

# Words in the brain's language

**Friedemann Pulvermüller**

Department of Psychology, University of Konstanz,

78434 Konstanz, Germany

friedemann.pulvermueller@uni-konstanz.de

www.clinical-psychology.uni-konstanz.de

**Abstract:** If the cortex is an associative memory, strongly connected cell assemblies will form when neurons in different cortical areas are frequently active at the same time. The cortical distributions of these assemblies must be a consequence of where in the cortex correlated neuronal activity occurred during learning. An assembly can be considered a functional unit exhibiting activity states such as full activation (“ignition”) after appropriate sensory stimulation (possibly related to perception) and continuous reverberation of excitation within the assembly (a putative memory process). This has implications for cortical topographies and activity dynamics of cell assemblies forming during language acquisition, in particular for those representing words. Cortical topographies of assemblies should be related to aspects of the meaning of the words they represent, and physiological signs of cell assembly ignition should be followed by possible indicators of reverberation. The following postulates are discussed in detail: (1) assemblies representing phonological word forms are strongly lateralized and distributed over perisylvian cortices; (2) assemblies representing highly abstract words such as grammatical function words are also strongly lateralized and restricted to these perisylvian regions; (3) assemblies representing concrete content words include additional neurons in both hemispheres; (4) assemblies representing words referring to visual stimuli include neurons in visual cortices; and (5) assemblies representing words referring to actions include neurons in motor cortices. Two main sources of evidence are used to evaluate these proposals: (a) imaging studies focusing on localizing word processing in the brain, based on stimulus-triggered event-related potentials (ERPs), positron emission tomography (PET), and functional magnetic resonance imaging (fMRI), and (b) studies of the temporal dynamics of fast activity changes in the brain, as revealed by high-frequency responses recorded in the electroencephalogram (EEG) and magnetoencephalogram (MEG). These data provide evidence for processing differences between words and matched meaningless pseudowords, and between word classes, such as concrete content and abstract function words, and words evoking visual or motor associations. There is evidence for early word class-specific spreading of neuronal activity and for equally specific high-frequency responses occurring later. These results support a neurobiological model of language in the Hebbian tradition. Competing large-scale neuronal theories of language are discussed in light of the data summarized. Neurobiological perspectives on the problem of serial order of words in syntactic strings are considered in closing.

**Keywords:** associative learning; cell assembly; cognition; cortex; ERP; EEG; fMRI; language; lexicon; MEG; PET; word category

## 1. Words in the brain: Where? Why? How?

Human language production is caused by neuronal activity and any speech signal necessarily activates neurons in the brains of listeners when being perceived. It is the purpose of language science to specify these processes and their underlying mechanisms. However, owing to the enormous complexity of language and the sparsity of our knowledge about brain functioning, neuroscientists, psychologists, and linguists have not attacked this goal directly. Indeed, biological knowledge currently available is still far from making it possible to spell out the great variety of language phenomena in terms of neurons. Nevertheless, it is possible to choose paradigmatic questions about language and to try to find answers for them based on biological principles. I will use this strategy here to approach the problem of language and the brain.

The issue I would like to address is that of different vocabulary classes. At school, one learns to categorize words into fifty or so lexical categories, such as noun or verb, and one may also be asked to categorize words on the basis of their meaning, according to semantic criteria. Of course it is useful, for didactic purposes, to make a large number of distinctions between classes of words, not only based on their meaning and their function in syntactic structures, but also based on criteria such as their intonation, syllable com-

plexity, number of letters or speech sounds, or the frequency with which they are used in ordinary language. However, one may wonder whether some of these distinctions reflect differences that are biologically real. This would mean that the members of word classes A and B, which can be distinguished on the basis of linguistic or didactic criteria, would also be represented differently in the human brain. In psycholinguistics, much effort has been expended to demonstrate processing differences between word classes, for example between the major lexical classes



FRIEDEMANN PULVERMÜLLER has an M.A. in Biology, a Ph.D. in Linguistics, a *Habilitation* in Psychology and Medicine, and is now *Privatdozent* of Psychology at the University of Konstanz, Germany. He is the author of over 80 publications in the area of cognitive neuroscience, including his recent book on neurobiology of language and the forthcoming “Neuronal Grammar.” Among his honors are an early career award from the Society for Psychophysiological Research and a Helmholtz and a Heisenberg Fellowship. His main scientific interest is to spell out language mechanisms in terms of neurons.

called *content words* (or open-class words, including nouns, verbs, and adjectives) and *function words* (or closed-class words, including articles, pronouns, auxiliary verbs, conjunctions, and so on. Some of these studies will be discussed in sect. 5.). It is good to know that two word groups are different; however, it is better to know (or to have an idea about) what the actual differences are. A biological approach aims at specifying the difference in terms of neurons and neuronal connections.

In recent years, more and more neuropsychological studies have been devoted to the investigation of cortical mechanisms necessary for word processing, and psychophysiological studies have been investigating the brain areas that “light up” when words are being produced or comprehended. Such studies are most welcome because they may contribute to an answer of the “where” question, that is, the question of where representations are housed and processes take place. However, even when questions such as “Which word classes will be selectively impaired after focal brain lesion in cortical area X?” or “Which brain areas will become active when words of class A are being produced or comprehended?” have been definitely answered, the question of why this is so may still be open. Why are words of class A processed in area X? An explanation of language mechanisms in the brain is only possible if such “why” questions can be answered from known biological principles. But even definite and exhaustive answers to “where” and “why” questions may still not be a satisfactory end point of cognitive neuroscientific research: If it is clear where in the brain particular language units are represented and processed, and if it is clear why this is so, one can still ask how language representations are laid down, and how they are activated when language units are being processed.

This target article will certainly not provide complete answers to “where,” “why,” and “how” questions related to language. It will provide preliminary answers to the “where” question as far as words of certain classes are concerned; it hopes to convince the reader that the “why” question can be answered in a few clear cases; and it tries to specify some very basic features of cortical representations and the way they become active and maintain their activity. All this is done on the basis of a brain model rooted in Hebb’s concept of cell assemblies. In fact, the purpose of this article is not only to discuss the issue of words in the brain, but to make it evident that the Hebbian approach is a powerful tool for cognitive neuroscience that may lead to a biological explanation of our language capacity and of other higher cognitive capacities as well.

## 2. The Hebbian model, recent modifications and evidence

In the late 1940s, Donald Hebb (1949) proposed a neuropsychological theory of cortical functioning that can be considered an alternative to both localizationist and holistic approaches. Localizationists would assume that small cortical areas are fully capable of performing complex cognitive operations. A localizationist would, for example, propose that an area of a few square centimeters of cortical surface is the locus of word comprehension (Broca 1861; Lichtheim 1885; Wernicke 1874). According to this view, the psychological process (word comprehension) is restricted to one area – that is, no other areas are assumed to contribute to

this specific process. Only under pathological conditions or during development may there be a shift of the process to another equally narrow area (Luria 1970; 1973). In contrast, a holistic approach would imply that the entire cortex exhibits equipotentiality with regard to all cognitive operations and that all cortical areas (or even brain parts) can contribute to sufficiently complex processes, such as those involved in language (for discussion, see Freud 1891, Lashley 1950, and, for an overview, Deacon 1989).

The Hebbian proposal is in sharp contrast to both of these views. Cell assemblies with defined cortical topographies are assumed to form the neurobiological representations of cognitive elements such as gestalt-like figures or words. This position is radically different from a localizationist approach, because it assumes that neurons in different cortical areas may be part of the same distributed functional unit. The Hebbian viewpoint is also different from the holistic view that “everything is equally distributed,” because it implies that the representation of, for example, an image may involve cortical areas entirely different from those contributing to the representation of, say, an odor. Accordingly, the representation of a word would not be restricted to a small cortical locus, but would be distributed over well-defined areas, for example over Broca’s, Wernicke’s, and some other areas.

The Hebbian model is based on three fundamental assumptions about cortical functioning, which can be summarized as follows:

1. Coactivated neurons become associated.
2. Associations can occur between adjacent or distant neurons; that is, the entire cortex is an associative memory.
3. If neurons become associated, they will develop into a functional unit, a cell assembly.

Hebb was frequently criticized, because his assumptions were considered too speculative and because some of his colleagues believed that his ideas would not be testable. Therefore, it is necessary to discuss his assumptions in light of evidence presently available.

Electrophysiological studies have demonstrated that having cortical neurons frequently active at the same time strengthens their connections. If a neuron, call it L, sends one connection to a second neuron, M, their synapse will strengthen when both are repeatedly active together, so that L will later have a stronger influence on M. Because this effect may last for many hours or days, or even longer, it has been termed *long-term potentiation* (LTP) (Ahissar et al. 1992; Gustafsson et al. 1987). After this kind of associative learning, connection strength will be a function of the frequency of coincident activity. Table 1 describes this kind of coincidence learning (Palm 1982).

One may object to this and similar learning rules that coincidence learning is only one form of associative learning known to take place between neocortical neurons. If only one of the two neurons is active while the other one remains silent, this could also have an effect on the strength of their connection. In fact, it was shown by electrophysiological experiments that activation of presynaptic neuron L alone, while the membrane potential of postsynaptic neuron M is stable (or only slightly depolarizes), leads to a weakening of their synaptic connection (Artola et al. 1990; Artola & Singer 1987; 1993; Rauschecker & Singer 1979). Because this reduction (or depression) of the influence of one neuron on the other is long-lasting, the phenomenon has been called *long-term depression* (LTD). There is also evidence for

Table 1. *Associative synaptic learning according to a Hebbian coincidence rule*

		neuron L	
		active	inactive
neuron M	active	+w*	--
	inactive	--	--

\*+w indicates an increase in connection strength between neurons L and M; hyphens indicate no change in connection strength.

LTD occurring when presynaptic neurons are silent while postsynaptic neurons fire frequently (Tsumoto 1992; Tsumoto & Suda 1979). Therefore, the original idea proposed by Hebb needs a slight but important modification: Connection strength is not only modified by coincident activity, it also changes if only one of two connected neurons is active while the other one is inactive. Table 2 describes this kind of learning, which will be called *correlation learning*, because after this kind of synaptic modification, the strength of the synaptic connection will include information not only about the frequency of coincident firing of neurons, but also about how strong the correlation was between their activations.

This formulation is very general. It does not make distinctions implied by more precise formulations of synaptic learning rules (Artola & Singer 1993; Bienenstock et al. 1982; Tsumoto 1992), in which, for example, the states called “active” and “inactive” above, have been replaced by gradual activity levels (quantified in terms of the frequency of action potentials or the membrane potential of the postsynaptic neuron). In addition, the above formulations leave open the questions of how the w-values should actually be chosen. Whereas  $w_1$  may be assumed to be larger than  $w_2$  and  $w_3$ , the exact values of the variables are unknown. These questions will not be addressed here, because they have been discussed in great detail based on what is known about synaptic dynamics in the neocortex (Tsumoto 1992) and in light of storage properties of artificial associative networks (Palm 1982; Palm & Sommer 1995; Willshaw & Dayan 1990). In the present context, it is most important to keep in mind that a correlation rule, rather than a coincidence rule, is a fundamental principle of synaptic learning in the cortex.

It appears uncontroversial that excitatory cortical neurons located close to each other are likely to have a synap-

Table 2. *Associative synaptic learning according to a correlation rule*

		neuron L	
		active	inactive
neuron M	active	+ $w_1$ *	- $w_2$
	inactive	- $w_3$	--

\*+ $w_1$ , - $w_2$ , and - $w_3$  indicate positive or negative changes in connection strength.

tic contact. Although this is not a 100% probability – it is actually far below (Braitenberg 1978a; Braitenberg & Schüz 1991) – it is evident that adjacent neurons are much more likely to be connected than neurons located far apart, that is, in distant cortical areas (Young et al. 1995). It is clear from neuroanatomical studies, however, that most cortical pyramidal cells have long axons reaching distant areas or subcortical structures, and that connections from one area project to several other areas. In the Macaca, for example, what may be considered the homologues of Broca’s and Wernicke’s areas are not only intensely connected to each other; they also exhibit connections to additional premotor, higher visual, and association cortices (Deacon 1992a; 1992b; Pandya & Vignolo 1971; Pandya & Yeterian 1985). Therefore, if correlated neuronal activity is present in a large number of neurons in different cortical areas, some of these neurons will exhibit direct connections to each other. These neurons will become more strongly associated even if they are located far apart. Thus, although the cortex is not a fully connected associative memory in which every processing unit is connected to every other one, it still appears to be an associative network well suited to allow for both local and between-area associative learning (Braitenberg & Schüz 1991; 1998; Fuster 1994; Palm 1982).

If neurons in an associative network exhibit correlated activity, they will be a stronger influence on each other. This implies that these neurons will be more likely to act together as a group. Hebb (1949) calls such anatomically and functionally connected neuron groups “cell assemblies.” The strong within-assembly connections are likely to have two important functional consequences: (1) If a sufficiently large number of the assembly neurons are stimulated by external input (either through sensory fibers or through cortico-cortical fibers), activity will spread to additional assembly members and, finally, the entire assembly will be active. This explosion-like process has been called ignition of the assembly (Braitenberg 1978b). (2) After an assembly has ignited, activity will not stop immediately (because of fatigue or regulation processes), but the strong connections within the assembly will allow activity for some time. Cell assemblies are sometimes conceptualized as packs of neurons without an ordered inner structure. However, according to Hebb’s (1949) proposal, assembly neurons are connected so that ordered spreading and reverberation of neuronal activity can occur.

