Recruitment and Consolidation of Cell Assemblies for Words by Way of Hebbian Learning and Competition in a Multi-Layer Neural Network

Max Garagnani · Thomas Wennekers ·
Friedemann Pulvermüller

Abstract Current cognitive theories postulate either localist representations of knowledge or fully overlapping, distributed ones. We use a connectionist model that closely replicates known anatomical properties of the cerebral cortex and neurophysiological principles to show that Hebbian learning in a multi-layer neural network leads to memory traces (cell assemblies) that are both distributed and anatomically distinct. Taking the example of word learning based on action-perception correlation, we document mechanisms underlying the emergence of these assemblies, especially (i) the recruitment of neurons and consolidation of connections defining the kernel of the assembly along with (ii) the pruning of the cell assembly's halo (consisting of very weakly connected cells). We found that, whereas a learning rule mapping covariance led to significant overlap and merging of assemblies, a neurobiologically grounded synaptic plasticity rule with fixed LTP/LTD thresholds produced minimal overlap and prevented merging, exhibiting competitive learning behaviour. Our results are discussed in light of current theories of language and memory. As simulations with neurobiologically realistic neural networks demonstrate here spontaneous emergence of lexical representations that are both cortically dispersed and anatomically distinct, both localist and distributed cognitive accounts receive partial support.

Keywords Competitive recruitment learning ·
Language acquisition · Memory trace · Consolidation ·
LTP/LTD · Localist distributed representations ·
Neurocomputation · Neurobiologically plausible
modelling · Synaptic plasticity · Perisylvian cortex ·
Perception-action associations

Introduction

The field of cognitive neuroscience has been inundated by a flood of experimental findings produced by new and exciting imaging techniques. The development of brain-based theories, however, has not been able to keep up with the large and unexpected flow of experimental data: although the mapping of cognitive processes to cortical areas may be motivated by empirical evidence, a fundamental issue that any theory should address is a principled explanation of *why* specific areas become active when specific cognitive processes are being performed. The major "label and conquer" approach to cognitive neuroscience has, in many cases, fallen short of providing such a mechanistic explanation.

Recently, researchers have started to use computational modelling in conjunction with experimental techniques with a view to combine cognitive and brain theories and link neuronal circuits to functional systems, especially in the domain of visual, auditory and language processing (e.g. [13, 14, 16, 33, 37, 86, 93]). When built so as to closely replicate neuroanatomical structure and neurophysiological characteristics of the cortex, computational models can make precise, quantitative predictions about

M. Garagnani (✉) · F. Pulvermüller
Medical Research Council, Cognition and Brain Sciences Unit,
15, Chaucer Road, Cambridge CB2 7EF, UK
e-mail: Max.Garagnani@mrc-cbu.cam.ac.uk

F. Pulvermüller
e-mail: Friedemann.Pulvermuller@mrc-cbu.cam.ac.uk

T. Wennekers
Centre for Theoretical and Computational Neuroscience,
School of Psychology, University of Plymouth, Plymouth, UK

when and where in the brain specific cognitive processes are expected to take place. Such predictions can be tested with experimental methods, which can provide evidence in support of the neurophysiological validity of the models or lead to their further refinement. Crucially, neurobiologically realistic models can help address a fundamental (and generally neglected) question in the field of cognitive neuroscience i.e. the "how/why" question: by shifting the level of investigation from that of abstract mechanisms down to that of cortical circuits, models can provide a mechanistic explanation of how human cognition might emerge from neurobiological structure and function. In the work described here this approach was successfully applied in the domain of language to simulate and explain, at the neuronal level, the mechanisms underlying early word acquisition.

In psycholinguistics, most existing computational approaches explain language processes either as the activation and long-term storage of localist elements [17, 18, 45, 51, 57, 58] or on the basis of fully distributed activity patterns [28, 39, 52, 64, 72, 73, 76]. Localist approaches typically assume, a priori, the existence of separate nodes for separate items (words), and of pre-established, "hard-wired" connections between them; the adoption of anatomically distinct nodes allows different item representations to be active at the same time while avoiding cross-talk. Distributed accounts, on the other hand, do not make such a priori assumptions: according to them, the representations of the relevant items emerge as distributed patterns of strengthened connections over all nodes in a layer (the hidden layer). In this approach, the same set of hidden nodes is used to encode different items as different patterns of graded activation; this, however, makes it impossible to maintain separate different item representations when these are simultaneously active. In general, cognitive arguments (e.g. our proven ability to maintain multiple item representations distinct) favour localist representations, whereas neuroscience arguments weight in favour of distributedness [22, 58].

The results presented here suggest that these two approaches are both partly correct and partly misleading: distributed *and* anatomically distinct representations can emerge spontaneously in the cortex solely as a result of Hebbian synaptic plasticity mechanisms, and do not need to be assumed a priori.

Background

We start from the hypothesis that the neural correlate of a word is a memory circuit ("trace") that develops during early language acquisition [66]. It is well known that articulation is controlled by neuronal activity in inferior-frontal (IF) areas; the articulatory gestures lead to acoustic signals which, in turn, cause stimulation of neurons in superior-temporal (ST) auditory areas (see Fig. 1a). Thus, during babbling and in the earliest stage of word learning [25, 67] activity in IF cortex is accompanied by near-simultaneous activity in ST. The same applies in adults: whenever we utter a word, correlated activity is present in the areas controlling speech output, IF, and those where neurons respond to the incoming sound being produced, ST. In the brain, IF and ST areas are connected reciprocally (mainly via the arcuate and uncinate fascicles and the extreme capsule [11, 47, 61, 75]). Therefore, repeated speech-related co-activation of neurons in these areas in presence of associative (Hebbian) synaptic mechanisms [35] should lead to the formation of strongly interconnected sets of cells distributed over IF and ST cortex [8, 66] constituting *sensory-motor* associations between co-occurring cortical patterns of activity, such that, for example, listening to speech sounds involving specific articulators leads to the activation of the corresponding motor representations. In early ontogeny, spontaneous articulatory gestures are generated by genetically pre-programed mechanisms; after babbling, auditory input from the environment can activate the pre-established circuits, facilitating and leading to repetition and early world learning. A significant body of experimental evidence confirms the presence of such speech-motor associations between left superior-temporal and inferior-frontal cortex [23, 66, 68–70, 90, 91, 95, 96] and their role in language processing. Throughout this article we refer to such distributed networks of strongly interconnected neurons as to cell assemblies (CAs) [8, 34, 35, 59, 92].

In order to test the mechanistic validity of this account, we implemented a brain-inspired neural network that replicates the areas in the left hemisphere involved in spoken language processing (here, "language cortex" for short) in close proximity of the sylvian fissure, along with their approximate connections as inferred from experimental research [60, 63, 74], and investigated the emergence and consolidation of such perception-action circuits in it. The network (see Fig. 1b) was specifically designed to mimic neuroanatomical, connectivity and neurophysiological properties of the left perisylvian language cortex, as summarized below (the model is described in detail in [27]):

(i) On the basis of neuroanatomical and imaging studies, six interconnected cortical areas are modelled: primary auditory cortex (A1), auditory belt (AB) and parabelt (PB) areas (Wernicke's area), inferior prefrontal (PF) and premotor (PM) cortex (Broca's area) and primary motor cortex (M1).

(ii) Neurons are modelled as graded-response cells with adaptation, whose output represents the average firing rate of a local pool of pyramidal cells.

Fig. 1 The relevant areas of the left perisylvian cortex involved in spoken language processing, the overall network architecture, and the mapping between the two, indicated by the colour code. **a** The six different areas modelled, grouped into ST areas (labelled M1, PM, PF) and IF areas (labelled A1, AB, PB). Long-distance cortico-cortical links between PF and PB are not shown. **b** The six-areas network model and an illustration of the type of distributed circuit that developed during learning of perception-action patterns. Each small oval ("cell") represents an excitatory neuronal pool (column); solid and dotted lines indicate, strong reciprocal and weak (and/or non-reciprocal) connections, respectively. Co-activated cells are depicted as black or grey ovals. Only forward and backward links between co-activated cells are shown. Pools of inhibitory inter-neurons are not depicted (adapted from [27])
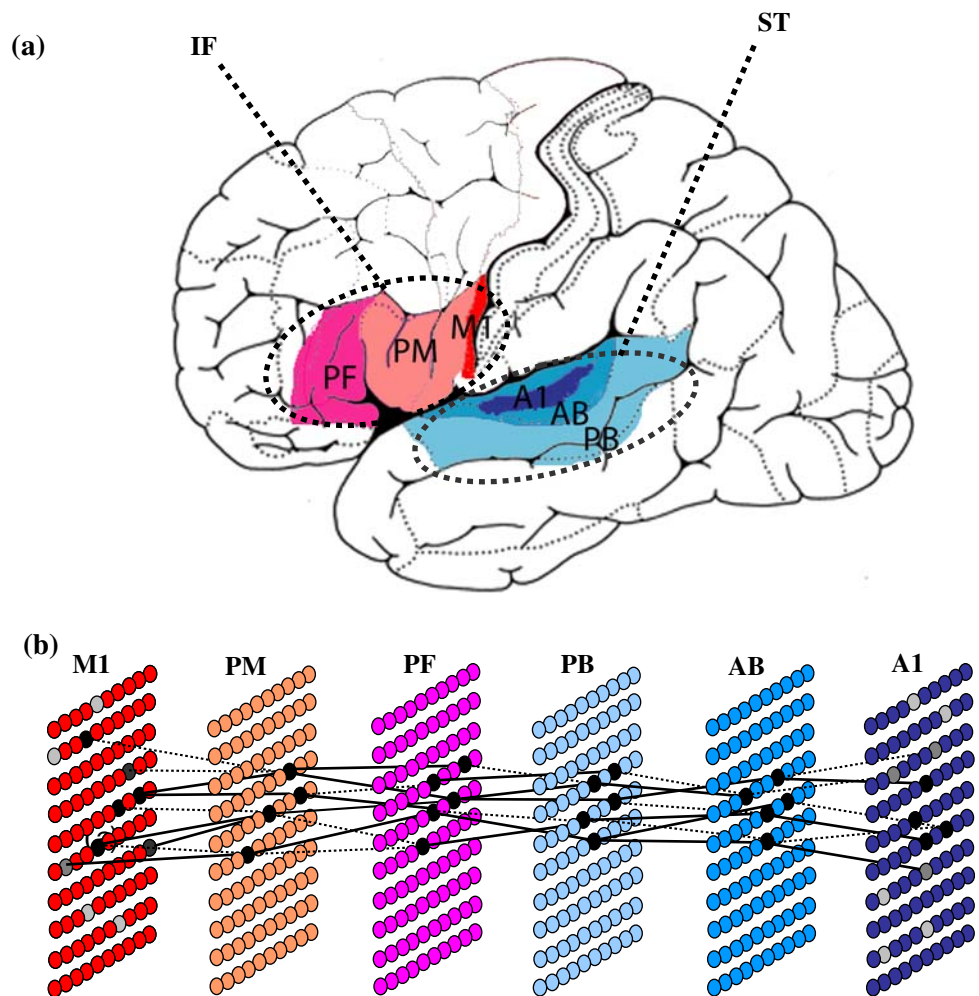
(iii) Within- (recurrent) and between-area connectivity is implemented via sparse, random, "patchy" next-neighbour synaptic links between cells, as typically found in the mammalian cortex [9, 29].

