97(22):11800–11806, October 2000.
44. C Distler, D Boussaoud, R Desimone, and L G Ungerleider. Cortical connections of inferior temporal area TEO in macaque monkeys. *J. Comp. Neurol.*, 334(1):125–150, August 1993.
45. H Nakamura, R Gattass, R Desimone, and L G Ungerleider. The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques. *J. Neurosci.*, 13(9): 3681–3691, September 1993.
46. Marco Catani, Derek K Jones, and Dominic H Ffytche. Perisylvian language networks of the human brain. *Ann. Neurol.*, 57(1):8–16, January 2005.
47. N Makris, J W Meyer, J F Bates, E H Yeterian, D N Kennedy, and V S Caviness. MRI-Based topographic parcellation of human cerebral white matter and nuclei II. rationale and applications with systematics of cerebral connectivity. *Neuroimage*, 9(1):18–45, January 1999.
48. Nikos Makris and Deepak N Pandya. The extreme capsule in humans and rethinking of the language circuitry. *Brain Struct. Funct.*, 213(3):343, December 2008.
49. Geoffrey J M Parker, Simona Luzzi, Daniel C Alexander, Claudia A M Wheeler-Kingshott, Olga Ciccarelli, and Matthew A Lambon Ralph. Lateralization of ventral and dorsal auditory-language pathways in the human brain. *Neuroimage*, 24(3):656–666, February 2005.
50. M Petrides and D N Pandya. Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *Eur. J. Neurosci.*, 16(2):291–310, July 2002.
51. Michael Petrides, Francesco Tomaiuolo, Edward H Yeterian, and Deepak N Pandya. The prefrontal cortex: Comparative architectonic organization in the human and the macaque monkey brains. *Cortex*, 48(1):46–57, January 2012.
52. James K Rilling. Comparative primate neuroimaging: insights into human brain evolution. *Trends Cogn. Sci.*, 18(1):46–55, January 2014.
53. James K Rilling, Matthew F Glasser, Todd M Preuss, Xiangyang Ma, Tiejun Zhao, Xiaoping Hu, and Timothy E J Behrens. The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat. Neurosci.*, 11(4):426–428, March 2008.
54. L M Romanski, B Tian, J Fritz, M Mishkin, P S Goldman-Rakic, and J P Rauschecker. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.*, 2(12):1131–1136, December 1999.
55. Lizabeth M Romanski. Representation and integration of auditory and visual stimuli in the primate ventral lateral prefrontal cortex. *Cereb. Cortex*, 17 Suppl 1:i61–9, September 2007.
56. Deepak N Pandya and Clifford L Barnes. Architecture and connections of the frontal lobe. In Ellen Perecman, editor, *The frontal lobes revisited , (pp*, volume 309, pages 41–72. The IRBN Press, xv, New York, NY, US, 1987.
57. L G Ungerleider, D Gaffan, and V S Pelak. Projections from inferior temporal cortex to prefrontal cortex via the uncinate fascicle in rhesus monkeys. *Exp. Brain Res.*, 76(3):473–484, August 1989.
58. M J Webster, J Bachevalier, and L G Ungerleider. Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cereb. Cortex*, 4(5): 470–483, September 1994.
59. Anthony Steven Dick, Byron Bernal, and Pascale Tremblay. The language connectome: new pathways, new concepts. *Neuroscientist*, 20(5):453–467, October 2014.
60. L M Romanski, J F Bates, and P S Goldman-Rakic. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.*, 403(2):141–157, January 1999.