The latter point needs further elaboration: Figure 1 is taken from Hebb’s 1949 book and depicts what the author believed to be a possible inner structure of an assembly. In this diagram, arrows represent subgroups of neurons included in the assembly. These subgroups would each become active at exactly the same point in time. Arrowheads indicate the other subgroups to which a given subgroup would project, and numbers denote a possible activity sequence. After synchronous activity of the neurons represented by the arrow labeled “1,” a wave of excitation will run through the assembly as indicated by the numbers, and activity will finally cease. Thus, it is evident that in Hebb’s early proposal, a cell assembly was already conceptualized as a highly structured entity. Whereas ignition of the assembly may simultaneously involve all assembly neurons, it is also possible to have a wave of excitation circulating and reverberating in the many loops of the assembly. The wave can be described as a *spatiotemporal pattern* of activity in which many cortical neurons participate.

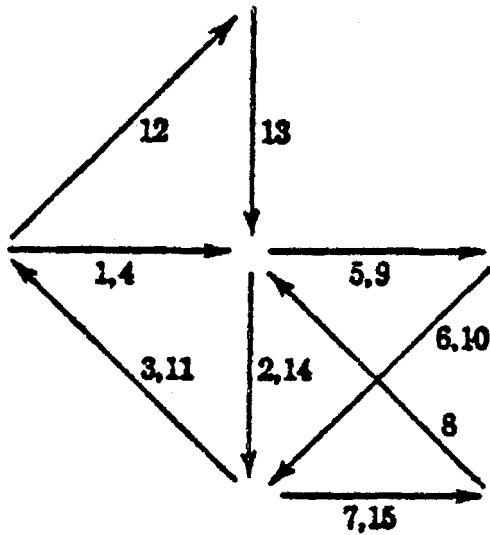


Figure 1. Hebb's (1949) illustration of the inner structure of a cell assembly consisting of several subgroups of neurons. Arrows represent subgroups of neurons that become active at exactly the same time. Numbers indicate the activation sequence following activity of the subgroup labelled 1. An ordered spatiotemporal pattern of activity is produced whenever a wave of excitation runs through the assembly.

The question of whether cell assemblies that represent stimuli and cognitive entities exist in the cortex has long been thought impossible to test by empirical research. As mentioned earlier, this belief was probably one of the main reasons why Hebb's theory was not generally accepted in the 1940s and 1950s. However, more recent experimental work has provided strong evidence for the Hebbian ideas. Neurophysiological work by Abeles, Aertsen, Gerstein, and their colleagues (Abeles 1982; 1991; Abeles et al. 1993; 1994; Aertsen et al. 1989; Gerstein et al. 1989) revealed exactly timed spatiotemporal firing patterns in cortical neurons. The specific neuronal connections these patterns are probably related to were labeled *synfire chains* by Abeles, because a subpopulation of neurons must synchronously activate the next subpopulation to keep the chain going. It is important to note that spatiotemporal activity patterns actually detected in cortical neurons frequently involve the repeated activation of a given neuron, thus suggesting reverberations caused by loops in the chain (Abeles et al. 1993). Evidently, the concept of a reverberating synfire chain emerging from recent neurophysiological data comes very close to Hebb's original proposal summarized in Figure 1. In contrast to the original proposal, it appears more realistic to postulate connections not only between consecutive subpopulations of neurons, but also connections that skip subgroups and directly link, for example, subgroups 1 and 3 in the example illustration (Fig. 1). Such bypass connections may be realized by relatively slowly-conducting cortico-cortical fibers (Miller 1996). Furthermore, Abeles's findings suggest that the neuron subgroups represented by arrows in Hebb's diagram overlap, so that a given neuron can be part of, say, subgroups 1 and 7.

In summary, after its full activation (ignition), neuronal activity may reverberate in the loops of an assembly. Ignition and reverberation may represent important functional states of Hebbian cell assemblies. On the cognitive level, ig-

nition may correspond to perception of a meaningful stimulus and to activation of its representation. The fact that an object partially hidden behind another one can frequently be identified can be explained by full ignition of a cell assembly after stimulation of only some of its neurons (Hebb 1949). Sustained activity of the assembly and reverberation of activity therein may represent an elementary process underlying short-term or active memory (Fuster 1989; 1995; Fuster & Jervey 1981). The latter view arises from studies that evidence a systematic relationship between the occurrence of defined spatio-temporal activity patterns in cortex and particular engrams an experimental animal has to keep in active memory (Fuster 1995; Villa & Fuster 1992).

Recent neurophysiological work not only revealed well-timed spatiotemporal activity patterns in cortical neurons related to memory processes but another line of research uncovered stimulus-specific synchronization of activity in cortical neurons related to perceptual processes. If an elementary visual stimulus, for example a bar moving in a particular direction, is presented to an experimental animal, numerous neurons in various visual cortices in both hemispheres start to synchronize their firing and, in many cases, exhibit coherent rhythmic activity in a relatively high frequency range, that is, above 20 Hz (Eckhorn et al. 1988; Engel et al. 1990; 1991b; Gray et al. 1989; Kreiter & Singer 1992).<sup>1</sup> This provides further evidence that neurons in different areas are strongly coupled and can act as a unit. Although synchronization phenomena have been observed in subcortical structures and even in the retina (Kirschfeld 1996; Neuenschwander & Singer 1996; Sillito et al. 1994; Steriade et al. 1993), cortico-cortical connections are apparently necessary for synchronization of neuron responses in cortex (Engel et al. 1991a; Gray et al. 1989; Singer & Gray 1995). Because synchronized responses change with stimulus features, for example the direction in which a bar moves (Eckhorn et al. 1988; Gray et al. 1989; Gray & Singer 1989), the idea receives support that there are stimulus-specific distributed neuron groups. It appears that these neurophysiological data can only be explained if cell assemblies are assumed that are (a) activated by specific external stimuli, (b) distributed over different cortical areas, and (c) connected through cortico-cortical fibers (and possibly additional subcortical connections).

These results can be interpreted as evidence for a version of Hebb's theory according to which cell assemblies must synchronously oscillate at high frequencies when active. However, synchronous oscillations are a special case of well-timed activity (Abeles et al. 1993; Aertsen & Arndt 1993). Therefore, these data are also consistent with the weaker position made explicit by Hebb that cell assemblies generate well-timed activity patterns in their many neurons. The latter position would imply that at least a fraction of the activated neurons (e.g., those forming one subgroup represented by an arrow in Fig. 1) exhibit synchronized activity when the assembly reverberates (see Pulvermüller et al. 1997 for further discussion).

If it is taken into account that most cortico-cortical fibers conduct action potentials with velocities around 5–10 m/s or faster (Aboitiz et al. 1992; Patton 1982), it becomes clear that a wave of activity running through and reverberating within an assembly will lead to rather fast activity changes. Suppose a large-scale physiological recording device (e.g., an electrode recording the local field potential, or even an EEG electrode or an MEG coil) is placed close to a frac-



tion of the neurons of the assembly sketched in Figure 1. In this case, a reverberating wave of activity in the assembly will cause rather fast activity changes at the recording device. If the neuronal subpopulations represented by arrows are assumed to be located in different cortical areas separated, say, by a few centimeters, it will take some hundredths of a second for neuronal activity to travel the loop labelled 1-2-3 and for the neurons denoted by the first arrow (the first and the fourth in the sequence) to become synchronously active for the second time. It follows that synchronous and fast reverberating activity in the assembly is most likely to lead to spectral dynamics in the high frequency range ( $>20$  Hz) recorded by the large-scale devices.<sup>2</sup>

If specific dynamics in high-frequency cortical activity are taken as an indicator of reverberating activity in Hebbian cell assemblies, the question of whether particular cognitive processes are related to high-frequency dynamics becomes particularly relevant for further testing the Hebbian ideas. It is known from animal experiments that if the receptive fields of two neurons in visual cortices are each stimulated by a moving bar and both stimuli are aligned and move together in the same direction, neuron responses can synchronize their fast rhythmic activity. However, if one neuron is stimulated by a bar moving in a particular direction, while the other is stimulated by a bar moving in the opposite direction, synchrony of rhythmic responses vanishes (Engel et al. 1991a). This result and similar findings indicate that synchrony of high-frequency neuronal activity reflects gestalt criteria, for example the fact that two objects move together (Singer 1995; Singer & Gray 1995). Consistent with this finding in animals, patterns of regularly moving bars have been found to evoke stronger high-frequency electrocortical responses recorded in the EEG compared to irregular bar patterns (Lutzenberger et al. 1995). Further support for the role of high-frequency cortical activity in cognitive processing comes from studies of electrocortical responses to attended and unattended stimuli (Tiitinen et al. 1993). Most important, gestalt-like figures such as Kanizsa's triangle have led to stronger high-frequency EEG responses around 30 Hz compared to physically similar stimuli that are not perceived as a coherent gestalt (Tallon et al. 1995; Tallon-Baudry et al. 1996). Thus, dynamics of high-frequency responses appear to be an indicator of the cognitive process of gestalt perception. These results are consistent with the idea that gestalts, such as a coherent bar pattern or a triangle, activate cortical cell assemblies that generate coherent high-frequency responses, while physically similar stimuli that are not perceived as coherent gestalts lack cortical representations and, therefore, evoke desynchronized electrocortical responses. Therefore, the idea that cell assemblies are relevant for cognitive processing not only receives support from recordings in animals' brains, but is consistent with noninvasive recordings of human brain activity using large-scale recording techniques such as EEG.

In summary, recent theoretical and empirical research provides support for the existence of Hebbian cell assemblies and for their importance for cognitive brain processes. It must be noted, however, that, based on experimental and theoretical work, the Hebbian concept and the assumptions connected with it have changed slightly. Some of these modifications are summarized in the following postulates (which are closely related to points (1) to (3) in sect. 2):

1'. Simultaneous pre- and postsynaptic activity of cortical neurons leads to synaptic strengthening. However, pre- or postsynaptic activity alone leads to synaptic weakening.

2'. Associations can occur between adjacent neurons and between cortical neurons located far apart, provided there is a synapse connecting them. The cortex is an associative memory although it is not fully connected.

3'. If synaptic strengthening occurs among many neurons, they will develop into an assembly that can ignite and exhibit well-timed reverberatory activity.

Future empirical testing of the modified Hebbian framework is, of course, necessary, and neuroimaging techniques make it possible to perform such testing, although techniques available at present do not allow for localizing each member of a widely distributed neuron set in different cortical areas. If an assembly ignites and stays active, signs of activity should be visible in single-cell and multiple-unit responses, local field potentials, and more global electrocortical activity, and possibly in metabolic changes in the brain as well. The cortical topography of these activity signs may allow for conclusions concerning assembly topographies. In addition to general signs of activity enhancement – enhanced blood flow, larger event-related potentials, more powerful single-cell responses – changes in well-timed high-frequency cortical responses may include information about reverberatory neuronal activity in cell assemblies.

It may be appropriate at this point to mention possible theoretical problems of the Hebbian approach, some of which have been summarized in a recent article by Milner (1996). If an ignition takes place, there is danger that activity will spread to additional assemblies and finally to the entire cortex or even brain, resulting in overactivity such as that seen during seizures. To avoid this, it is necessary to have a control device regulating the cortical equilibrium of activity. This device has been called “threshold control mechanism” (Braitenberg 1978b) and its neuroanatomical substrate has been proposed to be located in the basal ganglia (Miller & Wickens 1991; Wickens 1993) or, as an alternative, in the hippocampus (Fuster 1995). Furthermore, if a large number of cell assemblies are built up in the cortex, this may lead to an increase in average connection strength, and, in the worst case, to a clumping together of all assemblies. This would make it impossible to activate representations individually. However, this problem primarily occurs if a coincidence learning rule is assumed (Table 1). If LTD rules are added (e.g., in the case of correlation-based learning as sketched in Table 2), simultaneous activity of a set of cortical neurons will not only lead to synaptic strengthening between them, but also to a weakening of connections to neurons outside the set (Hetherington & Shapiro 1993; Palm 1990; Willshaw & Dayan 1990). In this case, the problem will occur only if w-parameters (see Table 2) are chosen inappropriately. It has also been argued that the cell assembly framework is not flexible enough to allow for a representation of complex objects. If a house includes a door and a window, how would the respective representations relate to each other? Here, it is necessary to allow for hierarchical organizations of cell assemblies: One assembly may be a subset of another one. This is also important for the semantic representations of words with similar meanings, for hyponyms and hyperonyms. Adjustment of the global activation threshold may account for whether the set or its subset is being activated (Braitenberg 1978b). Furthermore, concepts that have features in com-

mon may be represented in cell assemblies that share some of their neurons. These assemblies will, therefore, not be entirely different neuron sets, but they will overlap. The relations of inclusion and overlap can be realized quite naturally within a cell-assembly theory built on the Hebbian notion (Braitenberg 1978b; Palm 1982). Therefore, a modified version of the original Hebbian proposal appears to be well suited to provide neurobiological answers to important questions in cognitive science.

### 3. Cortical distribution of cell assemblies

In recent years, the Hebbian idea of distributed assemblies with defined cortical topographies has been incorporated into large-scale neuronal theories of language and other cognitive functions (Abeles 1991; Braitenberg & Schüz 1991; Damasio 1989a; Edelman 1992; Elbert & Rockstroh 1987; Fuster 1995; Gerstein et al. 1989; Mesulam 1990; Miller & Wickens 1991; Palm 1982; Posner & Raichle 1994; Pulvermüller 1992; Singer 1995; Wickens et al. 1994). At this point, there appears to be a consensus that neurons in distant cortical areas can work together as functional units. However, the Hebbian framework would not only postulate that there are large-scale neuronal networks, it also provides clear-cut criteria for the formation of cell assemblies and, therefore, straightforward predictions on assembly topographies.