(iv) Both local and global (non-specific) cortical inhibition mechanisms are realized:

    a. Inhibitory cells reciprocally connected with neighbouring excitatory cells simulate the action of a pool of inter-neurons surrounding a cortical pyramidal cell in serving as lateral inhibition and local activity control;

    b. Area-specific inhibitory loops implement a mechanism of self-regulation, preventing the overall network activity from falling into non-physiological states (total saturation or inactivity).

(v) Long-term potentiation (LTP) and depression (LTD) [10, 49] cortical mechanisms of synaptic plasticity are modelled.

In our previous work, this architecture has been used to investigate the emergence of cell assemblies for words and the effects of attention on language processes [26, 27]; here, we specifically focus on the mechanisms of cell assembly formation and on how different computational implementations of synaptic plasticity affect such mechanisms and the network's ability to spontaneously develop separate input pattern representations.

It should be noted that we did not model individual spiking neurons but chose to use a mean-field approach, where each cell of the network represents the average activity of a local pool of neurons, or cortical column [21, 94]. Although spiking neurons would have made the model more biologically realistic, their introduction would have produced a significant impact in terms of computational resources; thus, we decided to start simple, and leave the implementation of this level of detail to a possible second phase, if necessary. As it turned out, modelling the cortical interactions at the level of cortical columns was sufficient to reproduce the phenomena of interest here.

With regard to point (v), we postulate that the emergence of specialized cell assemblies for words is driven by the repeated presentation of the same action-perception patterns in presence of Hebbian mechanisms of associative learning. LTP and LTD, consisting of long-term increase or decrease in synaptic strength resulting from pairing pre-synaptic activity with specific levels of postsynaptic membrane potentials, are believed to play a key role in experience-dependent plasticity, memory and learning [48, 71]. The network implemented two different computational abstractions of LTP and LTD, one based on Sejnowski's covariance rule [77, 78], the other one on the Artola–Bröcher–Singer (ABS) model of LTP/LTD [3, 4]. Both algorithms are described in the next section.

Methods

To induce CA formation in the model, we repeatedly exposed the network to pairs of (sparse) activation configurations, each activation-pattern pair representing the model equivalent of an auditory-articulatory word form (Refer Fig. 2). More precisely, two predetermined, randomly generated sets of cells were activated at the same time in the primary auditory (A1) and motor (M1) areas of the model, simulating speech production and correlated perception of the same speech element. The number of cells (17) activated in each primary area equalled 2.72% of the total number of cells of one area (625). The training consisted of repeated presentation (in randomized order) of four different pairs of patterns, in alternation with periods of variable length of time during which no input was given and activity was driven by white noise. The training terminated when each of the four stimuli had been presented to the network for five thousand times. Throughout the training (including the period in which no input patterns were present) the weights of all the links between cells were left free to adapt according to the specific learning rule used.
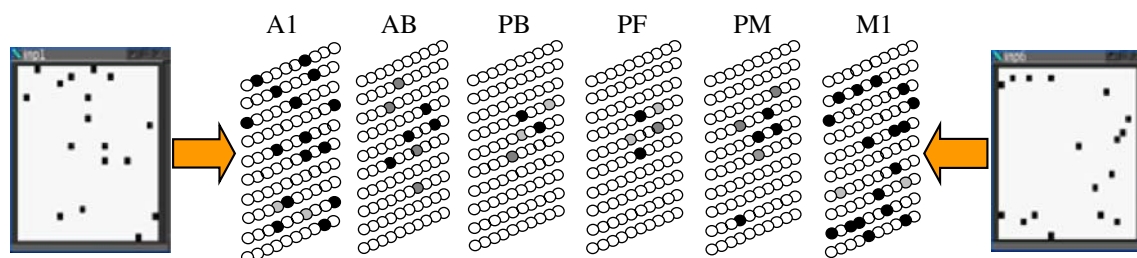
The first rule implemented, the co-variance rule [77], is a widely used [15, 46, 62, 93], neurobiologically inspired (e.g. [85, 87]; cf. [53] for a discussion) Hebbian rule. In it, the change of synaptic weight ω_{ij} of the excitatory link from pre-synaptic cell i to post-synaptic cell j per unit time is defined simply as:

$$\Delta\omega_{ij} = \alpha(x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \qquad (1)$$

where $\alpha \in \,]0,1]$ is a small constant specifying the learning rate, x_i is the current output of cell i, and $\langle x_i \rangle$ is the time-average output of cell i. While this rule captures well the essence of Hebbian learning (neurons that "fire-together, wire-together"), it was not originally built to accurately mimic known mechanisms of synaptic plasticity, like subsequent more realistic implementations have attempted to do (e.g. [7, 84]; see [6] for a useful account). The co-variance rule appears to be prone to the problem of CA merging [55], which, as discussed in detail in the section "Discussion", can be attributed to the imbalance between synaptic strengthening (LTP) and weakening (LTD) which is entailed by sparse neuronal activity. In the attempt to address this issue, we implemented and tested a second, more biologically accurate learning rule, based on the ABS model of LTP/LTD [3]. Such a model is derived from neurophysiological data suggesting that similar presynaptic activity (namely, brief activation of an excitatory pathway) can lead to synaptic LTD or LTP, depending on the level of postsynaptic depolarization co-occurring with the presynaptic activity. In particular, data from structures susceptible to both LTP and LTD [4, 20, 56] suggest that a stronger depolarization is required to induce LTP than to initiate LTD. Accordingly, the ABS model postulates the existence of two voltage-dependent thresholds in the postsynaptic cell, called θ_- and θ_+ (with $\theta_- < \theta_+$). The direction of change in synaptic efficacy depends on the membrane potential of the postsynaptic cell: if the potential reaches the first threshold (θ_-), all active synapses depress; if the second threshold (θ_+) is reached, all active synapses potentiate.

Fig. 2 Schematic illustration of network simulation of word learning processes: predefined stimulus patterns were presented simultaneously to areas A1 and M1, resulting in a temporary wave of activation that spread across the network. Black (grey) cells indicate strongly (weakly) activated cells. Between- and within-area synaptic links connecting cells are not depicted

We implemented a tractable version of the full ABS model [3]: the continuous range of possible synaptic efficacy changes was discretized into two possible step-changes, $+\Delta w$ and $-\Delta w$ (with $\Delta w \ll 1$ and fixed); also, we defined as "active" any link from a cell x such that the output $O(x,t)$ of cell x at time t is larger than θ_{pre}, where $\theta_{pre} \in]0,1]$ is an arbitrary threshold representing the minimum level of presynaptic activity required for LTP to occur.[1] Thus, given any two cells x and y currently linked with weight $w_t(x,y)$, the new weight $w_{t+1}(x,y)$ was calculated as follows:

$$w_{t+1}(x,y) = \begin{cases} w_t(x,y) + \Delta w & \text{if } O(x,t) \geq \theta_{pre} \text{ and } V(y,t) \geq \theta_+ \\ w_t(x,y) - \Delta w & \text{if } O(x,t) \geq \theta_{pre} \text{ and } \theta_- \leq V(y,t) < \theta_+ \\ w_t(x,y) - \Delta w & \text{if } O(x,t) < \theta_{pre} \text{ and } V(y,t) \geq \theta_+ \\ w_t(x,y) & \text{otherwise} \end{cases} \quad (2)$$

where $V(y,t)$ is the membrane potential of the postsynaptic cell y at time t, defined as the low-pass filtered sum of the total input to cell y (see Appendix 1). The total input to y represents all excitatory and inhibitory postsynaptic potentials (EPSPs, IPSPs) acting upon neuron pool y at time t (inhibitory inputs were given a negative sign). The three cases of Eq. 2 model, respectively, (i) homosynaptic and associative LTP, (ii) homosynaptic LTD and (iii) heterosynaptic LTD. The latter type of LTD involves synaptic change at inputs that are themselves inactive but that undergo depression due to depolarization spreading from adjacent active synapses [36]. It is important to note that post- and pre-synaptic thresholds θ_-, θ_+ and θ_{pre} are identical for all cells and remain unchanged throughout the simulation runs.

After the training, the network was tested to reveal the properties of the cell assemblies which had emerged for the given auditory-motor pattern pairs. More precisely, for each of the four patterns presented to the network, the time-average of the response (output value, or "firing rate") of each cell in the network was computed. These averages were used to identify the CAs that developed in the network in response to the four input pairs, as follows: a CA was defined simply as the subset of cells exhibiting average output above a given threshold $\gamma \in [0,1]$ during stimulus

presentation. Using the above functional definition, we then measured, for different values of γ, (i) CA size (averaged across the CAs) and (ii) distinctiveness of a CA, quantified as the average overlap (number of cells that two CAs shared) between one randomly chosen CA and the other three (this is also a measure of the amount of cross-talk between pairs of CAs). We repeated the above process and collected these measures for two sets of ten networks, each set trained using one of the two rules, and each network randomly initialized and trained with a different set of stimulus pairs.

Results

As the training progressed, we observed the emergence of *distributed* cell assemblies; the distinctiveness of the CAs, however, differed significantly depending on the learning rule adopted. This becomes apparent by examining the time-averaged response that each input pattern induced in the network at the different stages of the learning process.