61. Jeremy D Schmahmann, Deepak N Pandya, Ruopeng Wang, Guangping Dai, Helen E D'Arceuil, Alex J de Crespigny, and Van J Wedeen. Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography. *Brain*, 130(3): 630–653, February 2007.
62. Simona Maria Brambati, Serena Amici, Caroline A Racine, John Neuhaus, Zachary Miller, Jenny Ogar, Nina Dronkers, Bruce L Miller, Howard Rosen, and Maria Luisa Gorno-Tempini. Longitudinal gray matter contraction in three variants of primary progressive aphasia: A tenser-based morphometry study. *NeuroImage: Clinical*, 8:345–355, June 2015.
63. Max Garagnani, Thomas Wennekers, and Friedemann Pulvermüller. A neuroanatomically grounded hebbian-learning model of attention-language interactions in the human brain. *Eur. J. Neurosci.*, 27(2):492–513, January 2008.
64. Max Garagnani, Thomas Wennekers, and Friedemann Pulvermüller. Recruitment and consolidation of cell assemblies for words by way of hebbian learning and competition in a multi-layer neural network. *Cognit. Comput.*, 1(2):160–176, 2009.
65. Valentino Braitenberg. Cell assemblies in the cerebral cortex. In *Theoretical Approaches to Complex Systems*, pages 171–188. Springer Berlin Heidelberg, 1978.
66. Friedemann Pulvermüller and Max Garagnani. From sensorimotor learning to memory cells in prefrontal and temporal association cortex: A neurocomputational study of disembodiment. *Cortex*, 57:1–21, August 2014.
67. Max Garagnani and Friedemann Pulvermüller. Neuronal correlates of decisions to speak and act: Spontaneous emergence and dynamic topographies in a computational model of frontal and temporal areas. *Brain Lang.*, 127(1):75–85, October 2013.
68. Rosario Tomasello, Max Garagnani, Thomas Wennekers, and Friedemann Pulvermüller. A neurobiologically constrained cortex model of semantic grounding with spiking neurons and Brain-Like connectivity. *Front. Comput. Neurosci.*, 12:88, 2018.
69. Christine C Guo, Maria Luisa Gorno-Tempini, Benno Gesierich, Maya Henry, Andrew Trujillo, Tal Shany-Ur, Jorge Jovicich, Simon D Robinson, Joel H Kramer, Katherine P Rankin, Bruce L Miller, and William W Seeley. Anterior temporal lobe degeneration produces widespread network-driven dysfunction. *Brain*, 136(10):2979–2991, October 2013.
70. John R Hodges, Marina Martinos, Anna M Woollams, Karalyn Patterson, and Anna Lynne R Adlam. Repeat and point: Differentiating semantic dementia from progressive non-fluent aphasia. *Cortex*, 44(9):1265–1270, 2008.
71. Rebecca Fincher-Kiefer and Paul R D'Agostino. The role of visuospatial resources in generating predictive and bridging inferences. *Discourse Process.*, 37(3):205–224, May 2004.
72. Friedemann Pulvermüller. Brain mechanisms linking language and action. *Nat. Rev. Neurosci.*, 6(7):576–582, July 2005.
73. Rolf A Zwaan and Carol J Madden. Embodied sentence comprehension. In *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*, pages 224–245. Cambridge University Press, January 2005.
74. Matthew A Lambon Ralph, Christine Lowe, and Timothy T Rogers. Neural basis of category-specific semantic deficits for living things: Evidence from semantic dementia, HSVE and a neural network model. *Brain*, 130(4):1127–1137, April 2007.