For assembly formation, Hebb (1949) outlines the following scenario (pp. 235f): If a particular object is frequently being visually perceived, a set of neurons in visual cortices will repeatedly become active at the same time. Therefore, a cell assembly will form representing the shape of the object. This assembly is distributed over cortical regions where simultaneous neuronal activity is evoked by visual stimulation, that is, in primary and higher-order visual cortices in the occipital lobes, for example in Brodmann's (1909) areas 17, 18, 19, and 20. For convenience, Figure 2 displays a lateral view of the left cortical hemisphere on which the approximate locations of Brodmann's areas are indicated. If correlated neuronal activity is caused by input through other sensory modalities, or if it is related to motor output, the cortical distribution of the coactivated set of neurons will be different. For example, if motor behavior co-occurs with sensory stimulation, cell assemblies may form including neurons in motor and sensory cortices. To put it in a more general way, the cortical localization of a representation is a function of where in the cortex simultaneous activity occurred when the representation was acquired or learned.

Whereas correlated neuronal activity of a connected cortical neuron set is a sufficient condition for cell assembly formation, correlated occurrence of sensory stimuli is not. In the most extreme case, when an individual is asleep, correlated stimuli (e.g., in the somatosensory and acoustic modality) may not cause enough cortical activity to lead to synaptic strengthening. The same may be true in an individual exhibiting very low arousal. Furthermore, the amount of cortical activation caused by a stimulus depends on whether it is being attended (Heinze et al. 1994; Mangun 1995). Therefore, to make it possible for correlated stimuli to induce synaptic learning, sufficient arousal and attention to these stimuli appear necessary, and synaptic learning may depend on how much attention is being di-

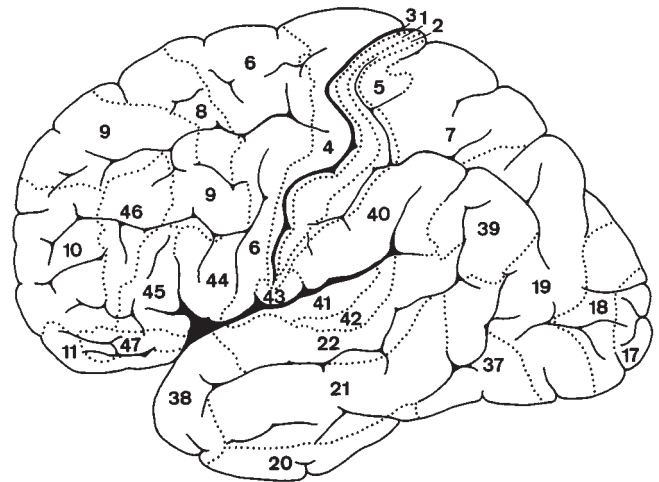


Figure 2. Lateral view of the left cortical hemisphere. Brodmann's (1909) areas are indicated. (Adopted from Pulvermüller & Preissl 1991.)

rected to relevant stimuli. In the following considerations it will be tacitly assumed that correlated stimuli receive a sufficient amount of attention from the learning individual to allow long-lasting changes of synaptic connections to occur.

#### 3.1. Assemblies representing word forms

Turning to language, it appears relevant to ask where in the cortex correlated neuronal activity occurs during verbal activities at early ontogenetic stages, when language learning takes place (Pulvermüller 1992; Pulvermüller & Schumann 1994). The infant's repeated articulations of syllables during the babbling phase are controlled by neuronal activity in inferior motor, premotor, and prefrontal cortices (Brodmann areas 4, 6, 44, 45). One may well envisage that one specific synfire chain controls the articulation of a given syllable and thus represents its articulatory program (Braitenberg & Pulvermüller 1992). In addition to and simultaneous with cortical activity related to motor programs, specific neurons in the auditory system are stimulated by the sounds produced during articulation (Braitenberg & Schüz 1992; Fry 1966). These neurons are localized in primary and higher-order auditory cortices (superior temporal lobe; Brodmann areas 41, 42, and 22). Furthermore, somatosensory self-stimulation during articulatory movements evokes activity in somatosensory cortices (inferior parietal lobe; areas 1–3 and 40). Therefore, neuronal activity can be assumed to be present almost simultaneously in defined primary and higher-order motor and sensory (auditory and somatosensory) cortices. All of these areas are within the first gyrus surrounding the sylvian fissure, the so-called perisylvian cortex (Bogen & Bogen 1976). Neuroanatomical evidence from monkeys suggests that the perisylvian areas are strongly and reciprocally connected, whereby long-distance connections between areas anterior to motor, adjacent to primary auditory, and posterior to primary somatosensory cortex are particularly relevant (Deacon 1992a; Pandya & Yeterian 1985; Young et al. 1995). Given that necessary long-distance connections are available, it follows by learning rule 1' (see also Table 2) that the coactivated neurons in the perisylvian areas develop into cell assemblies (Braitenberg 1980; Braitenberg & Pulvermüller

1992; Braitenberg & Schüz 1992; Pulvermüller 1992). Figure 3 represents an attempt to sketch such a perisylvian assembly. The individual circles in this diagram are thought to represent local clusters of strongly connected neurons. On the psychological level, the network may be considered the organic counterpart of a syllable frequently produced during babbling, or as the embodiment of the phonological form of a word acquired later during language acquisition.

The Hebbian framework implies that different gestalts and word forms have distinct cortical assemblies, because perception of these entities will activate different but possibly overlapping populations of neurons. If a language is not learned through the vocal and auditory modalities, but through the manual and visual modalities (sign languages), cortical localization of cell assemblies representing meaningful elements should be different. Because gestures are performed with both head and hands and perceived through the eyes, they are related to neuronal activity farther away from the sylvian fissure (more superior motor cortices and occipital visual cortices). Thus, it must be assumed that meaningful gestures included in sign languages involve these extra-perisylvian visual, motor, and association cortices (see Pulvermüller 1992 for further discussion).

In assuming cell assemblies distributed over perisylvian cortices, the Hebbian perspective is in apparent contrast to older localizationist models according to which motor and acoustic representations of words are stored separately in Broca's (areas 44 and 45) and Wernicke's regions (posterior part of area 22), respectively (Geschwind 1970; Lichtheim 1885; Wernicke 1874). The Hebbian view implies that the motor and acoustic representations of a word form are not separate, but that they are strongly connected so that they form a distributed functional unit. For this unit to function properly, both motor and acoustic parts need to be intact. This is important for the explanation of aphasias, in particular of the fact that in the majority of cases these organic language disturbances affect all modalities through which language is being transmitted. Whereas localizationist models have great difficulty explaining this (see, e.g., Lichtheim 1885 for discussion), a cell assembly model can account for the multimodality of most aphasias.<sup>3</sup> Furthermore, the as-

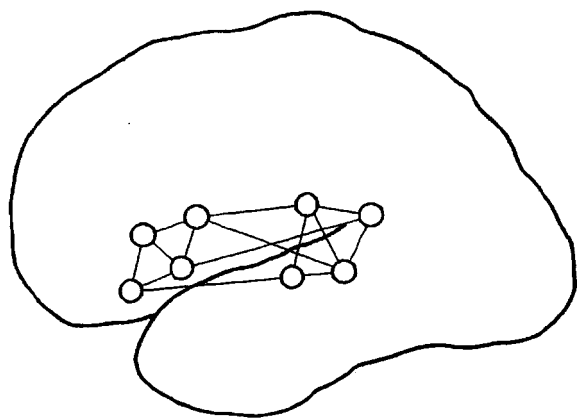


Figure 3. The cell assembly representing a phonological word form may be distributed over perisylvian areas. Circles represent local neuron clusters and lines represent reciprocal connections between such clusters. The connections are assumed to have strengthened because of correlated activity of neurons during articulation of the word form.

sumption that word form representations are distributed over inferior frontal and superior temporal areas receives support from imaging studies revealing simultaneous activation of both language areas when words or word-like elements are being perceived (Fiez et al. 1996; Mazoyer et al. 1993; Zatorre et al. 1992).

### 3.2. Cortical lateralization

From the Hebbian viewpoint, localization of language mechanisms is determined by associative learning and by the neuroanatomical and neurophysiological properties of the learning device (the cortex). The cortical loci where simultaneous activity occurs during motor performance and sensory stimulation follow from the wiring of efferent and afferent cortical connections, which are genetically determined. Genetic factors are also important for the formation of cortico-cortical fiber bundles, which are a necessary condition for long-distance association of coactivated neurons located in different areas. Furthermore, a pure associationist approach may have difficulty explaining why, in most right-handers, the left hemisphere – but not the right – is necessary for many aspects of language processing. Left hemispheric “language dominance” is evident from lesion studies in adults and in infants (Woods 1983) and from psychophysiological experiments in young children, demonstrating that stronger language-specific electrocortical activity can be recorded from the left hemisphere than from the right (Dehaene-Lambertz & Dehaene 1994; Molfese & Betz 1988). Neuroanatomical correlates of language laterality have been found in the size of perisylvian areas (Galaburda et al. 1978; 1991; Geschwind & Levitsky 1968; Steinmetz et al. 1990) and in size (Hayes & Lewis 1993), ordering (Seldon 1985), and dendritic arborization (Jacobs et al. 1993; Jacobs & Scheibel 1993; Scheibel et al. 1985) of pyramidal cells in the language areas. For differences in size of particular areas, epigenetic processes appear to be very important (Steinmetz et al. 1995). It is well known that differences in cell size and dendritic arborization may be influenced by sensory stimulation and motor output (Diamond 1990; Diamond et al. 1967) and, consistent with this view, language laterality has been proposed to be caused by environmental factors, such as lateralized auditory stimulation before birth (Previc 1991). Such stimulation may well underlie some of the morphological asymmetries mentioned. However, there are also arguments for a contribution of genetic factors to language lateralization (Annett 1979). At this point, it therefore appears safer not to dismiss a possible role of genetics here. For the Hebbian framework to operate, an anatomical substrate is necessary and this substrate is determined by genetic factors. Nevertheless, given the brain with its preprogrammed input and output pathways, its specific cortico-cortical projections, and its probably genetically determined left-hemispheric preference for language, the Hebbian approach leads to highly specific hypotheses about cortical distribution of language-related processing units.

One of these hypotheses concerns the cortical realization of laterality of language. According to localizationists, language processes take place in only one hemisphere. In contrast, the Hebbian framework suggests a different view. Although genetic and/or environmental factors lead to stronger language-related activation of left perisylvian cortex when language is being produced or perceived, articu-

lation of a word form is probably controlled by bi-hemispheric activity in motor regions, and acoustic perception of the word certainly leads to activation of bilateral auditory cortices. Because neurons in both hemispheres are coactivated when a word form is being produced or perceived, the cell assembly representing the word form should be distributed over bilateral perisylvian cortices (Mohr et al. 1994b; Pulvermüller & Mohr 1996; Pulvermüller & Schönle 1993). However, if the left hemisphere's neurons are more likely to respond to language stimuli and to control precisely timed articulations, cell assemblies representing word forms would be gradually lateralized to the left in the following sense: They include a large number of neurons in the left hemisphere and a smaller number of neurons in the right. According to this view, a lateralized cell assembly is not restricted to one hemisphere, but a greater percentage of its neurons would be in the "dominant" hemisphere and a smaller percentage in the "nondominant" hemisphere (Pulvermüller & Mohr 1996).

What would be the cause of this lateralization? Given that genetically programmed differences in the hemispheres' anatomical and physiological properties are the cause of lateralization of cognitive functions, it becomes important to develop ideas about how left/right differences in the "hardware" could influence the "software." Based on an extensive and profound review of neuroanatomical and neurophysiological asymmetries, Robert Miller (1987; 1996) recently proposed that axonal conduction times in the left hemisphere are slightly slower, on average, than those in the right hemisphere. According to Miller, this may lead to a bias in favor of the left hemisphere for storing short time delays, such as are important for distinguishing between certain phonemes (Liberman et al. 1967). For example, the probability of finding a neuron that responds specifically to a [p], but does not respond to a [b], may be greater in the left hemisphere than in the right, because neurons with slowly conducting axons that could be used as delay lines for hardwiring the long (>50 msec) voice onset time of the voiceless stop consonant would be more common in the left hemisphere. The availability of axons with particular conduction times may also be relevant for attributing additional distinctive features to acoustic input (Sussman 1988; 1989). If neurons sensitive to certain phonetic features have a higher probability of being housed in the left hemisphere, the neuron ensemble representing a phonological word form should finally be lateralized to the left. Although Miller's theory of cortical lateralization needs further support by empirical data, it clearly illustrates how hemispheric specialization at the cognitive and functional levels may arise from basic neuroanatomical and physiological differences between the hemispheres.

### 3.3. Word categories

Associative learning may not only be relevant for the cortical representation of word forms, it may also play an important role in the acquisition of word meanings. When the meaning of a concrete content word is being acquired, the learner may be exposed to stimuli of various modalities related to the word's meaning, or the learner may perform actions to which the word refers. Although such stimulus and response contingencies are certainly not sufficient for full acquisition of word meanings (Gleitman & Wanner 1982; Landau & Gleitman 1985) – they would not, for example,

allow the learner to distinguish between the morning and the evening star (Frege 1980) – they may nevertheless have important brain-internal consequences. From the Hebbian viewpoint, it is relevant that neurons related to a word form become active together with neurons related to perceptions and actions reflecting aspects of its meaning. If this coactivation happens frequently, it will change the assembly representing the word. Coactivated neurons in motor, visual, and other cortices and the perisylvian assembly representing the word form will develop into a higher-order assembly. A content word may thus be laid down in the cortex as an assembly including a phonological (perisylvian) and a semantic (mainly extra-perisylvian) part (Pulvermüller 1992).

After such an assembly has formed, the phonological signal will be sufficient for igniting the entire ensemble, including the semantic representation and, vice versa, the assembly may also become ignited by input only to its semantic part.<sup>4</sup> Thus, frequent co-occurrence and correlation of word form and meaning-related stimuli is only necessary at some point during the acquisition process. Later on, the strong connections within the higher-order assembly guarantee ignition of the entire assembly when part of it is being activated and, thus, they guarantee a high correlation of activity of all assembly parts, and, consequently, the endurance of the assembly.

When phonological word forms become meaningful, quite different cortical processes may take place, depending on what kind of information is being laid down in the associative network. Hebbian associationist logic suggests that cortical representations differ radically between words of different vocabulary types. In the following paragraphs, a few such differences will be discussed.