Impacts of the Learning Rule

Let us first consider the results obtained using the covariance learning rule. Figure 3 shows the time-averaged response of one (randomly chosen) network to the four input patterns W1–W4 (one for each row), at different points during the training (after 10, 50, 100 stimulus presentations in panel (a), and after 1,000, 2,000, 3,000 in panel (b)), averaged over the time during which the patterns W1–W4 were presented in input. Each input pattern Wi consisted of a pair (Wi_A, Wi_M), representing, respectively, auditory and motor form of that word. Initially, the presentation of the pattern produces only weak activation in the two secondary areas AB and PM, and no activation in the central (or associative) areas PB and PF. As the learning progresses, however, the average response produced by the same stimulus reaches further towards the central areas, where the binding of the sensory-motor patterns is expected to take place. As the CAs become stronger in such areas, however, an interesting phenomenon was observed (Fig. 3b). At some point between 1,000 and 2,000 presentations, the two CAs specific to the input

---

[1] The output value $O(x,t)$ of a cell x represents the cumulative output (number of action potentials per time unit) of neuronal cluster (column) x at time t, and is a piecewise-linear sigmoid function of the cell's membrane potential $V(x,t)$. See Appendix 1 for details.

Fig. 3 a Time-averaged response of a network (trained using the covariance rule) to the four input patterns ($W_1,\ldots W_4$) at different stages of training: after 10 (top), 50 (middle) and 100 (bottom) stimulus presentations. The brightness of a cell indicates its average output value, or firing rate (black areas consist of cells having output ~ 0.0). **b** Time-averaged response of a network (trained using the covariance rule) to the four input patterns after 1,000 (top), 2,000 (middle) and 3,000 (bottom) stimulus presentations. Note the merging of CAs for W_1 and W_2
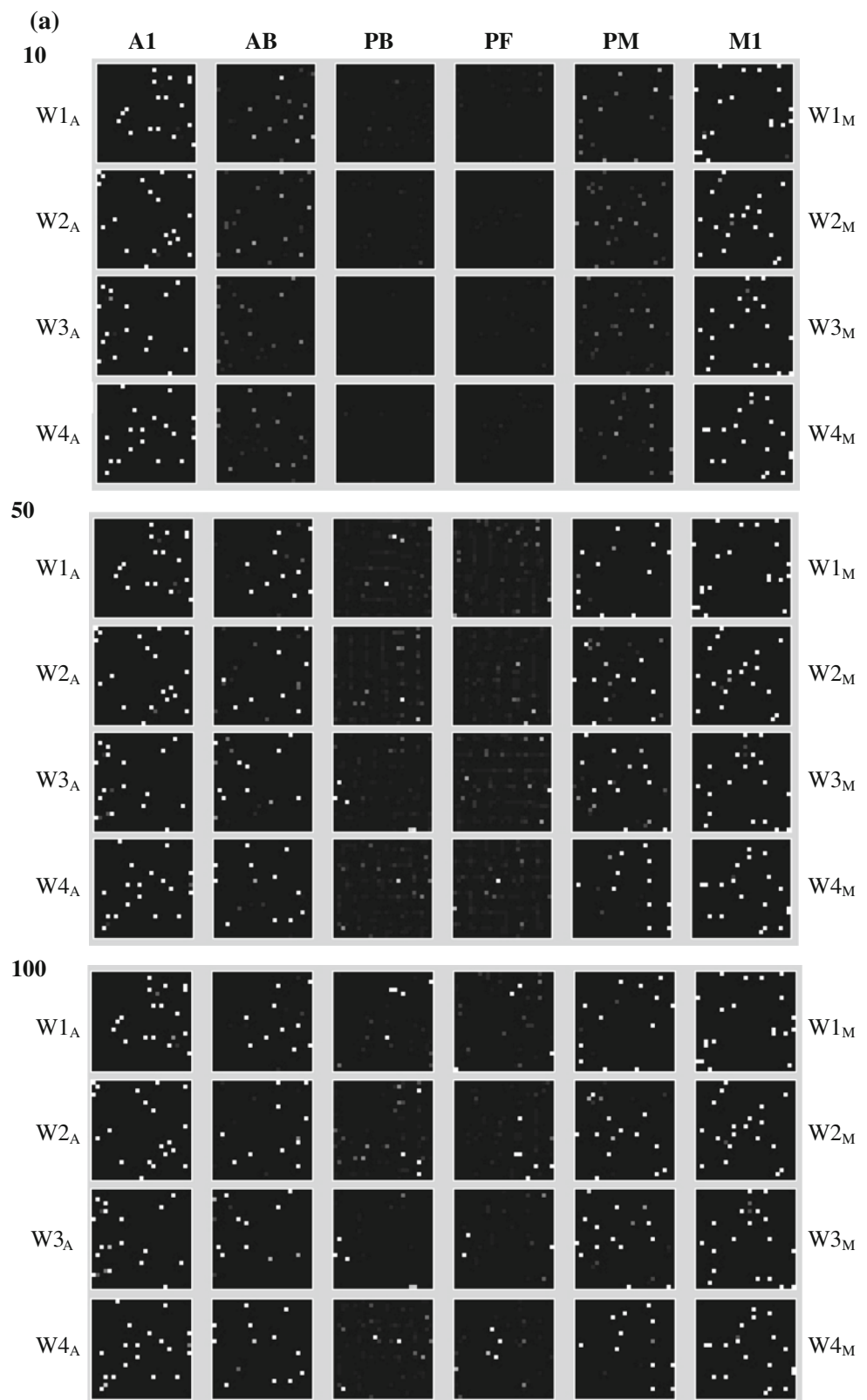
patterns W1–W2 "merge", becoming a single CA that responds to either of the two words: after 2,000 presentations, the responses to W1–W2 are almost identical across the four areas AB–PM. During the 1,000 stimulus presentations that follow, the CA development does not seem to progress any further.

Fig. 3 continued
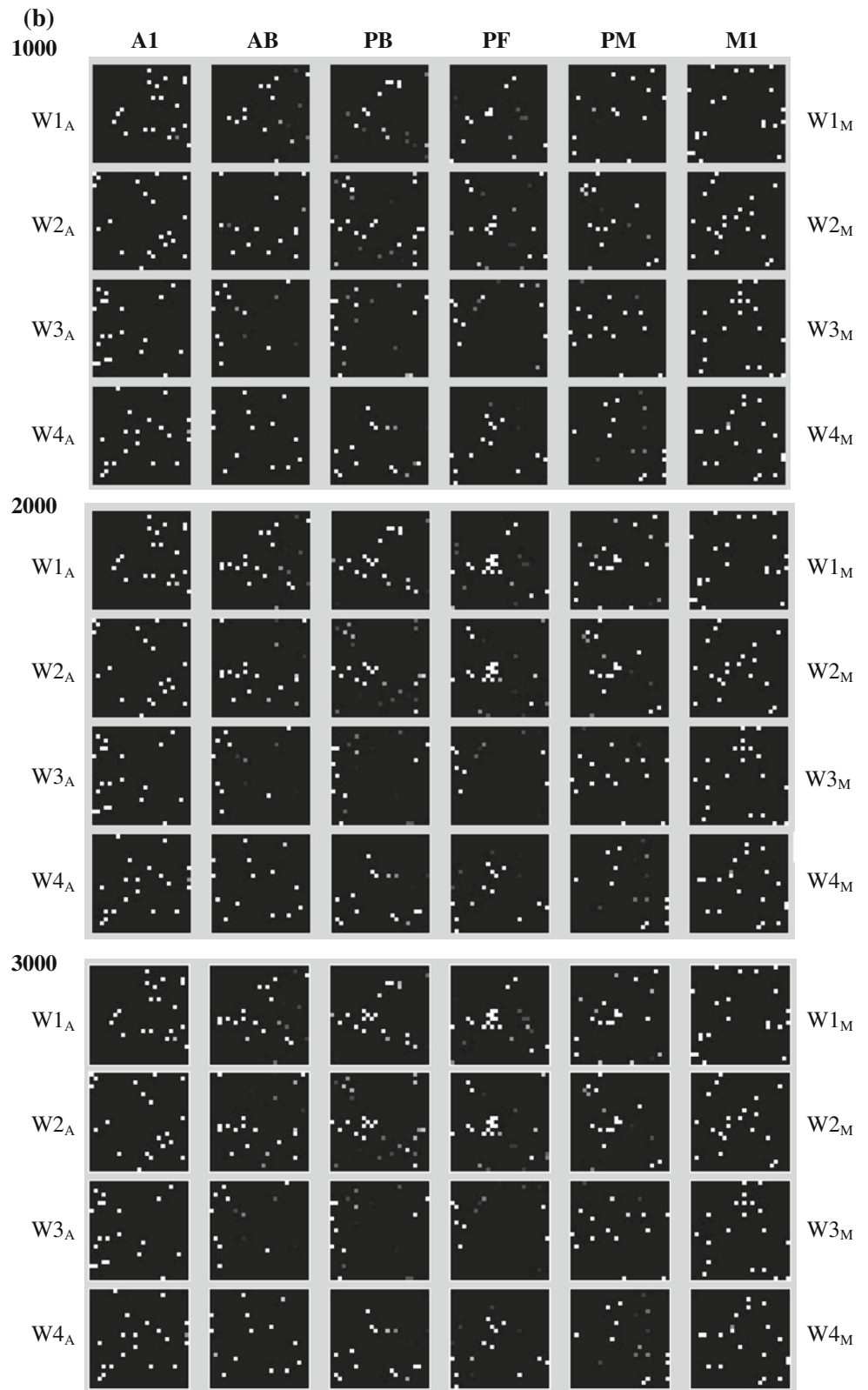
The merging phenomenon prevented the formation of distinct CAs in several of the networks, and was a symptom of the covariance learning rule's inability to separate, or "pull apart" the representations of two or more input patterns that happened to produce overlapping activations (see below and the "Discussion" section). This illustrates a

167
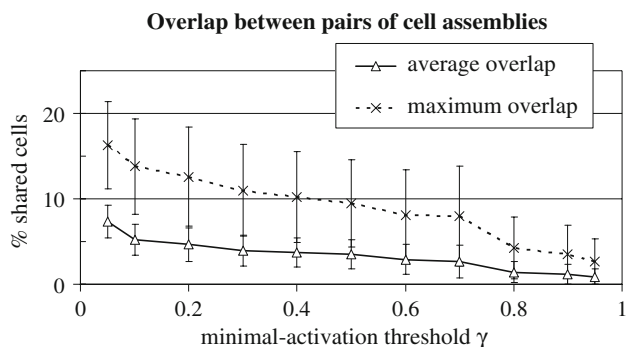
Fig. 4 Cell-assembly distinctiveness in networks trained using the covariance rule: average (SEM) overlap between pairs of CAs. Data are from the eight "best" networks (see text for details)

mechanism commonly used for nurturing reservations against the cell assembly theory: learning can "lump together" different representations (see, for example, [55]).