75. Sheila Robinson, Judit Druks, John Hodges, and Peter Garrard. The treatment of object naming, definition, and object use in semantic dementia: The effectiveness of errorless learning. *Aphasiology*, 23(6):749–775, June 2009.
76. D Tranel, H Damasio, and A R Damasio. A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia*, 35(10):1319–1327, October 1997.
77. Gabriella Vigliocco, David P Vinson, Judit Druks, Horacio Barber, and Stefano F Cappa. Nouns and verbs in the brain: a review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neurosci. Biobehav. Rev.*, 35(3):407–426, January 2011.

Supplementary Note 1: Appendix (Detailed Model Implementation)

Microstructure. Each area consists of two neuronal layers — an excitatory layer and an inhibitory one – with 625 (25x25) cells (e-cells for the former and i-cells for the latter) in each layer. To avoid any potential edge effects, layers have a toroidal structure: the top edge is adjacent to the bottom one, and the left edge is adjacent to the right one. I-cells and e-cells have one-to-one correspondence. A combination of an e-cell and an i-cell reflects a population of pyramidal excitatory neurons and inhibitory interneurons of one cortical column (grey matter under approximately 0.25 mm2 of the cortical surface). Cells are modeled as graded-response neurons.
Each e-cell is restricted to send its projections to the 19x19 e-cell patch in the same area, 19x19 e-cell patches in connected areas, and 5x5 i-cell patch in the same area (Fig.2). These projections are created during the first initialization of a network in a random manner. The probability that the projection is created is derived from the Gaussian probability density function (Table 2). The highest probability is in the patch center (although, a cell cannot send projection to itself) and it is decreasing with the distance from the center.

E-cell dynamics. The activity state of a cell is uniquely defined by its membrane potential. The change in the e-cell membrane potential is dependent on its current membrane potential, a sum of its inputs, and uniform distribution of white noise:

$$\tau_e \frac{dV(e,t)}{dt} = -V(e,t) + k_1(V_{in}(e,t) + k_2\eta(e,t)) \tag{2}$$

where $V(e,t)$ is a membrane potential of the E-cell
$V_{in}(e,t)$ is a sum of inputs to this E-cell (see equation 3)
$\eta(e,t)$ is noise
τ_e is the time constant
$k1$ and $k2$ are the scaling constants
(Membrane potential of the e-cell)

Note that each e-cell has noise as its property which represents the spontaneous activity of real neurons. The sum of inputs to the e-cell is governed by the equation:

$$V_{in} = \sum E/IPSPs - k_G\omega_G(e,t) \tag{3}$$

where $\sum E/IPSPs$ is the sum of excitatory and inhibitory postsynaptic potentials sent to the cell of interest
$\omega_G(e,t)$ is global inhibition (see equation 4)
k_G is the scaling constant
(Sum of inputs to the e-cell)

Postsynaptic potential is defined as the output of the cell (see equations 5 and 10) multiplied by the synaptic weight between this cell and the target cell. Each e-cell gets exactly one IPSP from the corresponding i-cell, which is included in this sum with a negative sign. Note the global inhibition mechanism that is an area-specific inhibitory loop that prevents overall network activity from falling into nonphysiological states. This inhibition is calculated as follows:

$$\tau_G \frac{d\omega_G(e,t)}{dt} = -\omega_G(e,t) + \sum_{e \in area} O(e,t) \tag{4}$$

where $\sum_{e \in area} O(e,t)$ is the sum of e-cells' outputs from the whole area (see equation 4)
τ_G is the time constant
(Global inhibition of the E-cell)

An output of the e-cell is some value between zero and one. It depends on the cell's membrane potential and the threshold that this potential needs to exceed for an output to have a non-zero value:

$$O(e,t) = \begin{cases} 0, & \text{if } V(e,t) \leq \phi(e,t) \\ V(e,t) - \phi(e,t), & \text{if } (V(e,t) - \phi(e,t)) \in (0,1] \\ 1, & \text{otherwise} \end{cases} \tag{5}$$

where $O(e,t)$ is the output of the e-cell
$\phi(e,t)$ is the threshold of the membrane potential (see equation 6)
(Output of the e-cell)

The threshold of the membrane potential is not constant, it depends on the cell's recent activity – the more active the cell was, the higher is the threshold:

$$\phi(e,t) = \alpha\omega(e,t) \tag{6}$$

where $\omega(e,t)$ is averaged recent activity of an e-cell (see equation 7)
α is adaptation strength
(Threshold of the membrane potential)

When the network is initiated, the recent activity of every cell is zero ($\omega(e,t) = 0$). Then it is updated on each time step with an increase/decrease calculated as follows:

$$\tau_A \frac{d\omega(e,t)}{dt} = -\omega(e,t) + O(e,t) \tag{7}$$

where τ_A is the time constant
(Averaged recent activity of the e-cell)

I-cell dynamics. Change in the i-cell membrane potential is calculated similarly to those of the E-cell, but without the white noise component:

$$\tau_i \frac{dV(i,t)}{dt} = -V(i,t) + k_1 V_{in}(i,t) \tag{8}$$

where $V(i,t)$ is the membrane potential of the I-cell
$V_in(i,t)$ is the sum of inputs to this I-cell (see equation 9)
τ_i is the time constant
k_1 is the scaling constant
(Membrane potential of the i-cell)