**3.3.1. Content and function words.** Neurons activated by stimuli related to the meaning of most concrete content words (nouns, adjectives, and verbs) are likely to be housed in both hemispheres. For example, the visual perceptions of objects that can be referred to as "mouse" will probably activate equal numbers of left- and right-hemispheric neurons because a corresponding visual stimulus is equally likely to be perceived in the right and left visual half-fields, and, in many cases, will be at fixation so that half of it is projected to the left visual field (right hemisphere) and the other half to the right visual field (left hemisphere). Therefore, if word form representations are strongly lateralized to the left, the assemblies representing content words (word form plus meaning) will be less strongly lateralized. Assemblies with different degrees of laterality are sketched in Figure 4.

In contrast to content words with concrete and well-imaginable meaning, function words such as pronouns, auxiliary verbs, conjunctions, and articles serve primarily a grammatical purpose. Many of them contribute significantly to the meaning of sentences, for example, "and," "or," "not," and "if." However, their meanings cannot be explained based on objects or actions to which the words refer. Rather, their meaning appears to be a more complex function of their use (Wittgenstein 1967) and can only be learned in highly variable linguistic and nonlinguistic contexts. Evidently, the correlation between the occurrence of a particular function word and certain stimuli or actions is low. Therefore, there is no reason why the perisylvian assembly representing the word form should incorporate additional neurons. If this is correct, assemblies representing



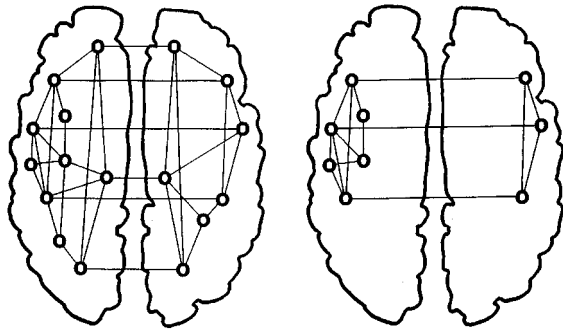


Figure 4. Cell assemblies relevant for cognitive processing may be distributed over both hemispheres and may be lateralized to different degrees. Whereas for cell assemblies representing phonological word forms and grammatical function words a high degree of laterality appears likely (right), an assembly representing a concrete content word may exhibit a reduced degree of laterality (left). (Adopted from Pulvermüller & Mohr 1996.)

function words remain limited to the perisylvian cortex and strongly left-lateralized in typical right-handers.

Note that this argument depends on the formulation of the cortical learning rule. If coincidence of neuronal activity was the factor causing synaptic modification, function words should have widely distributed cell assemblies because these words occur in a multitude of stimulus constellations and, in addition, they occur much more frequently than most content words (Francis & Kucera 1982; Ortmann 1975). When a function word (e.g., the article “the”) is being learned, it may be used with various content words (“the cat,” “the dog,” “the horse”) and, if there is a systematic relationship between the use of the content words and the occurrence of nonlinguistic stimuli (e.g., animal pictures), there will be a strong coincidence between the occurrences of each of these nonlinguistic stimuli and the word form. If only coincidence learning took place, cell assemblies representing function words should include even more neurons in visual cortices than most content word assemblies, because the assembly representing the function word would incorporate all neurons related to coincident visual nonlinguistic stimuli. However, because connections weaken if only pre- or only postsynaptic neurons fire (Table 2), the relatively infrequent co-occurrence of the function word with each of the visual stimuli will guarantee that its assembly does not become associated with representations of either visual stimulus. Correlation of neuronal activity is important for synaptic strengthening in the cortex, and this implies that function words are represented in cell assemblies restricted to perisylvian areas, or, at least, that they do not include large numbers of neurons outside.

**3.3.2. Abstract content words.** One may argue that the postulated difference in semantic meaning between content and function words does not apply for all members of these vocabulary classes. Rather, it appears that there is a continuum of meaning complexity between the “simple” concrete content words that have clearly defined entities they can refer to (so-called referents), more abstract items that may or may not be used to refer to objects and actions, and function words that cannot be used to refer to objects. It is therefore inappropriate to make a binary distinction between vocabulary classes based on semantic criteria. If

semantic criteria are crucial for intracortical representation, the suggested gradual differences in the correlation between word form and meaning-related stimuli or actions should be reflected in gradual differences in cortical lateralization and how assemblies are distributed. An abstract content word, such as “philosophy,” may therefore have an assembly somewhat in-between typical content and function word assemblies: It may exhibit an intermediate degree of laterality consisting mainly of perisylvian neurons, but including a few neuron clusters outside perisylvian areas.

Among the abstract content words are words referring to emotional states, for example “anger” and “joy.” For these words, it is not difficult to find characteristic visual stimuli related to their meaning – for example, angry or joyful faces. In addition, there are characteristic meaning-related patterns of muscle activity – namely, the contraction of the respective facial muscles – and autonomic nervous system activity (Ekman et al. 1983; Levenson et al. 1990). It should therefore be noted that, although these words do not refer to objects and actions in the sense in which the word “house” refers to an object, the likely co-occurrence of patterns of muscle contractions with the word forms may nevertheless lead to the formation of widely distributed cortical cell assemblies representing these words. In addition to cortical neurons added to the word form representations during learning, it has been proposed that these assemblies acquire additional links to subcortical neurons in structures of the limbic system related to emotional states (Pulvermüller & Schumann 1994). “Emotion words” may therefore be represented by a cortical assembly plus a limbic assembly-tail. The amygdala and the frontal septum may be the most important structures for linking the cortical assembly to its subcortical tail (Schumann 1990; 1997).

These considerations should make it clear that the degree of abstractness of an item is not the only factor influencing assembly topographies. According to the present proposal, the important criterion is the strength of the correlation between the occurrences of a given word form and a class of nonlinguistic stimuli or actions. In the clear cases, this likelihood is related to abstractness, but there are exceptions.

**3.3.3. Action words, perception words, and other word classes.** Content words are used to refer to odors, tastes, somatic sensations, sounds, visual perceptions, and motor activities. During language learning, word forms are frequently produced when stimuli to which the words refer are perceived or actions to which they refer are carried out by the infant. If the cortex is an associative memory, the modalities and processing channels through which meaning-related information is being transmitted must be important for formation of cortical assemblies. This has inspired models of word processing in the brain postulating distinct cortical representations for word classes that can be distinguished based on semantic criteria (Warrington & McCarthy 1987; Warrington & Shallice 1984).

If the modality through which meaning-related information is transmitted determines the cortical distribution of cell assemblies, a fundamental distinction between action and perception words can be made. Action words would refer to movements of one’s own body and would thus be used frequently when such actions are being performed. In this case, a perisylvian assembly representing the word form would become linked to neurons in motor, premotor, and

prefrontal cortices related to motor programs. Perception words, whose meaning can best be explained using prototypical stimuli, would consist of a perisylvian assembly plus neurons in posterior cortex. In many cases, visual stimuli are involved and the respective word category may therefore be labelled vision words. Assemblies representing words of this category would be distributed over perisylvian and visual cortices in parietal, temporal, and/or occipital lobes. Figure 5 presents sketches of the assembly types postulated for action and vision words. Examples of words whose meanings are related to the visual modality are concrete nouns with well-imaginable referents, such as animal names. The best examples of action words are in the category of action verbs.

This model draws too simple a picture of the relation between word forms and their meanings, because it does not explain homonymy (Bierwisch 1982; Miller 1991). If a phonological word form has two exclusive meanings – if it can, for example, be used as a noun with one meaning or as a verb with another meaning (the/to beat) – a mechanism must be assumed that realizes the exclusive-or relationship between the two meanings. As suggested earlier, homonyms could be represented by overlapping cell assemblies,

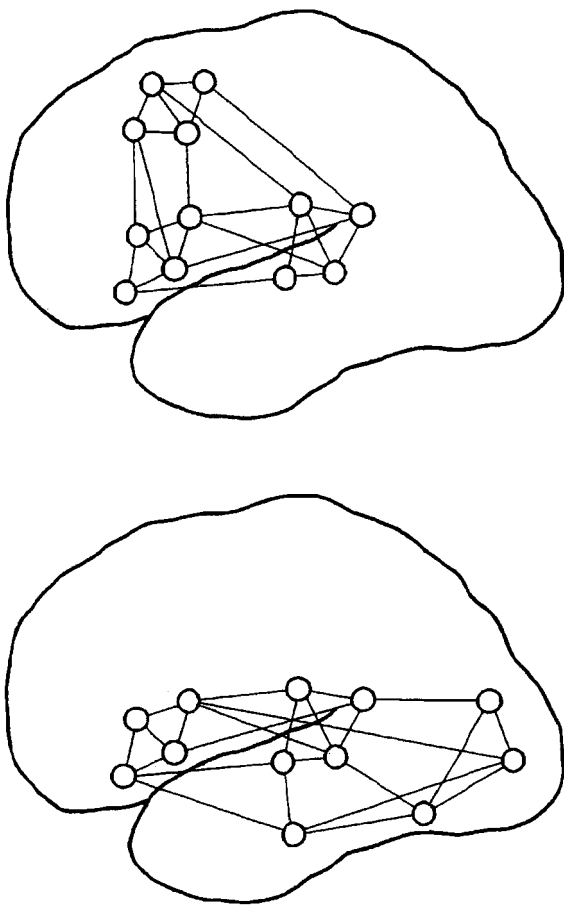


Figure 5. Whereas words eliciting strong visual associations (“vision words”) may be organized in assemblies distributed over perisylvian and additional visual cortices, words that remind one of movements of one’s own body (“action words”) may be organized in assemblies distributed over perisylvian and additional motor cortices. Many (but not all) concrete nouns are vision words and many verbs are action words.

that is, by two content word assemblies sharing one perisylvian phonological part. Inhibition between the semantic assembly parts is unlikely to be wired in cortex, because the percentage of cortical inhibitory neurons is low and these neurons are usually small (Braitenberg & Schüz 1991). Intracortical inhibitors would therefore be unlikely candidates for mediating inhibition between cortical areas – for example, between assembly parts in frontal and occipital lobes. However, such mutual inhibition between overlapping assemblies could be realized by striatal connections (Miller & Wickens 1991). Accordingly, homonymic content words may be realized as widely distributed assemblies sharing their perisylvian part while inhibiting each other through striatal connections. This wiring would allow the perisylvian word form representation to become active together with only one of its “semantic” assembly parts (see Pulvermüller 1992 for further discussion).<sup>5</sup>

The argument made above for action and visually-related words can be extended to words referring to stimuli perceived through other modalities. For those, additional word categories – odor, taste, pain, touch, and sound words – can be postulated. Members of these word classes should be represented in assemblies with specific cortical topographies. For example, whereas an assembly representing a pain or touch word may include substantial numbers of neurons in somatosensory cortices, sound words may have exceptionally high numbers of neurons in bilateral auditory cortices included in their assemblies. Again, it must be stressed that neurons responding to stimuli of various modalities and neurons controlling body movements and actions are located in both hemispheres. It is for this reason that cell assemblies representing these words are assumed to be distributed over both hemispheres and to be less strongly lateralized compared to assemblies representing function words (Pulvermüller & Mohr 1996).

The definition of action words is particularly delicate because not all action-related associations involve the motor modality. Here it is important to distinguish movements that are performed by the subject’s own body from movements that are only perceived visually. “To fly” or “the plane,” for example, are words that are frequently heard by children when they perceive certain moving visual stimuli. Although a relation of visual stimuli to the motor modality can hardly be denied – because perception of visual stimuli is usually accompanied by eye movements related to neuronal activity in frontal eye fields – this eye movement-related neuronal activity is probably not very stimulus-specific (similar saccades are made when different objects are looked at). Therefore, the correlation between visual input patterns and the occurrence of the word forms “fly” or “plane” may be highest and these words may thus be organized in assemblies including a significant number of neurons in visual cortices responding to specific moving contours. These words should therefore be classified not as action words but as visually-related words of a certain kind (as words referring to visually perceived movements). On the other hand, action words as defined above, that is, words usually referring to movements of one’s own body, may include movement detectors in visual cortices in their assemblies. Many body movements are visually perceived when they are performed, suggesting that sensory-motor assemblies are established for representing these actions – an idea for which there is ample support from recent studies (Fadiga et al. 1995; Gallese et al. 1996; Rizzolatti et al.

1996). These considerations indicate that Figure 5 draws too crude a picture of cell assemblies representing action words. Such assemblies can include additional neurons in visual cortices primarily processing movement information – many of which are probably located in the posterior part of the middle temporal gyrus (Watson et al 1993; Zeki et al. 1991). A similar point can be made for somatosensory stimulations caused by body movements, suggesting that neurons in parietal cortices may be added to the assembly representing an action word, as well.

Further word class-distinctions can be made based on the cortical areas active during meaning-related motor activity. Different kinds of action words can be distinguished considering the muscles most relevant for performing the actions (to chew, to write, to kick), the complexity of the movement (to knock, to write), and the number of muscles involved (to nod, to embrace). These factors may “shift” the neurons in frontal lobes added to the perisylvian assembly in the inferior/posterior (mouth/hand/foot representation) or anterior/posterior direction (complex/simple movements), or enlarge/reduce their cortical distribution (many/a few muscles involved in movement).