Figure 4 quantifies average cell assembly specificity in eight networks (the two networks showing the most extensive merging were discarded) as a function of the minimal-activation threshold γ, which was used for identifying the CAs (see "Methods" section). The graph shows the significant amount of overlap (or cross-talk) between pairs of CAs, expressed in % of shared cells between a randomly chosen CA and (i) the other three CAs (we plot the mean of the three overlaps) and (ii) the CA maximally overlapping with the chosen one.

Networks trained using the ABS rule produced significantly different results. Figure 5 plots CA distinctiveness for the 8 best networks. As opposed to the plot shown in Fig. 4, here the maximum overlap is above 5% only for values of $\gamma < 0.1$, and never above 10%. The average overlap is always below 5% and less than 2% for $\gamma > 0.2$.

Figure 6 shows an example of CA formation in one of the networks trained using the ABS rule. As before, activation is initially weak in the middle areas; however, after 100 presentations the CAs have already reached and
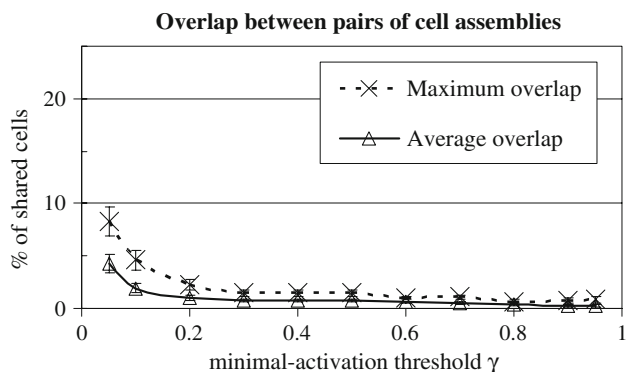
Fig. 5 Cell-assembly distinctiveness using the ABS learning rule: average (SEM) overlap between pairs of CAs as a function of the minimal-activation threshold γ (adapted from [27])

"filled" areas PB and PF, where the binding between sensory and motor patterns takes place.

The number of cells activated and involved in the binding is significantly larger than that observed in the previous simulations. In particular, a number of weakly active cells (widespread grey areas in the background) now accompany CA formation. Although a higher number of cells responding to an input pattern should increase the probability of CA overlap (and, thus, of CA merging), this did not happen (Fig. 5). In this example, too, a small overlap did develop between the CAs responding to W3 and W4 (e.g. compare the lower corners of area PB). This overlap, however, is maintained within limits (e.g. the responses to W3 and W4 still differ in areas AB and PM).

Although this may not be easily visible, the weakly active cells (which can be considered either as weak members of the CA or as part of what [8] called the "halo" of the assembly) become less numerous (and/or less active) between stages 100 and 5,000: this phenomenon is more apparent in the central areas (e.g. compare the responses to patterns W1–W3 in area PB at these two time points). On the other hand, cells that are already strongly active after 100 presentations (very bright or white dots) still respond equally (if not more) strongly after 5,000 presentations; this indicates that the CAs have reached a stable and robust configuration, with strongly and reciprocally connected sets of cells forming their "kernel" [8].

The reduction in the sizes of the CAs' haloes suggests that, subsequently to the initial period of CA growth, the links connecting a CA to the set of potential candidates (cells that could become part of the CA kernel i.e. be "recruited" by it) undergo a process of weakening, or "pruning". Such process could play a role in limiting CA merging and in "separating" initially overlapping CAs. To address this issue, it is necessary to look at the way in which CA overlap changes (in the areas where the pruning takes place) as a function of learning.

Pruning and CA Separation (with ABS Rule)

In this section we report evidence that the pruning and reduction in number of weakly active cells visually observed in the central areas is indeed a phenomenon that occurs reliably (on average) in all networks trained with the ABS rule, and that the amount of overlap between the CAs in these areas decreases as the learning progresses.

Figure 7 shows the average network responses to an input pattern at different stages of learning, expressed in number of cells responding to the input. All the cells in the network are grouped in different activation bins, according to the average output that they exhibit in response to a pattern (i.e. cells with output value between 0 and 0.01—1% of the maximal activation—are put in activation
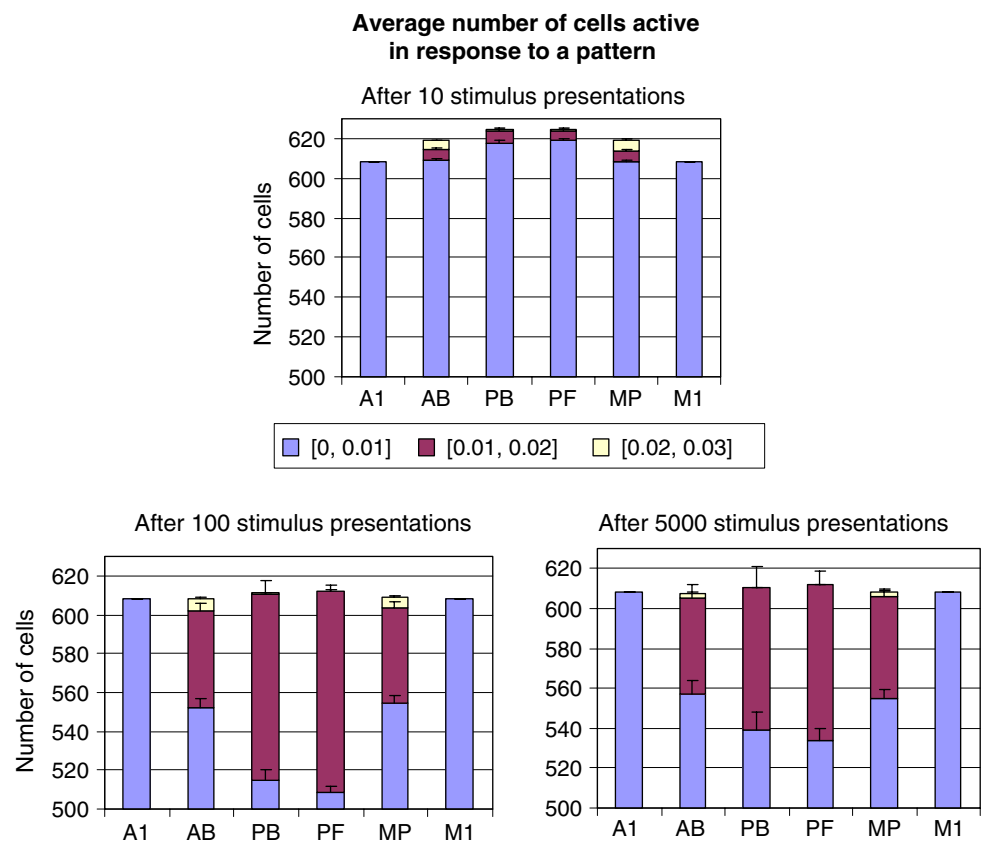
Fig. 6 Time-averaged response of a network trained using the ABS learning rule to the four input patterns after 10 (top), 100 (middle) and 5,000 (bottom) stimulus presentations. No global CA merging was observed
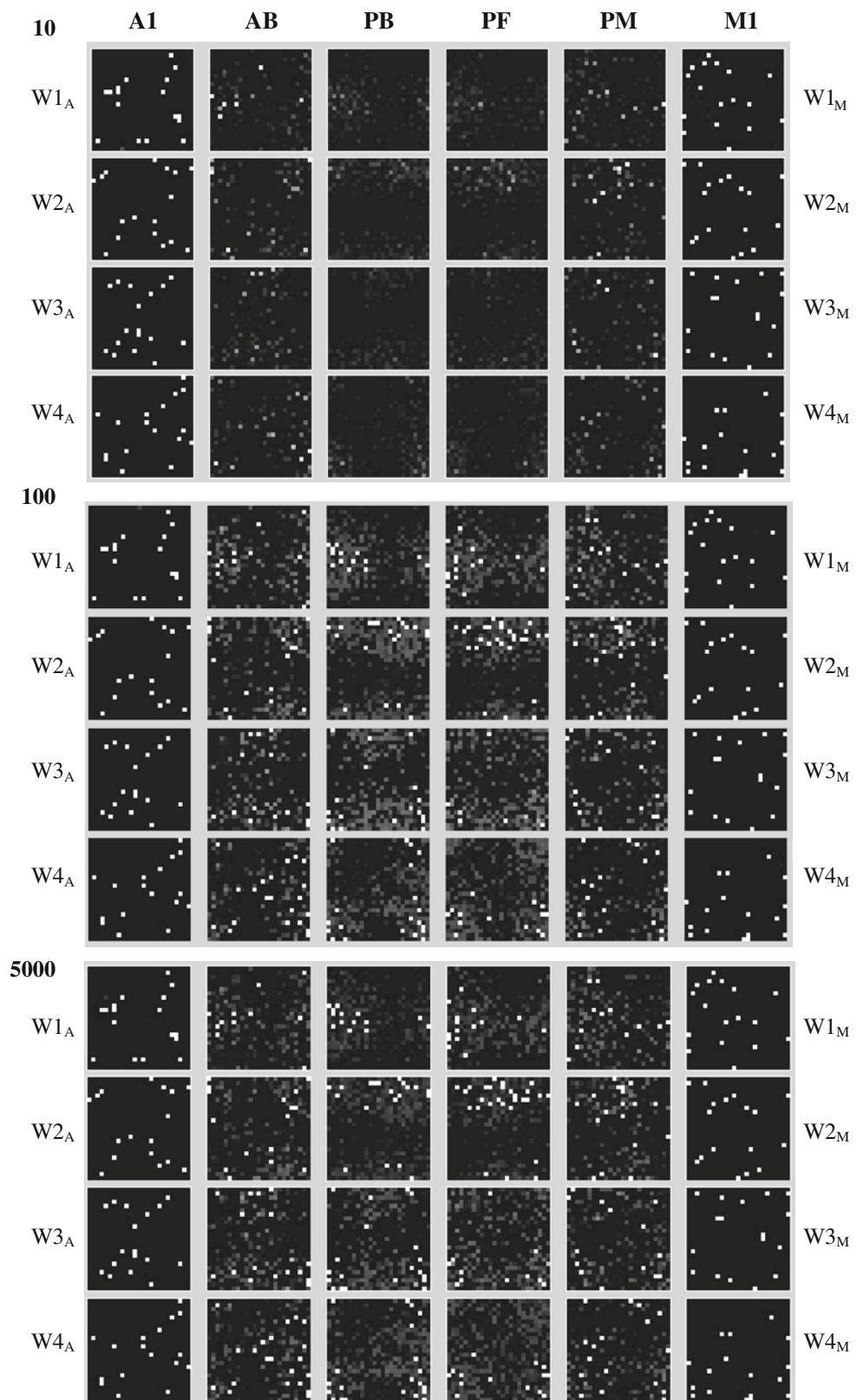
bin 0–0.01, etc.). Data for the activation bins 0–0.01, 0.01–0.02 and 0.02–0.03 are shown; note that more than 95% of the total number of cells in one area (625) fall within these bins.