There is no global inhibition mechanism for the i-cell as there is for the e-cell and its sum of inputs is just the sum of EPSPs it received:

$$V_{in} = \sum EPSPs \tag{9}$$

where $\sum EPSPs$ is the sum of EPSPs to the i-cell
(Sum of the inputs to the i-cell)

Note that the i-cell can receive only excitatory inputs from the 5x5 e-cells patch.
The output of the i-cell is simply its membrane potential if it is positive or zero otherwise:

$$O(i,t) = \begin{cases} 0, & \text{if } V(i,t) \leq 0 \\ V(i,t), & \text{otherwise} \end{cases} \tag{10}$$

where $O(i,t)$ is the output of the i-cell

Synaptic weight dynamics. Initially, random weights are assigned to all established connections with the weights of the links between two e-cells distributed uniformly and the weights of the links from e-cells to i-cells distributed normally, decreasing with the distance between the cells. Then, those weights are dynamically changing via the Hebbian learning mechanism:

$$\omega_{t+1}(x,y) = \begin{cases} \omega_t(x,y) + \Delta\omega, & \text{if } O(x,t) > \theta_{pre} \text{ and } V(y,t) > \theta_{post} \\ \omega_t(x,y) - \Delta\omega, & \text{if } O(x,t) \leq \theta_{pre} \text{ and } V(y,t) > \theta_{post} \\ \omega_t(x,y), & \text{otherwise} \end{cases} \tag{11}$$

where $\omega_t(x,y)$ is the weight of the synapse from the cell x to the cell y at time t
$\Delta\omega$ is the change in the synaptic weight
θ_{pre} is the presynaptic output activity threshold
θ_{post} is the postsynaptic membrane potential threshold

If both the output of the presynaptic cell and the membrane potential of the postsynaptic cell are high enough, synaptic weight is increased by constant number (long-term potentiation, LTP), which corresponds to the rule "fire together — wire together". If low presynaptic output is correlated with high postsynaptic potential, synaptic weight is weakened by the same amount (long-term depression, LTD), which corresponds to the rule "out of sync— out of link".

Table 1. Parameters of the model

$\tau_e = 2.5$	time constant for membrane potential of the e-cell (Eq.2)
$\tau_i = 5$	time constant for membrane potential of the i-cell (Eq.8)
$\tau_G = 12$	time constant for global inhibition (Eq.4)
$\tau_A = 10$	time constant for averaged recent activity of the e-cell (Eq.7)
$k_1 = 0.01$	scaling constant (Eq.2)
$k_2 = 100\sqrt{3}$	scaling constant (Eq.2)
$k_G = 95$	scaling constant (Eq.3)
$\alpha = 0.01$	adaptation strength (Eq.6)
$\eta \approx U[-0.5, 0.5]$	noise distribution (Eq.2)
$\Delta\omega = 0.0008$	change in synaptic weight during Hebbian learning (Eq.11)
$\theta_{pre} = 0.05$	threshold for presynaptic activity (Eq.11)
$\theta_{post} = 0.15$	threshold for postsynaptic activity (Eq.11)
$\gamma = 0.5$	fraction of the time-averaged output of the most active e-cell to determine the CA threshold (Eq.1)

Table 2. Gaussian distribution parameters. The probability that links between two cells will be created follows the Gaussian distribution, decreasing with the distance between cells.

$P_rec = 0.15$	central probability of the Gaussian distribution for recurrent links
$P_between = 0.28$	central probability of the Gaussian distribution for links between different areas
$\sigma_rec = 4.5$	standard deviation for the Gaussian distribution for recurrent links
$\sigma_between = 6.5$	standard deviation for the Gaussian distribution for links between different areas
$\omega_initial \approx U[0, 0.1]$	initial weight distribution for links between two e-cells
$\omega_inh \approx N[0.295, 0.2]$	weight distribution for links from the e-cell to a 5x5 patch of i-cells, decreasing with the distance between cells