Similar, more fine-grained distinctions are desirable for visually-related words. Some vision words refer to static objects (house), others to moving objects (train), some refer to colors or colored objects (iguana), others to objects lacking colors (penguin). Furthermore, some visual stimuli are very simple (line), others are more complex (square, cube, house, town, megalopolis). This suggests that different sets of neurons are being added to the assembly when contingencies between words and different kinds of visual stimuli are being learned. The assembly of a word used to refer to colors or colored objects may include neurons maximally responding to color, and, as discussed above, neurons sensitive to moving visual stimuli may be included in the assemblies representing words referring to such stimuli. Recently, cortical processing streams have been discovered in temporal lobes that are primarily concerned with movement or color information from the visual input (Corbetta et al. 1990; Watson et al. 1993; Zeki et al. 1991). If movement-detecting cells are more frequent in one area, for example in the posterior middle temporal gyrus, and neurons in primary and secondary visual cortex that respond to color preferentially project to other areas, for example in the inferior temporal lobe, this would suggest that words referring to colors or colored objects are realized as assemblies including additional neurons in color areas (e.g., in the inferior temporal gyrus), and that words referring to visually perceived movements have assemblies that comprise additional neurons in visual movement areas (in the middle temporal gyrus).

It is important to stress that (1) word types defined in this way<sup>6</sup> do not necessarily have a congruent lexical category; most – but not all – verbs are action words, and there may be action words from other lexical categories; and (2) it is not always clear from theoretical consideration to which category a particular word should be assigned. Most concrete content words probably exhibit a high correlation with stimuli of more than one modality, and their presentation may therefore remind subjects of multimodal stimuli. Whereas verbs referring to body movements are probably action words, and concrete nouns (such as animal names) are almost certainly related to vision, other word groups – for example, nouns referring to tools – probably lead to

both visual and motor associations. Therefore, when evaluating the present ideas about word class-differences related to word meaning in neuroscientific experiments, it is most important to assess quantitatively semantic associations elicited by word stimuli.

#### 4. Cortical activation during word processing: Predictions and methodological remarks

Cognitive brain theories lead to empirical predictions in psychophysiological studies. Testing such predictions is not trivial, however. In the case of language, it is particularly difficult to design experiments and interpret their results because there are so many possible confounds to which, for example, a physiological processing difference between two stimulus words could be attributed. Furthermore, the subtraction logic used in many imaging studies of cognitive processes has frequently been criticized, and one may prefer designs that could prove more useful in testing precise predictions on cognitive processes of comparable complexities.

After summarizing selected predictions derived from the Hebbian model (sect. 4.1), the subtraction logic underlying many imaging studies will be contrasted to what will be called the double dissociation approach to neuroimaging (sect. 4.2), and, finally, methodological issues specific to the investigation of word processing will be addressed (sect. 4.3).

##### 4.1. Predictions about where and how

Hebbian logic suggests that content and function words, and words referring to actions and perceptions, have different neurobiological counterparts. The cell assemblies representing these lexical elements may differ with regard to their laterality and cortical topography. Whereas all assemblies representing words are assumed to include a strongly lateralized perisylvian part, neurons outside perisylvian language areas (and in both hemispheres) would only be added to the assembly if words refer to actions and perceivable objects. If assembly topographies are a function of semantic word properties, signs of cortical activity should differ when these different assemblies are being activated.<sup>7</sup> Based on these ideas, one would expect:

1. function words to evoke strongly left-lateralized signs of cortical activity restricted to perisylvian cortices,
2. content words to evoke less lateralized signs of cortical activity in perisylvian areas and outside,
3. action words to evoke additional activity signs in motor cortices of frontal lobes,<sup>8</sup> and
4. visually-related words to evoke additional activity signs in visual cortices of occipital and inferior temporal lobes.

These are some of the predictions obvious from the above considerations (sect. 3) that relate to the *where* question. When the assumptions leading to these predictions were discussed in section 3, the *why* question was traced back, in each case, to a Hebbian learning rule postulating that correlated neuronal activity is the driving force of assembly formation. With regard to the *how* question, it is important to recall that cell assemblies were assumed to exhibit two functional states, namely, ignition (or full activation) and reverberation (or sustained partial activity). When outlining

empirical tests of the cell assembly framework and its application to language, one may not only be interested in testing predictions about assembly topographies, but one may also want to think about how to distinguish and detect possible physiological signs of ignition and reverberation. As detailed in section 2, ignition may be reflected in a sudden spreading of neuronal activity shortly after stimulation, and reverberation would follow ignition and could become visible in high-frequency brain responses. Therefore, the following additional predictions are possible:

5. shortly after stimulation, signs of cell assembly ignition are simultaneously present at the cortical loci where the assembly is located, and

6. after a longer delay, signs of reverberation emerge in the same areas.

It is not possible to deduce the exact point in time when these putative physiological processes take place. However, because words are recognized rather quickly – for example, lexical decisions, that is, judgments on letter strings according to whether they are real words or not, can be made as early as  $\frac{1}{2}$  second after the onset of written stimuli – it is clear that the postulated physiological process of cell assembly activation must take place during the first few hundreds of milliseconds after the stimulus has been presented.

Although numerous additional predictions can be derived from the discussion in section 3, sections 5 and 6 will focus on hypotheses 1–6. These hypotheses will be discussed based on the results of psychophysiological and neuroimaging experiments.

#### **4.2. Subtractions versus double dissociations in psychophysiology**

In psychophysiology, numerous neuroimaging techniques are available for investigating higher cognitive processes. Activity of large neuron ensembles can be visualized using electrophysiological recording techniques, such as electroencephalography (EEG) and magnetoencephalography (MEG). These techniques provide exact information about temporal dynamics of electrophysiological activation and deactivation processes that occur in the millisecond range. They also allow for localization of sources, although such localization is usually much less precise than imaging of brain metabolism. Metabolic imaging techniques with high spatial resolution, such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), are extremely valuable for localizing brain structures that maximally become active, thereby increasing their metabolic rates during cognitive tasks. However, the metabolic methods give only a rough picture of temporal dynamics of brain processes, and it is therefore important to use both electrophysiological and metabolic imaging techniques when investigating brain processes of cognitive functions.

It is necessary to recall that important information about where, why, and how cognitive processes take place in the human brain was obtained before modern imaging techniques were available. Most of these studies used the individuals' behavior as the dependent measure. In addition, studies of neurological patients with focal lesions can answer the question of which brain structures are necessary for particular cognitive operations (Jackson 1878; 1879). Studies of healthy individuals in whom stimulus information reaches only one hemisphere – for example, using

the technique of lateralized tachistoscopic presentation of visual stimuli – can provide important insights into the hemispheres' roles in language processing (Hellige 1993; Pulvermüller & Mohr 1996). Together with such neuropsychological evidence, modern neuroimaging and psychophysiological data can provide even stronger conclusions about language mechanisms in the human brain (Posner & Raichle 1994).

In recent years, a large number of imaging studies of word processing have been carried out, many of which are relevant for evaluating the Hebbian model outlined above. When interpreting these results, it is necessary to consider basic methodological issues. Giving an overview of all possible methodological problems that may become relevant is outside the scope of the present article (see, e.g., Posner & Raichle 1995 and comments therein). Rather, two important points will be mentioned briefly, the so-called subtraction logic and the question of stimulus matching, which are both crucial for investigating word class-differences.

Various dependent measures recorded by large-scale imaging techniques are usually interpreted as signs of cortical activity. However, the exact mechanisms by which an increase in cortical activation (i.e., the frequency of excitatory postsynaptic potentials in a set of neurons) may lead to an increase in the CO<sub>2</sub> concentration in numerous blood vessels, to an increase in intracellular glucose levels, to an enhancement of biomagnetic signals, or to a more positive or negative event-related brain potential are not sufficiently understood to make quantitative predictions possible. For example, one may predict that higher glucose metabolism or event-related potential amplitudes are present in or close to the inferior prefrontal cortex during processing of a given word class, but quantification of the expected difference, in terms of microvolts, for example, would not be possible. Ultimately, even the rationale underlying the more/less logic may be flawed, because an increase in biomagnetic activity or enhancement of cortical metabolism may be caused by the activation of inhibitory neurons (Mitzdorf 1985; Posner & Raichle 1995). Nevertheless, at least in the cortex, excitatory neurons represent the majority (>85% of cortical neurons are excitatory), and they are, on average, much larger than inhibitory neurons (Braitenberg & Schüz 1998). Furthermore, their function is probably to control excitatory activity in cortex, rather than to process more specific information. It is therefore possible, but not likely, that an enhancement of large-scale measures of cortical activity exclusively reflects inhibitory processes on the neuronal level. (This may be more likely for structures with high percentages of inhibitory neurons, such as the striatum.) Therefore, in the majority of cases, it appears reasonable to use large-scale neuroimaging measures to draw conclusions on activity changes in large numbers of excitatory neurons in the cortex.

The logic underlying all imaging work is that a dependent measure indicates a difference in brain activity between two conditions. In most cases, a critical condition is compared to a baseline or control condition. In the simplest case, looking at an empty computer screen or at a fixation cross may be compared to reading words or to making lexical decisions on these stimuli. Using a more complex design, the task of silently reading a word may be compared to the generation of a verb that somehow relates to the meaning of a displayed word. If an area of cortex is found to “light up” in such an experiment, one can conclude that



the perceptual, cognitive, or motor operations induced by the two conditions differ with regard to neuronal activity in this particular area.

Unfortunately, however, in many experiments there are several differences between critical and control conditions. For example, the tasks of looking at an empty screen and of making lexical decisions about words appearing on the screen differ with regard to several aspects: (1) perceptions – either a word or nothing is being perceived; (2) higher cognitive processes – the stimulus has to be classified as a real word or as a meaningless element, or nothing has to be done; and (3) motor activities – a button press is either required or not. In addition, silently reading a noun (e.g., cow) and silently generating a word that refers to an activity related to the object to which the noun refers (e.g., to milk, to buy) involve quite different cognitive processes. Although identical words may be displayed in the two conditions and no overt response may be required, the two conditions differ because only one of them requires strong attention and involves search processes, semantic inferences, repeated lexical access, and so on (see also the discussion in Posner & Raichle 1995). Finally, another difference between the reading and the generation tasks is that only in the latter are verbs involved (but nouns are being read in both conditions). Given that an area is found to “light up” in the generation condition if compared to the reading condition, it is not clear which of the many different cognitive processes relates to the difference in brain activity. The difference may even be used to evaluate prediction 3 (sect. 4.1) because action verbs are relevant in only one of the conditions, but, of course, if the prediction is met, the experimental result would not provide strong support for it because of the many confounds.

A solution to the problem may lie in a more careful selection of the conditions and stimuli that are being compared. If, for example, silently reading words is compared to reading random letter strings made up of the same letters, one may argue that in this case the critical and control conditions differ only with regard to well-defined linguistic processes, such as word form identification and processing of semantic information. However, the objection can be raised that processing of words is not even necessary under such conditions because random letter strings can frequently be distinguished from real words merely by looking at the first three letters of the items and deciding whether these letters can be combined according to the phonological or orthographic rules of the language from which the real words are taken. Thus, word processing could be avoided by experiment participants in these conditions. To allow conclusions on processes specific for words, even more similarity between the stimulus classes should be required. For example, only letter strings that are in accord with the phonological rules of the language could be allowed as pseudowords, and lexical decisions could be required so that experiment participants would be forced to attend to and process the stimuli. In this case, a neuroimaging difference between conditions could be attributed to the difference between word and pseudoword processing, although from a psycholinguistic perspective these processes may differ under various aspects (including word form identification, semantic processes, and the use of a “time out” strategy for rejecting pseudowords; Grainger & Jacobs 1996; Jacobs & Grainger 1994; Mohr et al. 1994b). Nevertheless, a difference in brain activity between these

conditions would allow stronger conclusions on the cortical processes induced by the words.

In many cases, two conditions are being compared in which condition 1 is considered to induce a subset of the processes induced in condition 2. The subtraction of the brain responses would then be interpreted as reflecting the psychological processes that condition 2 exhibits but condition 1 lacks. Subtractions can be performed repeatedly, so that a hierarchy of conditions corresponds to a set of subtractions (Posner & Raichle 1995). However, the principal problems remain, namely, (I) that a difference in more than one psychological process may be attributed to each pair of conditions, making it difficult to attribute a physiological contrast to one of them, and (II) that statistical criteria for the comparison of two conditions are difficult to choose if multiple pairs of physiological data are compared. If many comparisons are being made (when data from tens of channels or thousands of voxels are contrasted), the likelihood of a difference occurring by chance is high. On the other hand, if critical significance levels are adjusted to reduce the likelihood of significant results (e.g., by following Bonferroni logic), an actual difference between brain responses in two conditions may be masked because the too rigid statistical criterion is almost impossible to reach (Wise et al. 1991).

The only way to avoid problem (I) appears to be to choose maximally similar experimental conditions. To investigate word class-specific processes, a good option appears to be a comparison of two psycholinguistically similar stimulus classes while the experimental task is kept constant in conditions 1 and 2. To reduce the risk of obtaining by-chance results with standard significance criteria (II), more risky predictions can be derived and tested. One way to do this is to predict interactions between topographical variables and stimulus classes, rather than only more or less activity at a not-yet-specified locus. In the best case, condition 1 and condition 2 would induce quite similar cognitive processes, but condition 1 would induce a process not induced by 2, and, conversely, condition 2 would induce a specific process not induced by 1. Based on theoretical predictions, processing of stimuli of class 1 in the task chosen may then be assumed to activate a set A of cortical loci not activated by class 2, whereas stimuli of class 2 processed in the same task would be assumed to activate a different set B of areas not activated by 1. (Of course there may be additional areas C activated by both classes.) The brain areas activated by the two conditions or stimulus types would be distinct, and each set of areas would include loci not included in the other. This can be called a *physiological double dissociation*. The prediction to be tested by analysis of variance would be that direct comparison of the two activity patterns leads to a significant interaction of the task variable with the topography variable. It is unlikely that such a prediction is being verified by chance in a neuroimaging experiment, in particular if the loci where differences are actually found have been specified before the experiment based on theoretical considerations. The rationale underlying this is very similar to the logic used in neuropsychology, where double dissociations are taken as strong evidence for processing differences (Shallice 1988; 1989), although the dependent measure is behavioral in neuropsychology, but physiological in psychophysiology.