By far the largest group (shown in cyan) consists of cells that are either not active or have very weak output (below 0.01). A direct comparison of the bar graphs for 100 and

Fig. 7 Changes in the CAs' haloes size: average (SEM) network response to an input pattern at different stages of training using the ABS rule (average of eight networks, each trained with four input patterns). The number of cells having average output value in the intervals 0–0.01, 0.01–0.02 and 0.02–0.03 are shown

Average number of cells active in response to a pattern

After 10 stimulus presentations

□ [0, 0.01]   ■ [0.01, 0.02]   □ [0.02, 0.03]

After 100 stimulus presentations

After 5000 stimulus presentations

5,000 training phases shows that, in the two central areas (PB and PF) a significant number of cells initially active in the 0.01–0.02 interval are pushed to the lower activation level (0–0.01). This result indicates that the weakly active cells (grey areas in Fig. 6) indeed become less numerous (and/or less active) between stages 100 and 5,000 in the two central areas. As the weaker activations can only be explained by synaptic-weight reduction, these numbers illustrate the phenomenon of pruning during the learning process.

Finally, the graphs in Fig. 8 demonstrate that the pruning leads to the gradual separation ("pulling apart") of the CAs: in addition to the expected decrease in CA size (panel (a)), both the maximum (panel (b)) and average (panel (c)) overlap between CAs decrease significantly (in areas PF and PB) as the learning progresses.

Emergence (Cell Recruitment) and Consolidation of CAs

The data presented in Figs. 7 and 8 only concern groups of weakly active cells, forming the CA's halo; what happens to the more strongly active ones, constituting the CA's kernel? Figure 9 shows the emergence and development of both the weakly and strongly active groups of cells. The

bar graphs plot the number of active cells per activation interval (bins 0.01–0.02 to 0.09–0.1 and 0.1–0.2 to 0.9–1.0) per area, in response to an input pattern, at three points during the training. After 10 presentations, strongly active cells are only present in the primary areas A1 and M1, where the input patterns are presented; the central areas contain a fair number of weakly active cells. After 100 presentations, three things can be observed: (i) the number of cells active between 0.01 and 0.02 exceeds the graph ceiling (as can be seen from Fig. 7, the actual numbers vary in the range 50–100); (ii) the slightly more active cells (still part of the CA's halo) have decreased in number and (iii) the number of strongly active cells (from 0.5–0.6 to 0.9–1.0) has significantly grown. As we know from the data in Fig. 7 that the numbers for the 0–0.01 bin decreased and that the total number of cells with activity below 0–0.03 was either unchanged or reduced across areas, we can conclude that the fate of the weakly active cells was to be "recruited" by the CA and to become strongly active i.e. part of the CAs' kernel. After 5,000 stimulus presentations, the same cells have shifted their activation even closer to the maximum: the training has produced further strengthening of the synaptic links between the cells recruited at the beginning, which now form a stable and robust CA.

(a)

(b)

(c)
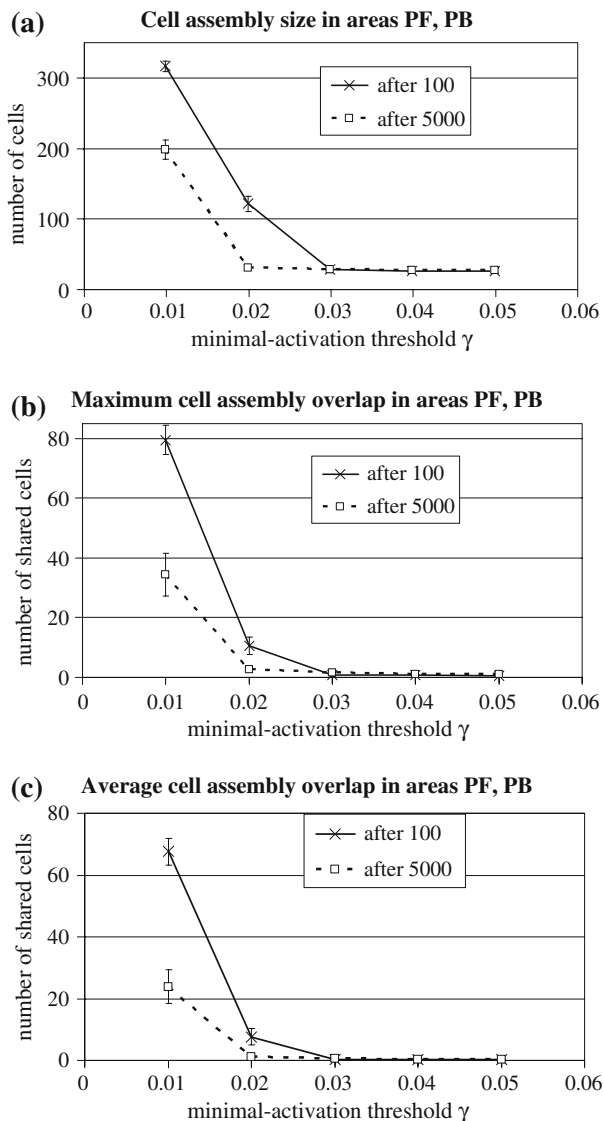
Fig. 8 Pruning of CAs' haloes: average (SEM) cell assembly size (**a**), maximum (**b**) and average overlap (**c**) between one CA and the other three in the central areas (PF and PB), for two different training points. Note the significant drop in the size and overlap between 100 and 5,000 stimulus presentations

Discussion

We report here four observations apparent from word learning simulations in a brain-inspired neural architecture:

1. Action-perception patterns are stored as distributed sets of neurons in multilayer networks with neuroanatomical constraints;

2. During learning, neurons are recruited and gradually bound together into a single CA; such recruitment and CA consolidation processes proceed from the sensory and motor areas inwards, towards the central, or "amodal", associative areas;

3. Whereas networks adopting the covariance rule [77] struggle to produce input-specific, distinct lexical representations, the adoption of a neurobiologically grounded Hebbian rule with fixed thresholds, based on the Artola–Bröcher–Singer [3] model of LTP/LTD, leads to CA overlap minimization (<5%) and anatomically distinct CAs;

4. In the networks adopting the ABS rule, the process of growth and merging of CAs is countered by a process of competition and pruning of the CAs' halo; we conjecture that this reduction in CA size and overlap reflects a strong weight decrease in the connections between kernel and halo.

It should be added that the strong binding of the articulatory-acoustic activation patterns and distinctiveness of CAs was confirmed by additional simulations (not reported here) in which the network was stimulated, in area A1 only, with the auditory component of a word pattern; the results showed that, while the input-specific CA was strongly activated and led to partial reconstruction of the associated motor pattern in M1 (on average, approximately 30% of the motor pattern was reproduced), the other CAs remained almost completely silent (see Garagnani et al. [27], their Figs. 10 and 11).

The simulation results indicate that the formation of cell assemblies (or, more generally, of memory traces) begins with an initial period of CA expansion, during which neurons in progressively more central areas are recruited and bound strongly into the CA's kernel. As soon as the CAs' haloes start to overlap (this happens already after 100 stimulus presentations), competition for the recruitment of cells in such overlaps begins, gradually leading to the survival of only the strongest connections and pruning of the weakest ones. This reduction effectively means that the representations of the input patterns are being separated, a phenomenon typically observed in networks that implement competitive learning mechanisms [31, 32, 42, 43]. At the same time, the neurons initially recruited are further consolidated in the CA, preventing, to some degree, their re-use by different CAs. By "competitive learning" here we mean competition between the incoming patterns, rather than just between synapses; this behaviour is often considered a hallmark of many forms of developmental plasticity [10, 40, 82]. Notice that the presence of synaptic competition in a learning rule (implemented, e.g. via heterosynaptic LTD or weight normalization) does not, in itself, guarantee competition between the incoming patterns. Indeed, the covariance rule tested here [77, 78] implements both LTP (Table 1, case (a)) and heterosynaptic LTD (Table 1, case (b)), but cannot achieve competitive learning [54].

To understand why this is so, and what may be underlying the ABS rule's competitive behaviour, consider the

Fig. 9 Emergence and consolidation of CAs: number of cells active in response to an input pattern (averaged across eight networks), grouped by activation level (bins) and area, at three different points of training (using the ABS rule). Data for the activation bin 0–0.01 are not shown (but see Fig. 7). Cells from the CA's halo are recruited during the initial phase and become part of the kernel; as learning progresses, the strongly active cells in all areas become more firmly bound into the CA
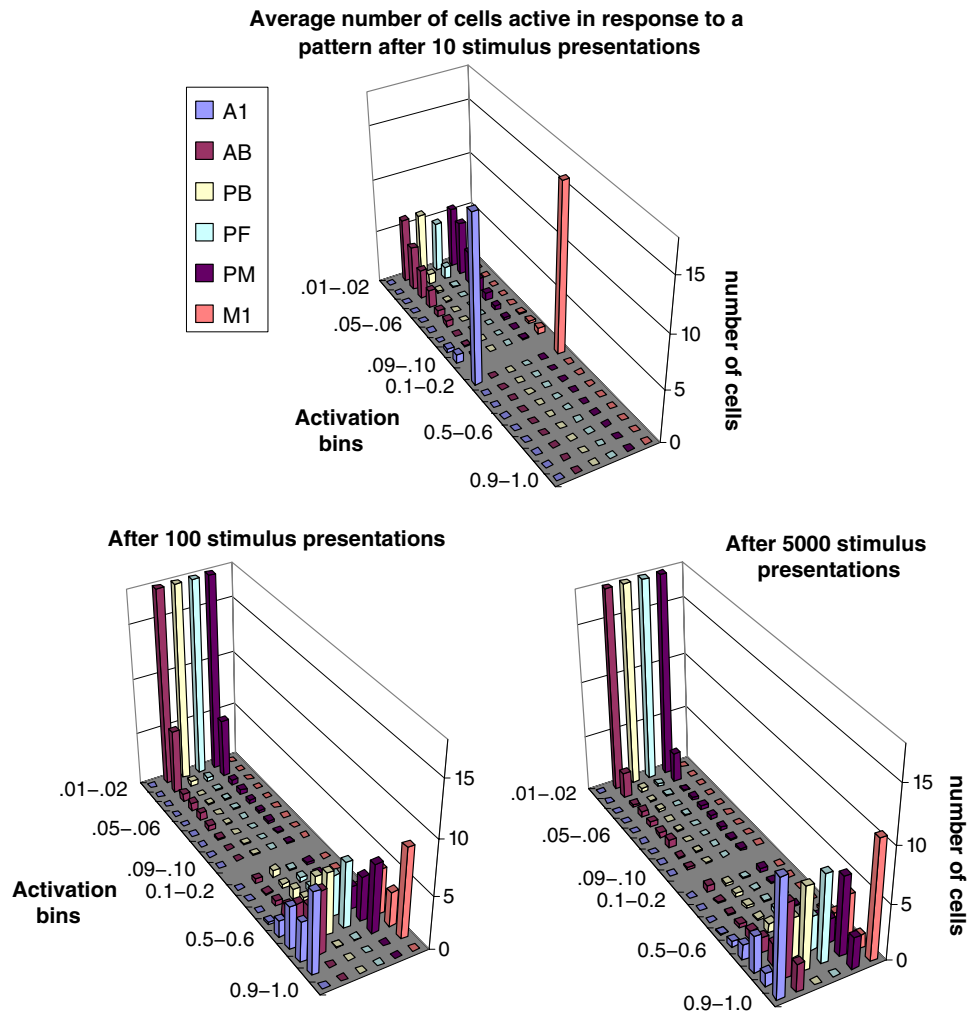
2-area network of cells depicted in Fig. 10 below. Let us assume that the network uses sparse coding, and that the cells in area 1 are repeatedly confronted with different patterns of activation: two input patterns (called A and B) strongly activate cells A_1, A_2, C_1, C_2 and B_1, B_2, C_2, C_3, respectively.
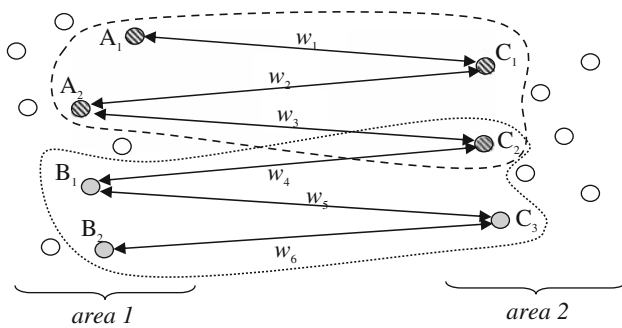
Fig. 10 Example of overlapping cell assemblies. Nodes simultaneously active are depicted using the same fill pattern. The dashed and dotted lines identify the two CAs activated by two different input patterns

Assume that during training, the weights are modified according to the co-variance rule, as summarized by Table 1. The difference between current and average activity of one cell is larger when cells are fully active than when they are silent, as in a sparsely active network, a

Table 1 Summary of [77] covariance rule

Pre-synaptic cell		Post-synaptic cell		Δw= pre x post	
Active	↑	Active	↑	↑	(*a*)
Silent	↓	Active	↑	↓	(*b*)
Active	↑	Silent	↓	↓	(*c*)
Silent	↓	Silent	↓	↑	(*d*)

The size of the arrows in columns 2 and 4 indicates the approximate magnitude of the difference between current and average activity of that cell (assuming sparse neuronal coding); the orientation indicates the sign of the difference (up: positive, i.e. the cell is active above average; down: negative). Arrows in the last column indicate the amount of change in weight (Δw) that the rule produces in a link when the pre- and post-synaptic cells exhibit the specified activations (obtained as the "product" of columns 2 and 4)

cell's average activity is much closer to zero than to its maximum level of activation.

Note that links between two cells that are simultaneously silent are strengthened (case (*d*) in Table 1). This leads to an overall merging effect. Setting $\Delta w = 0$ in case (*d*) is not sufficient to solve this problem. In fact, because of the differences in magnitude, the net effect produced by the alternated strengthening (*a*) and weakening (cases (*b*) or (*c*)) of a link is an *increase* in strength. In the example of Fig. 10, alternation of inputs A and B means alternated increase (homosynaptic LTP, (*a*)) and decrease (heterosynaptic LTD, (*b*)) of w_3 and w_4: the net effect is a weight increase in both, which, in the long run, causes the two cell assemblies to merge into a single one.

The Artola–Bröcher–Singer rule differs from the covariance rule at least in the following ways:

- It uses the same amount of weight change Δw per unit time for both LTP and LTD (this implies that weakening and strengthening produce weight changes of equal magnitude);
- It does not strengthen links between cells that are simultaneously silent;
- It uses a single parameter's value (the postsynaptic membrane potential) to determine whether LTP or LTD should occur—see Eq. 2.

The last feature (based on neurobiological evidence) allows one to precisely define the ranges of values of the postsynaptic membrane potential for which either LTP or LTD will occur. We speculate that the ratio between the widths of these ranges allows the modulation of the total amount of competitive learning that takes place in the network. The validity of this hypothesis requires further computational testing, motivated also by the presence of a residual CA overlap in the ABS rule simulations.

Also note that as words usually have "neighbours" with which they share part of their form [12, 50], the model correlate of words should also partially overlap in their articulatory-acoustic signatures. Input patterns with overlapping activations, however, are expected to lead to an increase in CA overlap and merging; to maintain CA distinctiveness and neurobiological realism, we envisage the use, in future, of temporally dynamic patterns, spiking neurons and time-dependent synaptic plasticity rules.

The fact that the computational abstraction of the ABS model implemented here exhibits both competitive and recruitment learning behaviour [19, 24, 80] is worth of note. Perhaps the most well-known example of Hebbian learning rule exhibiting both of these properties is the Bienenstock–Cooper–Munro (BCM) rule [7], which has been successfully used to model and explain the spontaneous emergence of orientation selectivity and ocular dominance in the visual cortex [83]. It should be noted that,

although many of the BCM rule properties have been shown to arise from spike-time dependent plasticity rules (e.g. [38]), this rule had been originally developed to account for cortical organization and receptive field properties during *development*. Instead, the ABS model was derived from neurophysiological data obtained in the mature cortex. Below we discuss in detail additional aspects that distinguish the ABS rule implemented here from the classical BCM rule.

First of all, in the BCM rule the LTP/LTD threshold—corresponding to parameter θ_+ in Eq. 2—is not, like here, a predefined, fixed value, but a sliding threshold that changes according to the running average of the postsynaptic cell's activity.[2] As pointed out by [53], although evidence suggesting that the LTP/LTD threshold may be affected by the activity of the cell does exist (e.g. [5, 41]), it has been established that this effect is input (i.e. synapse) specific, and that it depends on the pattern of *pre*-synaptic rather than *post*-synaptic activity [2]. Thus, the assumption of a *single*, postsynaptic-driven LTP/LTD threshold that applies to *all* the synapses of a cell is not entirely justified.[3] Second, in the BCM rule LTD occurs even with very small postsynaptic potentials, whereas experimental evidence suggests that if postsynaptic depolarization remains below a certain threshold, the synaptic efficacy should remain unchanged, regardless of any presynaptic activity [4]. This aspect was implemented in the present ABS rule using a second (fixed) threshold, parameter θ_- in Eq. 2. Finally, the BCM rule is unable to model heterosynaptic LTD (the weakening of synaptic inputs that are themselves inactive), as it requires at least *some* presynaptic activity to be present at a synapse for LTD to take place. This form of LTD has been observed in the hippocampus and neocortex [36]; the induction protocols require strong postsynaptic activation (e.g. high frequency stimulation of the cell through excitatory inputs). Accordingly, the ABS rule implemented here allows heterosynaptic LTD to occur, subject to the postsynaptic cell being strongly depolarized (condition $V(y,t) \geq \theta_+$ in Eq. 2).

In view of the above point, we submit that the ABS rule that we adopted is more neurobiologically accurate than the BCM rule; furthermore, it does not make any assumptions

---

[2] More precisely, for the BCM rule to exhibit stable learning behaviour, the threshold must be a more-than-linear function of the cell's average output rate (simulations typically use the power of 2).

[3] Although evidence in support of the existence of homeostatic plasticity mechanisms exists (see [88] for a review), phenomena such as that of synaptic scaling—showing that prolonged changes in the cell's activity lead to the multiplicative scaling of all the amplitudes of the miniature excitatory postsynaptic currents [89]—do not constitute direct evidence for the presence of a single sliding LTP/LTD cell-threshold. Equally, synaptic scaling does not justify assuming that the norm of the vector of the synaptic strengths is conserved and equal for all cells, as often presupposed by neurobiologically inspired implementations of Hebbian learning (e.g. [44]).

about the existence of a global sliding threshold (or conservation of the cell's total synaptic strength). At the time of writing, we are not aware of any other examples of biologically realistic learning rules with *fixed* (non-sliding), input-specific (local) LTP/LTD thresholds (and no synaptic-weight conservation) that have been reported to exhibit both recruitment and competitive learning (but see [81] for a biologically grounded recruitment learning algorithm).