In summary, one perspective on overcoming some of the problems of a simple subtraction logic in neuroimaging ex-

periments is offered by a double dissociation approach to psychophysiology. In this approach, physiological signs induced by maximally similar tasks – or even patterns of brain activation caused by matched stimuli in the same task – are being compared, and the prediction would be that class 1 of stimuli activates cortical loci A more strongly than class 2, whereas class 2 induces stronger activity signs than class 1 at distinct loci B. With regard to the present discussion, classes 1 and 2 may represent different word categories – for example, action and visually-related words – and loci A and B would then be large sets of cortical areas – for example, motor versus visual cortices.

### **4.3. Word properties affecting brain processes**

Given that comparable stimulus materials are used in an imaging experiment on processing differences between word classes, the expectation would be that defined cortical areas “light up” when members of a given word class are being processed (see predictions 1–4). But what would “comparable” mean in this case? Behavioral studies in which response times and accuracies of responses were measured precisely have clearly shown that various properties of stimuli influence information processing in the brain, and many of the results from behavioral studies could be confirmed by psychophysiological experiments. Imaging techniques with good spatial resolution have only been used for a few years and, therefore, many methodological studies on the influence of stimulus properties have not yet been performed using these techniques. When evaluating imaging studies of word processing, it is essential to keep in mind the stimulus properties for which behavioral and earlier psychophysiological studies have demonstrated strong effects on brain processes.

Words can vary on various scales. The naive observation that long words are more difficult to read than short ones is paralleled in the observation that words of different length elicit different electrocortical responses measured in the EEG. This appears to be the case regardless of whether the items are presented acoustically (Woodward et al. 1990) or visually (Kaufman 1994). A second important factor influencing behavioral and physiological responses to words is whether they are common or exceptional. In contrast to pictures or real objects for which it is difficult to estimate whether they are frequently or rarely being perceived, the frequency of words can be exactly determined based on the evaluation of large corpora of spoken or written text. Word frequency is well known to have a strong influence on response times and accuracies of word processing (see, e.g., Bradley 1978; Mohr et al. 1996). In addition, word frequency has a strong influence on cortical potentials evoked by word presentation (Polich & Donchin 1988; Rugg 1990; Rugg & Doyle 1992). Because certain word classes exhibit enormous differences in word frequencies, this variable may affect the outcome of studies of word class-differences. For example, whereas most function words are in the highest frequency range, only a small percentage of the content words can be found in this high range, and most content words are used only rarely. Thus, word frequency is a likely confounding factor of experimental results about differences between word classes.

Additional possible confounds of word category differences are related to psychological processes induced by the stimuli. Some words are more arousing than others: The

word “spider” may lead to much more pronounced brain activity in an arachnophobic patient compared to “beetle,” and normal individuals may exhibit similar differences in brain responses. That event-related potentials reliably differ between more or less arousing words has been shown by numerous studies (Chapman et al. 1980; Johnston et al. 1986; Naumann et al. 1992; Williamson et al. 1991), and there is also evidence that a variable called “valence,” that is, the degree to which the stimulus is evaluated as positive or negative, can have an effect on event-related potentials. Therefore, there is some reason to believe that what has been called the “affective meaning” of words (Osgood et al. 1975) can influence the brain processes these stimuli induce. Stimulus matching for the variables’ valence and arousal therefore appears desirable – except, of course, if the role of these variables in word processing is the subject of the experiment.

Another variable strongly affecting behavioral and physiological responses to word stimuli is the context in which they are being presented. There are different types of context effects. They can be elicited not only if words are presented in well-formed or ill-formed sentences, but also when words are presented one by one. If a word occurs twice in the same experiment, event-related potentials are usually more positive-going for the second occurrence (see, e.g., Rugg 1985; Smith & Halgren 1987). The repetition effect appears to be quite complex and can interact with other variables, for example word frequency (Rugg 1990). Therefore, if a physiological difference is observed between words of different frequencies that are repeatedly presented in the same experiment, it cannot be decided to which variable the difference should be attributed.

Context effects can also occur between different words that are semantically related (semantic priming). Presentation of a prime word changes electrocortical signs of activity elicited by a subsequently presented target that is semantically related to the prime (Holcomb & Neville 1990; Nobre & McCarthy 1994; Rugg 1985). Similar priming effects may also occur when a word is being presented in sentence context. A pronounced negative deflection is seen when meaningful words appear at the end of a sentence where they are highly uncommon (Kutas & Hillyard 1980a), and different brain waves have been identified that may indicate different forms of syntactic or semantic violations (Neville et al. 1991; Osterhout & Holcomb 1992). Although there are several different effects of sentence context on word-evoked potentials, at least one of these effects appears to be quite similar to the effect induced by semantic priming (Van Petten 1993). Most importantly, context effects are not necessarily the same for all word classes (Besson et al. 1992). As mentioned above for the effects of word frequency and word repetition, sentence context effects may vary between word classes as well. Event-related potentials elicited by content words are attenuated by a sentence context, provided that semantic and syntactic restrictions are met by the sentence. In contrast, function words also show attenuation of event-related potentials when presented in semantically deviant strings that still preserve some basic sentence-like structure (Van Petten & Kutas 1991). If words are presented in sentences or in sentence-like word strings, it may well be that not only the effect of a stimulus word is seen in the neurophysiological response, but a complex blend of the effects of the critical word, its preceding words, and their semantic and syntactic rela-

tions. The various context effects may therefore either artificially produce word class-difference, or they may mask real processing differences between word classes.

When brain processes distinguishing between word classes are investigated, it appears necessary to keep in mind these effects of word length, word frequency, emotional (arousal and valence) properties of the stimuli, as well as those of word repetition, priming, and syntactic and semantic sentence context. These properties of word stimuli and strings may confound results of any imaging study investigating differences in brain activity evoked by two word groups. Only if such confounds are excluded can a strong conclusion on differences between lexical or semantic word categories be drawn.<sup>9</sup>

## 5. Brain activity during word processing: Where?

In this section, studies on the cortical areas activated during word processing will be discussed. The main question will be whether there is evidence for or against predictions 1–4. Studies on differences between content and function words will be dealt with in section 5.1, and section 5.2 will be concerned with action and visually-related words and related categories.

### 5.1. Content and function words

Neuropsychological work clearly indicates that different brain areas are necessary for processing content and function words. Whereas aphasic patients with anomia have difficulty finding content words (Benson 1979), for patients with agrammatic aphasia function words are more difficult to produce (Caramazza & Berndt 1985; Pick 1913). In addition, aspects of agrammatics' deficit in language comprehension can be explained based on the assumption that they have a selective deficit in processing these lexical items (Pulvermüller 1995a). Lesions within the entirety of the perisylvian region can be the cause of the agrammatic language disturbance (Vanier & Caplan 1990). In contrast, lesions at various cortical sites outside left-hemispheric perisylvian cortices can lead to selective impairment in using or comprehending word categories included in the content word vocabulary (see the discussion in sect. 5.2). If function word representations are assumed to be restricted to perisylvian cortices (see Fig. 3), and content word representations are assumed to be more widely distributed (see examples in Fig. 5), a perisylvian lesion will destroy a large percentage of neurons included in function word representations, but will only remove a smaller part of the representations of content words. In contrast, lesions outside the perisylvian region will only affect representations of content words. Thus, different cortical distributions of cell assemblies representing content and function words can account for the double dissociation in processing content and function words in specific aphasic impairments such as agrammatism and anomia (Pulvermüller 1995a; Pulvermüller & Preissl 1991).

In addition, evidence from behavioral experiments in healthy individuals using lateralized tachistoscopic presentation have provided further support for processing differences between content and function words. It is well known that words presented either in the left visual hemifield (and, thus, to the right hemisphere) or in the right visual hemi-

field (to the left hemisphere) of right-handed individuals exhibit a processing advantage after presentation in the right visual field ("right visual field advantage"; see, e.g., Bradley 1978). In behavioral experiments, these effects can be quantified exactly in terms of response times and accuracies. A frequently applied paradigm is lexical decision, where words and matched meaningless pseudowords are presented in random order and study participants have to indicate whether an item is a legal word or not. In lexical decision experiments, the "right visual field advantage" has been found to be stronger for function words compared to content words matched for word frequency and length (Chiarello & Nuding 1987; Mohr et al. 1994b). For function words, direct stimulation of the left hemisphere leads to faster or more accurate responses compared to stimulation of the right hemisphere. This is consistent with the idea that cell assemblies representing function words are strongly lateralized to the left (sect. 3.3.1). The weaker or even absent right visual field advantage for content words supports the idea that cell assemblies underlying content word processing are less lateralized (Mohr et al. 1994b).

Several studies investigating event-related potentials (ERPs) have been conducted in search of differential brain activity induced by content and function words. Garnsey's (1985) early experiment revealed a fine-grained word class-difference in event-related potentials uncovered by principal component analysis. Neville et al. (1992) presented content and function words in sentence context and had subjects indicate whether the sentences made sense or not. Words of the two classes were not matched for word length or frequency. These authors reported a left-lateralized component evoked by function words which peaked at 280 msec after stimulus onset, whereas a peak more symmetrical over the hemispheres was evoked by content words at 350 msec. A similar result was obtained by Nobre and McCarthy (1994), who used stimuli matched for word length but not for word frequency. These authors presented words one by one and their subjects studied the sequence while trying to detect words of a particular semantic class. Again, a left-lateralized negative peak followed function word presentation (latency: 288 msec), whereas content words led to an enhanced negativity (latency: 364 msec) that was more symmetrical over the hemispheres. Gevins et al. (1995) used a cued two-stimulus paradigm and asked subjects to indicate whether two stimuli were similar according to phonological, syntactic, or semantic criteria. These authors reported a lateralized positivity (latency: 445 msec) elicited by function words which was most pronounced over left frontal regions, whereas content words failed to elicit a late lateralized component. These authors did not report stimulus lengths or frequencies, however, and it is therefore not possible to exclude the most likely confounds. In an experiment comparing brain responses to content and function words matched for word frequency and word length (Pulvermüller et al. 1995a) while study participants had to make speeded lexical decisions, a negative-going wave that peaked around 160 msec after the onset of visual stimuli revealed a significant interaction of the word class and hemisphere factors. The peak in the event-related potential was equally visible over both hemispheres after presentation of content words, but it was pronounced over the left hemisphere and reduced over the right when function words were processed. Mean event-related potentials obtained between 150 and 300 msec after stimulus onset also re-

vealed a significant interaction of the factors hemisphere and word class (left/right difference strong for function words, but minimal or absent for content words).

It is important to point out some of the differences between these studies. For example, the tasks to be performed by participants differed (lexical decision, sentence judgment, etc.). In spite of these differences, all of these experiments revealed differences in electrocortical responses between the major vocabulary types. Results were very similar in the studies by Neville et al. and by Nobre and McCarthy. In both cases an early left-lateralized component was found after function words and a component symmetrical over the hemispheres followed content words after a longer delay. In Gevins et al.'s results, function words led to a left-lateralized component that occurred much later compared to both earlier studies, and, again, no such lateralized component was present for content words. In our study, we found no word class-difference in latencies of event-related potentials, but this study again confirmed the observation of a left-lateralized component evoked by function words and a component symmetrical over the hemispheres evoked by content words. Thus, all of these studies agreed on the finding of left-lateralized electrocortical responses to function word presentation and less or even absent lateralization of potentials evoked by content words.

Checking these studies against possible confounds reveals the following: Words were presented in sentence context only in Neville's experiment, whereas context effects are likely to play a minor role in the remaining studies. Matching of stimuli for word length was performed for Nobre and McCarthy's and for Pulvermüller et al.'s experiment. Only the latter study used content and function words matched for word frequency. As already pointed out in section 4.3, the issue of frequency matching is of particular relevance for electrocortical content/function word differences, because there are data (reported by King & Kutas 1995) indicating that latency differences may be the result of different word frequencies of the stimuli chosen from the two vocabulary classes. After frequency-matching of stimuli, word class-differences in latencies of event-related potentials indeed vanished. However, the differences in laterality of electrocortical responses to content and function words were still present with frequency-matched stimuli (Pulvermüller et al. 1995a). Therefore, the difference in laterality – rather than the difference in latency – appears to be characteristic of the major word classes.<sup>10</sup>

These studies are consistent with predictions 1 and 2 proposed in section 4.1. A possible explanation for the differences in cortical laterality of brain responses to content and function words is that specific cortical representations of these stimuli have different degrees of laterality. At present, there is no strong evidence from neuroimaging that content and function word representations are differently distributed within each hemisphere, although neuropsychological data support this view (Pulvermüller et al. 1996c; Vanier & Caplan 1990). However, recent preliminary PET data indicate that this prediction may also be correct (Nobre et al. 1997).

The Hebbian viewpoint suggests that differences in cortical loci involved in representing and processing words depend on semantic word properties. However, the summarized studies do not include information about which of the many properties distinguishing content and function words are crucial for differential brain activation induced by these

stimuli. Content and function words not only differ with regard to semantic criteria (e.g., only the former can be used to refer to objects and actions), they also belong to different lexical categories, and even their phonological structure may be different. To find out whether semantic factors are indeed crucial, it is necessary to compare words that share phonological and lexical properties and differ only in their meaning. In a study comparing nouns with concrete and abstract meaning, electrocortical responses were also found to be different over the hemispheres (Kounios & Holcomb 1994). Abstract nouns led to an interhemispheric difference in electrocortical activity, whereas concrete nouns evoked similar responses over both hemispheres. This is consistent with the assumption that semantic differences underlie differential laterality of event-related potentials to concrete and abstract nouns. One may argue that this result makes it plausible that the same is true for the difference between content and function words, although this suggestion cannot be proven to be correct at present. Consistent with this view, however, the high degree of abstractness of function words is paralleled by a strong interhemispheric difference in event-related potentials, and the smaller degree of abstractness of abstract nouns is paralleled by a weaker interhemispheric difference evoked in a lexical decision task.<sup>11</sup> This pattern of results is in agreement with the assumption of strongly lateralized cell assemblies representing function words, weakly lateralized assemblies representing concrete content words, and a moderate degree of laterality for assemblies representing abstract content words (see sect. 3.3.2). Therefore, the view that the degree of laterality of brain responses to words reflects semantic stimulus properties receives support from the summarized psychophysiological studies.

## 5.2. Words related to action and vision

If the cortical distribution of word representations is determined by the cortical pathways through which meaning-related information is being transmitted, differences in cortical localization should not only distinguish representations of content and function words, but, in addition, words that differ in their motor and visual associations, such as nouns and verbs or animal and tool names, should have cell assemblies with different cortical topographies. The Hebbian model, and probably any associationist approach, suggests that semantic word class-differences determine differences in cortical representations. Most importantly, however, based on a Hebbian associationist model the semantic differences between word categories can be used to generate predictions on cortical areas that are involved in processing words of such categories. As discussed in section 3.3.3, we can expect words eliciting strong visual associations to be represented and processed in perisylvian and additional visual cortices in inferior temporal and occipital areas, whereas words with strong motor associations would be expected to involve additional motor areas in the frontal lobe. Concrete nouns referring to animals or large man-made objects appear to be examples of typical vision words, verbs referring to actions usually performed by humans are probably typical action words, and words referring to tools may evoke both strong motor and visual associations.

Neuropsychological data clearly indicate that focal brain lesions can affect these word categories to different degrees. Whereas lesions in temporal and/or occipital regions



sometimes selectively impair processing of nouns, lesions in frontal areas have been reported to be associated with deficits in processing verbs (Damasio & Tranel 1993; Daniele et al. 1994; Goodglass et al. 1966; Miceli et al. 1984). There is also evidence for more fine-grained disturbances primarily affecting, for example, words referring to small man-made objects, such as tools, or words referring to living entities, such as animals (Damasio et al. 1996; Warrington & McCarthy 1983; 1987; Warrington & Shallice 1984). The relationship between anatomical lesion site and category-specific deficit has not yet been investigated systematically for all cortical lobes. However, studies of lesions in the left temporal lobe indicate that damage to the middle part of the inferior temporal gyrus most strongly impairs naming of animals whereas more posterior lesions involving inferior and middle temporal gyri result in a more pronounced deficit in naming tools (Damasio et al. 1996). The idea that cell assemblies representing words of different semantic and lexical categories have different cortical distributions therefore receives support from neuropsychological research, although it is not yet clear whether all of the more exact predictions on the cortical loci involved can be verified.

Imaging work that might reveal clues about processing differences between nouns and verbs was frequently carried out after Petersen et al. (1989) and Wise et al. (1991) reported that verb generation involved cortical areas less activated during noun reading. These authors and several more recent investigations used PET to measure brain activity while experiment participants either read visually presented nouns (reading task) or tried to generate verbs that “go with” the nouns (verb generation task).<sup>12</sup> If “car” is being presented, generation of “drive” or “race” may be expected. For evaluation, brain activity maps from the reading task were subtracted from those from the verb generation task. Significantly enhanced brain metabolism in a particular area during the generation task was attributed to cognitive processes necessary for verb generation and not necessary for reading nouns.

Although not all of the studies agree on the cortical loci of activity enhancement during verb generation, it appears that increased blood flow in prefrontal and temporal cortices can be observed.<sup>13</sup> Activity enhancement in the left frontal lobe has been reported in Broca’s area and anterior and superior to it (McCarthy et al. 1993; Petersen et al. 1989). Also Wernicke’s region (posterior area 22; Wise et al. 1991) and the middle temporal gyrus (Fiez et al. 1996) showed increased blood flow. Thus, during verb generation stronger activity in perisylvian language cortices and in additional premotor, prefrontal, and temporal areas was found. Figure 6 presents results from one study revealing both prefrontal and middle temporal activation during verb generation relative to the reading condition.

When interpreting these results to draw conclusions on cognitive processes, such as processing of a particular class of words, the following should be noted. As the above example clearly demonstrates (“car” leading to generation of “drive” or “race”), the generated words are not necessarily verbs, in particular if the experimental language is English where many verbs can also be used as nouns and vice versa. From this point of view, it does not seem appropriate to call it a “verb generation task,” but rather a task to generate action words. However, even this may not be correct, because subjects may have been instructed to describe “what the

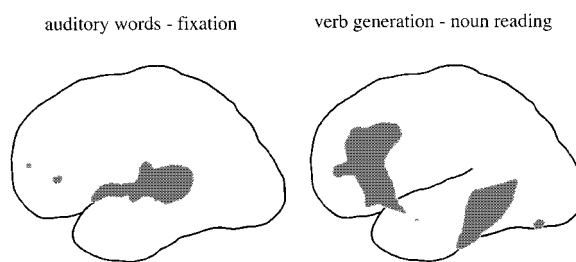


Figure 6. Presentation of spoken words leads to increased blood flow in perisylvian areas relative to looking at a fixation point (left diagram). In contrast, verb generation can activate additional prefrontal areas and the middle temporal gyrus (right diagram). (Modified from Fiez et al. 1996.)

nouns might be used for or what they might do” (Fiez et al. 1996, p. 1), thus allowing for the generation of both action words and vision words related to perceived movements. In addition, arguments raised in section 4.2 become relevant here, namely, that it is difficult to interpret these results in psychological terms because comparison of word generation to the reading task reveals several differences on the cognitive level. Recall that the generation of action words makes not only semantic processes necessary, it also requires, for example, lexical search and stronger attention compared to the highly automatized process of reading common words. Furthermore, in most cases no information about stimulus or response properties is given that would allow for evaluation of possible confounds as pointed out in section 4.3. Based on these PET results alone, it is therefore not possible to attribute blood-flow changes to verb or action word processing. Nevertheless, assuming that action words were frequently produced by experiment participants, these results appear consistent with the following view. During the generation of action words, an additional cell assembly was activated (compared to the reading task) that included neurons not only in perisylvian cortices but also in prefrontal, premotor, and middle temporal areas. This is probably not too far from what could be expected based on the associationist framework discussed in section 3 (see also prediction 3 above). However, from a methodological more rigorous point of view it appears necessary to compare brain activity when action and vision words are being processed in the same task (see sect. 4).

In a recent PET study, Martin et al. (1995) presented achromatic line drawings of objects and had subjects generate action names and color words associated with the objects. Direct comparison of activity patterns evoked during generation of these word categories revealed increased metabolic rates in the ventral temporal lobe when color words were generated. In contrast, generation of action words led to stronger activity in more superior temporal areas on the middle temporal gyrus, and in inferior frontal areas, but not in additional motor cortices. This failure may reflect the fact that, as these authors emphasize (footnote 26, p. 105), many of the words actually generated by experimental subjects did not refer to movements the subjects would perform themselves, but rather to movements of objects that are perceived visually. Examples of responses listed by these authors include the verbs “fly,” “see,” and “sleep,” for which visual associations are plausible, but a classification as action words may appear inappropriate. If many verbs without motor associations were produced, this

may be the reason why visual areas were activated instead of additional motor areas relevant for controlling hand or foot movements. This point further evidences the necessity to control both stimulus and response properties carefully. It is important to note, however, that part of the left middle temporal gyrus was active during verb generation in the study by Martin et al. and that either the same or a closely adjacent area has been found to be active during verb generation from visually presented nouns (Fiez et al. 1996; Petersen et al. 1989).

Differences between action and vision words were also investigated using event-related potentials calculated from EEG recordings. Most of these studies compared electrocortical responses to nouns and verbs. Whereas an early study (Samar & Berent 1986) reported generally more positive potentials following verbs (compared to nouns), more recent work using larger electrode arrays (32 or 64 channels) and more sophisticated analysis techniques (e.g., current source density analysis) suggest word class-differences in cortical topographies of event-related potentials. In a study investigating potentials evoked by several word classes, Dehaene (1995) presented numerals, nouns (animals' and persons' names), and verbs matched for word length.<sup>14</sup> Whereas word-evoked potentials were generally larger over the left hemisphere, word class-differences were discovered over both hemispheres around 300 msec after stimulus onset (see p. 2155). Verbs elicited a left-lateralized positive component maximal over inferior frontal cortical sites which was not found for nouns. Both nouns referring to animals and verbs led to almost identical left-temporal negativities. These results are consistent with the assumption of additional left-frontal activity during processing of verbs, but do not indicate any noun/verb processing differences in more posterior cortical loci. With regard to the methods, however, it should be noted that no matching for word frequency or arousal and valence values was performed for nouns and verbs, one third of the verb stimuli had homophonous common nouns, and stimuli were repeated in the experiment. The first point makes a replication with matched stimulus materials desirable.

Presenting nouns and verbs matched for word frequency, length, arousal, and valence in a lexical decision task, Preissl and colleagues (Preissl et al. 1995; Pulvermüller 1996a) found electrocortical differences as early as 200 msec after the onset of visual stimuli. When average noun- and verb-evoked potentials (between 200 and 230 msec) were compared, significant differences were seen only over the frontal cortex. After submission of data to current source density analysis in order to maximize the contribution of local generators to the signal (Hjorth 1975; Perrin et al. 1989), stronger electrocortical signs of activity were found after verb presentation over bilateral motor cortices, but more pronounced event-related potentials over visual cortices in the occipital lobes were seen after nouns. It is significant that stimuli were carefully evaluated for motor and visual associations. Ratings of experiment participants confirmed differences in associations of body movements and visual scenes elicited by stimulus words. Verbs were judged to elicit significantly stronger motor associations than nouns, and nouns were judged to elicit stronger visual associations than verbs.<sup>15</sup> The electrocortical differences seen over motor and visual cortices paralleled these differences in conscious motor and visual associations. The left diagram in Figure 9 presents these differences in event-related poten-

tial topographies elicited by well-matched nouns and verbs. These data are in agreement with predictions 3 and 4 listed in section 4.1. They can be explained by the assumption that action words activate additional neuronal generators close to motor cortices, whereas vision words spark additional neuron populations in or close to primary visual areas in the occipital lobes.

It could be argued that although an influence of the confounding factors discussed in section 4.3 appears unlikely in this case, it is not clear whether the electrocortical word class-differences are related to semantic associations elicited by the stimuli, or to the fact that stimuli belong to different lexical categories (noun and verb). However, because the assumption that semantic differences are crucial can explain the topographical differences found in electrocortical responses, this view should probably be preferred. Differential involvement of motor and visual cortices could be predicted based on associationist principles. In contrast, there is no a priori reason why members of different lexical categories should involve different cortical lobes. However, to further confirm the idea that semantic properties of words, not their lexical categories, are crucial for differences in the topography of cortical activation, it is appropriate to look at stimuli from the same lexical category (nouns) that nevertheless evoke either primarily visual associations (e.g., animal names) or associations of body movements (e.g., tool names).<sup>16</sup>

Recently, Damasio and colleagues (1996) examined differences in brain activity during naming of animals and tools. In a PET study investigating activity changes in the temporal lobes, they found strong activation of the middle part of the left inferior temporal gyrus during animal naming (compared to a baseline condition), whereas enhancement of activity in more posterior cortices in the inferior and middle temporal gyri were found when naming of tools was compared to the baseline. These results suggest that different neuronal populations and cortical areas in the left temporal lobe contribute to processing of action words compared to words with additional visual associations.

Differences in brain activation during naming of tools and animals were also investigated in a PET study by Martin and colleagues (1996). In this case, subjects had to silently name objects depicted either in line drawing or in silhouette (to eliminate differences in internal detail of drawings).<sup>17</sup> The names of these objects were matched for word frequency. Direct statistical comparison of activity patterns elicited by animal and tool naming revealed the following. Animal naming led to relatively enhanced blood flow in primary and higher visual cortices in the calcarine sulcus in the left hemisphere (and to small activity foci in the prefrontal lobe). In contrast, tool naming was accompanied by activity enhancement in left premotor areas, plus an activity increase in the middle temporal gyrus. These data provide additional evidence that areas outside the perisylvian cortices contribute to processing of animal and tool names. Consistent with earlier studies using the verb generation task, a cortical locus in the left middle temporal gyrus was activated when words with strong motor associations (tool names, action verbs) were generated. In contrast to the results of the Damasio study, activity enhancement during animal naming involved occipital visual cortices rather than inferior temporal sites (which is consistent with prediction 4). Most importantly, however, naming of tools led to an additional activity focus in the premotor area con-

trolling hand movements (Fig. 7). This is consistent with the assumption that processing of words with motor associations activates motor cortices involved in programming such movements.

Although this study has several methodological advantages over other PET studies (e.g., matching of stimuli, of responses, calculation of significant differences between critical conditions rather than only between critical condition and baseline), it should be kept in mind that a naming study was carried out and differences between naming conditions may be related to several cognitive processes. Looking at the list of methodological desiderata from section 4, it is relevant that for most PET studies it is not clear whether and to what degree complexity, frequency, arousal or valence values, and repetition of stimuli or responses influenced the results.<sup>18</sup> Furthermore, when naming of depicted animals and tools is being compared, it must be noted that whereas animal pictures include many curved lines, are usually rather complex, and can include various colors or shadings, tools can be drawn with a few straight lines and usually lack extensive coloring or shading. If matching of visual stimuli for visual complexity has not been performed, physical differences of stimuli may account for differential activation of visual pathways specialized for processing of particular aspects of stimuli.