Relevant here is the innovative recent study by [79], which explores the effects of the BCM learning rule on the dynamics of neural populations in a model of the hippocampus and surrounding cortical areas. Like the simulations presented here, Seth and Edelman's network uses mean-firing rate units and Hebbian synaptic plasticity, and carefully replicates neuroanatomical structure and connectivity of the relevant areas. In this study, however, the authors investigate the spontaneous emergence of mesoscale (i.e. cell assembly level) causal structures during the execution of a spatial navigation task. To analyse the network's behaviour in terms of population dynamics, for each specific network output, or neural reference (NR),[4] a corresponding "context network" is identified, consisting of the set of neuronal interactions that could have potentially caused the observed NR (i.e. the set of cells that were active before the NR and are connected to it, either directly or via a number of synaptic links smaller than six). From this context, a "causal core" of interactions is then extracted, comprised of only those links that are *causally significant* (according to Granger's concept of causality [30]) and that form a chain of activations causing (or *predicting*) the specific NR. Interestingly, the results of this study indicate that the size of the causal cores diminishes as the learning progresses, a process which the authors refer to as of "refinement", and which is interpreted by them as possibly reflecting the selection of a specific causal pathway from diverse neuronal repertoires.

In view of the apparent similarities with the present work, one might be tempted to draw parallels between the concepts of causal core and cell assembly, or, even further, between the process of refinement and that of pruning of a CA's halo. However, this analogy would not be very appropriate. To begin with, the concept of causal core (or context network) introduced by Seth and Edelman is not equivalent to that of cell assembly kernel (or halo). In fact, a CA kernel consists of a set of strongly and reciprocally connected cells; the presence of positive feedback loops within the CA's circuits is a crucial feature, as it allows

reverberation and persistence of activity even in absence of a stimulus. Instead, causal cores, more akin to Abeles' synfire chains [1], are formed by mono-directional chains of (not necessarily strongly connected) cells, whose sequential activation is a good predictor of the activation of a *single* output cell at a *particular* time point (due to the combinatorial growth in the number of cells to be included in the context network, the causal core analysis cannot be easily extended to include a *set* of NRs rather than just one [79]). Second, while in the results reported here the pruning process that gradually separates the CAs takes place in the CA's halo, in Seth & Edelman's simulations the refinement takes place in the causal core (the equivalent of the CA kernel), and not in the context network, or "periphery" of the circuit. Third, Seth and Edelman's results do not provide evidence of an ongoing process of recruitment of cells from the periphery (or halo) and consequent consolidation of the links between them: the context network is mostly unaffected by the learning process [79], their Fig. 3a). Finally, to achieve competitive learning, in addition to adopting a BCM-based rule (which, as discussed earlier, makes some assumptions that are not biologically motivated), Seth & Edelman's model of synaptic plasticity uses weight normalization [44], a mechanisms not fully justified on neurobiological grounds. In spite of these differences, Seth & Edelman's significant contribution is still close, in spirit, to the present and other works (e.g. [16, 33, 37, 86]), which attempt to explain the emergence of high-level behavioural and cognitive processes in terms of neural population dynamics in neurobiologically realistic neural networks.

Finally, the present results provide evidence in support of the hypothesis that words, similar to other units of cognitive processing (e.g. objects, faces), are represented in the human brain as distributed *and* anatomically distinct action-perception circuits. Existing theoretical and computational accounts of knowledge representation in the brain explain memory either as the activation of a priori-established localist elements, or on the basis of fully overlapping, distributed patterns (see "Introduction" section). However, neither of these accounts is entirely compatible with an approach grounded in neuroanatomy and neurophysiology: localist networks with one cell (neuronal pool, or cortical column) coding for one cognitive trace may have difficulty in explaining (or making predictions about) the experimentally observed spreading of activity in cortex when words or concepts are recognized. Fully distributed networks, on the other hand, predict very global brain activity if their layers are not firmly related to specific cortical areas, and struggle to explain our ability to maintain active more than one representation at the same time within the same sensory modality. The present results suggest that anatomically

---

[4] A "neural reference" is defined as the activation, at a particular time, of a particular cell in area CA1; in the model in question, the activity in area CA1 directly affected motor output.

distinct *and* distributed action-perception circuits can emerge spontaneously in the cortex as a result of synaptic plasticity. Our model predicts and explains the formation of lexical representations consisting of strongly interconnected, anatomically distinct cortical circuits distributed across multiple cortical areas, allowing two or more lexical items to be active at the same time. Crucially, our simulations provide a principled, mechanistic explanation of *where* and *why* such representations should emerge in the brain, making predictions about the spreading of activity in large neuronal assemblies distributed over precisely defined areas, thus paving the way for an investigation of the physiology of language and memory guided by neurocomputational and brain theory.

Appendix 1

The membrane potential $V(x,t)$ at time t of a model cell x with membrane time-constant τ is governed by the equation:

$$\tau \cdot \frac{dV(x,t)}{dt} = -V(x,t) + V_{In}(x,t) \qquad (3)$$

where $V_{In}(x,t)$ is the total input to cell x. The output $O(x,t)$ of an excitatory cell x at time t is defined as:

$$O(x,t) = \begin{cases} 0 & \text{if } V(x,t) \leq \varphi \\ (V(x,t) - \varphi) & \text{if } 0 < (V(x,t) - \varphi) \leq 1 \\ 1 & \text{otherwise} \end{cases}$$

$$(4)$$

The cell adaptation φ is initialized to 0 and varies in time according to Eq. (5):

$$\tau_A \cdot \frac{d\varphi(x,t)}{dt} = -\varphi(x,t) + O(x,t) \qquad (5)$$

The output $O(x,t)$ of an inhibitory cell x at time t is 0 if $V(x,t) < 0$, and $V(x,t)$ otherwise.

Parameter values used in the simulations were:

(Eq. 3) Excitatory cells: $\tau = 2.5$; Inhibitory cells: $\tau = 5$ (in simulation time-steps)

(Eq. 5) $\tau_A = 15$

(Eq. 1) $\alpha = 0.004$

(Eq. 2) $\theta_- = 0.15$; $\theta_+ = 0.25$; $\theta_{pre} = 0.05$; $\Delta w = 0.0005$

The low-pass dynamics of the network cells (Eqs. 3 and 5) were integrated using the Euler scheme with step size Δt [65], with $\Delta t = 0.5$ (in arbitrary units of time).

References

1. Abeles M. Corticonics—neural circuits of the cerebral cortex. Cambridge: Cambridge University Press; 1991.
2. Abraham WC, Bear MF. Metaplasticity: the plasticity of synaptic plasticity. Trends Neurosci. 1996;19(4):126–30.
3. Artola A, Singer W. Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. Trends Neurosci. 1993;16:480–7.
4. Artola A, Bröcher S, Singer W. Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. Nature. 1990;347:69–72.
5. Bear MF. Mechanism for a sliding synaptic modification threshold. Neuron. 1995;15(1):1–4.
6. Bi GQ, Poo MM. Synaptic modification by correlated activity: Hebb's postulate revisited. Annu Rev Neurosci. 2001;24:139–66.
7. Bienenstock EL, Cooper LN, Munro PW. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. J Neurosci. 1982;2:32–48.
8. Braitenberg V. Cell assemblies in the cerebral cortex. In: Heim R, Palm G, editors. Theoretical approaches to complex systems, vol. 21. Berlin: Springer; 1978. p. 171–88.
9. Braitenberg V, Schüz A. Cortex: statistics and geometry of neuronal connectivity. 2nd ed. Berlin: Springer; 1998.
10. Buonomano DV, Merzenich MM. Cortical plasticity: from synapses to maps. Annu Rev Neurosci. 1998;21:149–86.
11. Catani M, Jones DK, Ffytche DH. Perisylvian language networks of the human brain. Ann Neurol. 2005;57(1):8–16.
12. Coltheart M, Davelaar E, Jonasson J, Besner D. Access to the internal lexicon. In: Dornic S, editor. Attention and performance VI. London: Academic Press; 1977. p. 535–55.
13. Corchs S, Deco G. Large-scale neural model for visual attention: integration of experimental single-cell and fMRI data. Cereb Cortex. 2002;12(4):339–48.
14. David O, Friston KJ. A neural mass model for MEG/EEG: coupling and neuronal dynamics. Neuroimage. 2003;20(3):1743–55.
15. Dayan P, Sejnowski TJ. The variance of covariance rules for associative matrix memories and reinforcement learning. Neural Comput. 1993;5(2):205–9.
16. Deco G, Rolls ET, Horwitz B. "What'' and "where'' in visual working memory: a computational neurodynamical perspective for integrating fMRI and single-neuron data. J Cogn Neurosci. 2004;16(4):683–701.
17. Dell GS. A spreading-activation theory of retrieval in sentence production. Psychol Rev. 1986;93(3):283–321.
18. Dell GS, Chang F, Griffin ZM. Connectionist models of language production: lexical access and grammatical encoding. Cogn Sci Multidiscipl J. 1999;23(4):517–42.

19. Diederich J. Instruction and high-level learning in connectionist networks. Connect Sci. 1989;1:161–80.

20. Dudek SM, Bear MF. Homosynaptic long-term depression in area Ca1 of hippocampus and effects of N-methyl-D-aspartate receptor blockade. Proc Natl Acad Sci USA. 1992;89(10):4363–7.

21. Eggert J, van Hemmen JL. Unifying framework for neuronal assembly dynamics. Phys Rev E. 2000;61(2):1855–74.

22. Elman JL, Bates EA, Johnson MH, Karmiloff-Smith A, Parisi D, Plunkett K. Rethinking innateness: a connectionist perspective on development. Cambridge, MA: MIT Press; 1996.

23. Fadiga L, Craighero L, Buccino G, Rizzolatti G. Speech listening specifically modulates the excitability of tongue muscles: a TMS study. Eur J NeuroSci. 2002;15(2):399–402.

24. Feldman J. Dynamic connections in neural networks. Biol Cybern. 1982;46:27–39.

25. Fry DB. The development of the phonological system in the normal and deaf child. In: Smith F, Miller GA, editors. The genesis of language. Cambridge, MA: MIT Press; 1966. p. 187–206.

26. Garagnani M, Wennekers T, Pulvermüller F. A neuronal model of the language cortex. Neurocomputing. 2007;70:1914–9.

27. Garagnani M, Wennekers T, Pulvermüller F. A neuroanatomically grounded Hebbian-learning model of attention-language interactions in the human brain. Eur J NeuroSci. 2008;27(2):492–513.

28. Gaskell MG, Hare M, Marslen-Wilson WD. A connectionist model of phonological representation in speech perception. Cogn Sci: Multidis J. 1995;19(4):407–39.

29. Gilbert CD, Wiesel TN. Clustered intrinsic connections in cat visual-cortex. J Neurosci. 1983;3(5):1116–33.

30. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. Econometrica. 1969;37(3):424–38.

31. Grossberg S. Adaptive pattern-classification and universal recoding.1. Parallel development and coding of neural feature detectors. Biol Cybern. 1976;23(3):121–34.

32. Grossberg S. Adaptive pattern-classification and universal recoding.2. Feedback, expectation, olfaction, illusions. Biol Cybern. 1976;23(4):187–202.

33. Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang. 2006;96(3):280–301.

34. Harris KD. Neural signatures of cell assembly organization. Nat Rev Neurosci. 2005;6(5):399–407.

35. Hebb DO. The organization of behavior. New York: John Wiley; 1949.

36. Hirsch JC, Barrionuevo G, Crepel F. Homo- and heterosynaptic changes in efficacy are expressed in prefrontal neurons: an in vitro study in the rat. Synapse. 1992;12(1):82–5.

37. Husain FT, Tagamets MA, Fromm SJ, Braun AR, Horwitz B. Relating neuronal dynamics for auditory object processing to neuroimaging activity: a computational modeling and an fMRI study. Neuroimage. 2004;21(4):1701–20.

38. Izhikevich EM, Desai NS. Relating STDP to BCM. Neural Comput. 2003;15(7):1511–23.

39. Joanisse MF, Seidenberg MS. Impairments in verb morphology after brain injury: a connectionist model. Proc Natl Acad Sci USA. 1999;96(13):7592–7.

40. Katz LC, Shatz CJ. Synaptic activity and the construction of cortical circuits. Science. 1996;274(5290):1133–8.

41. Kirkwood A, Rioult MC, Bear MF. Experience-dependent modification of synaptic plasticity in visual cortex. Nature. 1996;381(6582):526–8.

42. Kohonen T. Self-organisation and associative memory. Berlin: Springer; 1984.

43. Kohonen T, Makisara K. The self-organizing feature maps. Phys Scripta. 1989;39(1):168–72.

44. Krichmar JL, Seth AK, Nitz DA, Fleischer JG, Edelman GM. Spatial navigation and causal analysis in a brain-based device modeling cortical-hippocampal interactions. Neuroinformatics. 2005;3(3):197–221.

45. Levelt WJM, Roelofs A, Meyer AS. A theory of lexical access in speech production. Behav Brain Sci. 1999;22:1–75.

46. Linsker R. Self-organization in a perceptual network. Computer. 1988;21(3):105–17.

47. Makris N, Meyer JW, Bates JF, Yeterian EH, Kennedy DN, Caviness VS. MRI-based topographic parcellation of human cerebral white matter and nuclei II. Rationale and applications with systematics of cerebral connectivity. Neuroimage. 1999;9(1):18–45.

48. Malenka RC, Bear MF. LTP and LTD: an embarrassment of riches. Neuron. 2004;44(1):5–21.

49. Malenka RC, Nicoll RA. Neuroscience—long-term potentiation—a decade of progress? Science. 1999;285(5435):1870–4.

50. Marslen-Wilson WD, Tyler LK. The temporal structure of spoken language understanding. Cognition. 1980;8:1–71.

51. McClelland JL, Elman JL. The trace model of speech perception. Cogn Psychol. 1986;18:1–86.

52. McClelland JL, Rumelhart DE. Distributed memory and the representation of general and specific information. J Exp Psychol Gen. 1985;114:159–88.

53. Miller KD. Synaptic economics: competition and cooperation in synaptic plasticity. Neuron. 1996;17(3):371–4.

54. Miller KD, Mackay DJC. The role of constraints in Hebbian learning. Neural Comput. 1994;6(1):100–26.

55. Milner PM. Neural representation: some old problems revisited. J Cogn Neurosci. 1996;8:69–77.

56. Ngezahayo A, Schachner M, Artola A. Synaptic activity modulates the induction of bidirectional synaptic changes in adult mouse hippocampus. J Neurosci. 2000;20(7):2451–8.

57. Norris D. Shortlist—a connectionist model of continuous speech recognition. Cognition. 1994;52(3):189–234.

58. Page M. Connectionist modelling in psychology: a localist manifesto. Behav Brain Sci. 2000;23(4):443–67.

59. Palm G. Neural assemblies. Berlin: Springer; 1982.

60. Pandya DN, Yeterian EH. Architecture and connections of cortical association areas. In: Peters A, Jones EG, editors. Cerebral cortex. Vol. 4. Association and auditory cortices. London: Plenum Press; 1985. p. 3–61.

61. Parker GJ, Luzzi S, Alexander DC, Wheeler-Kingshott CA, Ciccarelli O, Lambon Ralph MA. Lateralization of ventral and dorsal auditory-language pathways in the human brain. Neuroimage. 2005;24(3):656–66.

62. Penke M, Westermann G. Broca's area and inflectional morphology: evidence from broca's aphasia and computer modeling. Cortex. 2006;42(4):563–76.

63. Petrides M, Pandya DN. Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. Eur J NeuroSci. 2002;16(2):291–310.

64. Plaut DC, McClelland JL, Seidenberg MS, Patterson K. Understanding normal and impaired word reading: computational principles in quasi-regular domains. Psychol Rev. 1996;103:56–115.

65. Press WH, Teukolski SA, Vetterling WT, Flannery BP. Numerical recipes in C: the art of scientific computing. 2nd ed. Cambridge: Cambridge University Press; 1992.

66. Pulvermüller F. Words in the brain's language. Behav Brain Sci. 1999;22:253–336.

67. Pulvermüller F, Preissl H. A cell assembly model of language. Netw Comput Neural Syst. 1991;2:455–68.

68. Pulvermüller F, Shtyrov Y. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. Prog Neurobiol. 2006;79(1):49–71.

69. Pulvermüller F, Kujala T, Shtyrov Y, Simola J, Tiitinen H, Alku P, et al. Memory traces for words as revealed by the mismatch negativity. Neuroimage. 2001;14(3):607–16.

70. Pulvermüller F, Huss M, Kherif F, Moscoso del Prado Martin F, Hauk O, Shtyrov Y. Motor cortex maps articulatory features of speech sounds. Proc Natl Acad Sci USA. 2006;103(20):7865–70.

71. Rioult-Pedotti MS, Friedman D, Donoghue JP. Learning-induced LTP in neocortex. Science. 2000;290(5491):533–6.

72. Rogers TT, McClelland JL. Semantic cognition. Cambridge, MA: MIT Press; 1994.

73. Rogers TT, Lambon Ralph MA, Garrard P, Bozeat S, McClelland JL, Hodges JR, et al. Structure and deterioration of semantic memory: a neuropsychological and computational investigation. Psychol Rev. 2004;111(1):205–35.

74. Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci. 1999;2(12):1131–6.

75. Saur D, Kreher BW, Schnell S, Kummerer D, Kellmeyer P, Vry MS, et al. Ventral and dorsal pathways for language. Proc Natl Acad Sci USA. 2008;105(46):18035–40.

76. Seidenberg MS, McClelland JL. A distributed, developmental model of word recognition and naming. Psychol Rev. 1989;96(4):523–68.

77. Sejnowski TJ. Storing covariance with nonlinearly interacting neurons. J Math Biol. 1977;4(4):303–21.

78. Sejnowski TJ, Tesauro G. The Hebb rule for synaptic plasticity: algorithms and implementations. In: Byrne JH, Berry WO, editors. Neural models of plasticity. Sandiego: Academic Press; 1989. p. 94–103.

79. Seth AK, Edelman GM. Distinguishing causal interactions in neural populations. Neural Comput. 2007;19(4):910–33.

80. Shastri L. Semanti networks: an evidential formalization and its connectionist realization. Los Altos/London: Morgan Kaufmann/Pitman Publishing Company; 1988.

81. Shastri L. Biological grounding of recruitment learning and vicinal algorithms in long-term potentiation. In: Austin J, Wermter S, Willshaw D, editors. Emergent neural computational architectures based on neuroscience, lecture notes in computer science, vol. 2036. Berlin: Springer-Verlag; 2001. p. 348–67.

82. Shatz CJ. Impulse activity and the patterning of connections during CNS development. Neuron. 1990;5(6):745–56.

83. Shouval H, Intrator N, Cooper LN. BCM network develops orientation selectivity and ocular dominance in natural scene environment. Vision Res. 1997;37(23):3339–42.

84. Song S, Miller KD, Abbott LF. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nat Neurosci. 2000;3(9):919–26.

85. Stanton PK, Sejnowski TJ. Associative long-term depression in the hippocampus induced by hebbian covariance. Nature. 1989;339(6221):215–8.

86. Tagamets MA, Horwitz B. Integrating electrophysiological and anatomical experimental data to create a large-scale model that simulates a delayed match-to-sample human brain imaging study. Cereb Cortex. 1998;8(4):310–20.

87. Tsumoto T. Long-term potentiation and long-term depression in the neocortex. Prog Neurobiol. 1992;39:209–28.

88. Turrigiano GG, Nelson SB. Homeostatic plasticity in the developing nervous system. Nat Rev Neurosci. 2004;5(2):97–107.

89. Turrigiano GG, Leslie KR, Desai NS, Rutherford LC, Nelson SB. Activity-dependent scaling of quantal amplitude in neocortical neurons. Nature. 1998;391(6670):892–6.

90. Watkins KE, Paus T. Modulation of motor excitability during speech perception: the role of Broca's area. J Cogn Neurosci. 2004;16(6):978–87.

91. Watkins KE, Strafella AP, Paus T. Seeing and hearing speech excites the motor system involved in speech production. Neuropsychologia. 2003;41(8):989–94.

92. Wennekers T, Sommer F, Aertsen A. Theories in bioscience—editorial: cell assemblies. Theory Biosci. 2003;122(1):1–4.

93. Westermann G, Miranda ER. A new model of sensorimotor coupling in the development of speech. Brain Lang. 2004;89(2):393–400.

94. Wilson HR, Cowan JD. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik. 1973;13:35–80.

95. Wilson SM, Saygin AP, Sereno MI, Iacoboni M. Listening to speech activates motor areas involved in speech production. Nat Neurosci. 2004;7(7):701–2.

96. Zatorre RJ, Meyer E, Gjedde A, Evans AC. PET studies of phonetic processing of speech: review, replication, and reanalysis. Cereb Cortex. 1996;6(1):21–30.