The possible merit of exact investigation of psychological properties of stimuli and responses can be further illustrated based on results from the Damasio study mentioned above (Damasio et al. 1996). In that investigation, highest activation values during naming of famous persons' faces were observed in the temporal poles of both hemispheres. It is unclear to which psychological variable this activity enhancement relates. However, it is clear from psychophysiological investigations that faces are among the most arousing stimuli (Lang 1979; Lang et al. 1990), and words referring to such stimuli are very likely to exhibit comparatively high arousal values, as well. It has been proposed that high-arousal words (i.e., words evoking strong emotional as-

sociations) are represented in cell assemblies that include additional neurons in the amygdala and subcortical structures (e.g., midbrain dopamine system; Pulvermüller & Schumann 1994; Schumann 1990).<sup>19</sup> This provides a tentative explanation why Damasio and colleagues found enhanced activity in temporal poles during naming of famous persons. When persons' names were retrieved, it may be that cell assemblies including large numbers of amygdala neurons became active, and, therefore, blood flow increases were found in adjacent cortical areas strongly connected to the amygdalae (see Amaral et al. 1992), that is, in temporal poles. Thus, differential arousal values of words and/or pictorial stimuli may explain differential involvement of temporal poles during naming of pictures of famous persons.<sup>20</sup>

In summary, these studies include the following results relevant to the idea of different cortical representation and processing of action and vision words:

1. PET and fMRI studies using the verb generation task revealed enhanced activity in perisylvian language areas and adjacent temporal and prefrontal cortices in the left hemisphere. Perisylvian activity enhancement may be accounted for by assuming that an additional word form representation is being activated in the generation task (relative to the baseline, usually noun reading). Activation of additional cortical areas outside the perisylvian region may indicate psychological processes coupled to word form processing. Whereas prefrontal activity increases dorsal to Broca's area may relate to body movements to which the words refer, activity enhancement in middle temporal gyrus may be related to visual imagination of movements.

2. ERP studies indicate that nouns with strong visual associations and verbs with strong motor associations activate different cortical generators in both hemispheres. Whereas stronger signs of electrocortical activity following action verbs have been recorded from anterior and central regions, nouns led to more pronounced activity signs over occipital visual cortices. These differences appear to be related to neuronal activity in or close to primary motor or visual cortices underlying movement and visual associations, respectively.

3. PET studies of animal and tool naming provide additional evidence for processing differences between action and vision words. Tool naming with nouns that probably elicit motor associations activated premotor cortices and additional sites in the middle temporal gyrus, and naming animals using visual nouns led to activity enhancement in inferior temporal cortices and in occipital cortices close to the primary visual area.

Although these studies are subject to methodological problems to different degrees (as pointed out in great detail above), a coherent picture can nevertheless be drawn on their basis. Both ERP and PET studies support a contribution of occipital areas close to primary visual cortices to the processing of vision words. There is also PET evidence for a specific contribution of inferior temporal cortices to the processing of vision words. Thus, the processing of words that evoke visual associations appears to be related to neuronal activity in visual cortices (see postulate 3).

In contrast, processing words eliciting motor associations such as certain action verbs and tool names activates areas in the frontal lobe close to motor cortices, as revealed by ERP, PET, and fMRI studies (see postulate 4). PET studies revealed an additional area of activation in the middle

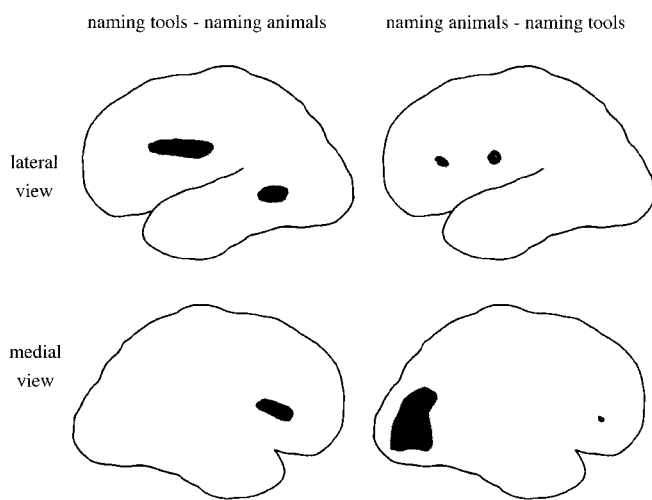


Figure 7. Naming of tools and animals can activate quite different cortical regions. Diagrams indicate significant differences in brain metabolism between animal and tool naming. In the left occipital cortex, enhanced metabolism is seen during animal naming. In the left premotor cortex close to the hand representation and in the middle temporal gyrus activation is stronger when tools are being named. (Modified from Martin et al. 1996.)

temporal gyrus which may be related to visual imagination of movements induced by action words. This focus may be localized in such a way that its activity is difficult to record in the EEG, or that even higher spatial resolution and, therefore, larger numbers of electrodes are necessary to monitor it. Thus, the data are consistent with the view that the processing of words that remind experiment participants of movements of their own bodies and the visual perceptions associated with these movements activate cortical loci in frontal lobes and temporal lobes related to the perception and imagination of such movements.

From a methodological point of view, it appears particularly important to stress that, in some cases, not only were significant differences in the activation of critical areas found when word processing was compared to a resting condition, but physiological double dissociations (see sect. 4.2) between the processing of action and vision words were discovered and substantiated by interactions of the factors topography and vocabulary type. These interactions could even be obtained for action and vision words well matched for various psychological variables specified in section 4.3.

Although there is considerable agreement between electrophysiological and metabolic studies of action and vision words, there is one major incompatibility that calls for brief discussion. Whereas, in most cases, PET studies revealed processing differences between word class-differences only in the left hemisphere,<sup>21</sup> ERP studies also indicated word class-differences in the right hemisphere. This was not only so in ERP studies of noun and verb processing, it was also true for most of the studies revealing electrocortical differences between content and function words (see sect. 4.2). It is possible that some of the differences recorded over the right hemisphere were actually caused by generators in the left hemisphere. For example, larger, right-sided, event-related potentials over the occipital cortex related to noun processing (Pulvermüller 1996a) may be caused by left-hemispheric neuronal generators within the interhemispheric sulcus (which can be activated during processing of visual nouns, as suggested by the study of Martin et al. 1996; see Fig. 7). However, at this point it appears unlikely that all electrocortical differences seen at recording sites over the right hemispheres are caused by left-hemispheric generators (see discussion in sect. 5.1). Lack of right-hemispheric word class-differences reported in PET studies does not prove that the null hypothesis of equal activity patterns is correct. Fine-grained differences may be ruled out by too rigid a criterion of significance (Wise et al. 1991). Furthermore, some metabolic imaging studies provided direct support for processing differences between word classes in the right hemisphere (footnote 21). Thus, a final decision regarding the right hemisphere's role in word class-specific processes must be left for future investigation.

These results are consistent with the Hebbian postulate of different cortical distributions of widespread cortical assemblies related to cognitive processing. In addition to neurons in the perisylvian language areas of Broca and Wernicke, assemblies representing action words may comprise neurons in motor, premotor, and prefrontal cortices and in middle temporal gyrus.<sup>22</sup> Vision words may be organized as assemblies distributed over perisylvian and additional higher-order cortices in temporal and occipital lobes, and even over primary visual cortices. Thus, postulates 3 and 4 formulated above receive support from the data discussed.

However, the Hebbian model implies that physiological word class-differences should be present in both hemispheres – a postulate that can, at present, be supported only by some of the studies.

## 6. Word processing in the brain: How?

The question of *where* in the brain cognitive processes take place is only one of the interesting issues to be addressed in cognitive neuroscience. Perhaps equally relevant are the questions of *how* the building blocks of cognitive operations – for example, meaningful words and gestalt-like figures<sup>23</sup> – are represented in the brain and in which way these representations are activated during processing of meaningful information. As discussed in section 2, the Hebbian framework provides tentative answers to both of these questions: The elements of representation are strongly connected but distributed populations of several thousand nerve cells. If such an assembly is being activated, an ignition takes place and, subsequently, neuronal activity will be retained in the assembly, leading to fast and coherent reverberation of activity in many of its neurons. As already mentioned, these ideas have received empirical support from investigations of high-frequency coherent brain activity and precisely timed spatiotemporal patterns of neuronal firing in cats, nonhuman primates, and humans. Stimulus-specific coherent and rhythmic activity in distant cortical neurons is usually seen in the high frequency range above 20 Hz. Spatiotemporal patterns are precisely timed with millisecond precision and reverberations may occur after short time lags of a few hundredths of a second. Reverberation of a cortical cell assembly may therefore lead to dynamics in high-frequency responses. If signals are large enough, they may well be picked up using large-scale neuroimaging techniques. It is clear that only EEG and MEG have a temporal resolution fine-grained enough to reveal spectral dynamics in high frequencies.

Starting from the working hypothesis that cell assemblies produce well-timed fast activity changes in many neurons, it appears most crucial to investigate high-frequency cortical responses in the EEG and MEG to further test the Hebbian model of word processing summarized above. Words – but not meaningless pseudowords – are assumed to be cortically represented as Hebbian cell assemblies. If high-frequency cortical activity reflects reverberation of neuronal activity in assemblies, such high-frequency responses during word processing should be stronger compared to processing of pseudowords. In addition, differences in topographies of high-frequency responses can be predicted for words with different meanings. Furthermore, because reverberatory activity can be assumed to occur only after ignition of cell assemblies, stimulus-related differences in high-frequency activity should occur only after differences observed using other measures of electrocortical activity, such as event-related potentials (predictions 5 and 6 in sect. 4.2).

Several experiments were carried out to test these hypotheses. In the earliest study, the EEG was recorded while subjects performed lexical decisions on German words and matched pronounceable and orthographically regular pseudowords displayed on a monitor. EEG data were submitted to a technique called Current Source Density Analysis to minimize the contribution of distant sources to the



signals and to maximize the contribution of generators close to the recording electrodes (Law et al. 1993; Perrin et al. 1989).<sup>24</sup> Spectral responses were obtained by filtering in various frequency bands. These spectral responses were rectified and averaged over trials. Results revealed a stimulus-specific difference in high-frequency responses around 30 Hz. Words evoked stronger responses in the band between 25 and 35 Hz compared to matched pseudowords. The difference was not only specific to this frequency band, it also showed topographic specificity in that it was visible only at recording sites over left perisylvian cortices, and it was only significant in the time interval 300–500 msec after stimulus onset (Lutzenberger et al. 1994; Pulvermüller et al. 1995b). No changes or differences between stimulus classes were visible in even higher parts of the spectrum, for example around 60 to 80 Hz where muscle activity produces most pronounced effects (Cacioppo et al. 1990).

Differences in high-frequency spectral responses between words and pseudowords were most pronounced and significant only at perisylvian recording sites. Because EEG signals were first submitted to Current Source Density Analysis (to emphasize the contribution of local generators), it appears that neuronal dynamics of large neuronal populations housed primarily in left-perisylvian cortices differed during word and pseudoword processing. The difference in spectral responses to words and pseudowords can be explained based on the assumption that additional high-frequency neuronal processes not involved after pseudoword presentation were sparked by word stimuli. Relatively strong 30 Hz activity during word processing may indicate continuous activity of large cortical assemblies producing well-timed and rapidly changing neuronal activity. Pseudowords – for which no cortical assemblies are assumed – would fail to ignite a specific neuronal representation, and, therefore, would finally cause less well-timed activity. These findings support predictions made based on the Hebbian model (Pulvermüller et al. 1994b).

It may be argued that the difference in high-frequency responses to words and pseudowords is related to the language used in the EEG experiment (German), to the EEG recording techniques, to features of the procedures used for signal analysis, to the modality of stimulus presentation (visual), or to the motor responses study participants had to perform to express their lexical decisions (yes/no button presses with the left hand; see Pulvermüller et al. 1995b). Therefore, an experiment was conducted in which all of these features of experimental setting and evaluation procedure were changed. A 74-channel biomagnetometer was used to record biomagnetic signals from both hemispheres of right-handed native speakers of English who heard English words and pseudowords spoken by a professional speaker. This time, subjects did not have to respond to the stimuli, but they were asked to memorize all stimuli in order to pass a recognition test later. For calculating spectral responses the method described by Makeig (1993) was used. For a variety of frequency bands, spectral power was determined in overlapping time windows of 0.3 seconds.

Figure 8 presents results obtained from one participant in the MEG experiment. In this subject, as well as in the group as a whole, words again evoked stronger high-frequency cortical responses than pseudowords (Pulvermüller et al. 1996a). Consistent with results from the EEG experiment, differential high-frequency responses could be ob-

served in the 20–35 Hz range. No significant difference in spectral power was seen in any of the other bands examined. The difference was only significant for recordings from anterior channels placed over inferior frontal areas of the left hemisphere. At these channels, biomagnetic responses evoked by words and pseudowords were larger compared to all other channels, so that a maximal signal-to-noise-ratio can be assumed. The consistency of results between EEG and MEG experiments indicates that the result does not depend on the language from which stimuli are taken, on special features of the methods for analyzing spectral responses, on whether high-frequency neuronal activity is recorded in the EEG or MEG, or on the task performed by experiment participants. It should be noted, however, that in a more recent study, Eulitz et al. (1996) found left-hemispheric differences between words and nonwords in the 60–70 Hz range when experiment participants had to decide whether words included nonlinguistic signals (incomplete letters, noise). Differences in the frequency range where high-frequency dynamics occur may therefore depend on the experimental task applied (see Pulvermüller et al. 1997 for further discussion).

Although high-frequency cortical responses differed between words and pseudowords as predicted based on the Hebbian notion of a cell assembly (see Fig. 1), one may wonder why this difference was reliable only in recordings from electrodes and coils placed close to left-perisylvian cortices. In section 3, cell assemblies representing words were assumed to be distributed over wide cortical areas and over both hemispheres. One may therefore ask how the present results of word/pseudoword differences in high-frequency responses recorded only from the left hemisphere would fit the model. The answer to this question is quite straightforward: Whereas all assemblies representing word forms are assumed to include a large percentage of their neurons in left-perisylvian areas, additional neurons outside these areas would be included in the assembly only if word forms frequently co-occur with multimodal nonlinguistic stimuli. These additional areas involved are likely to

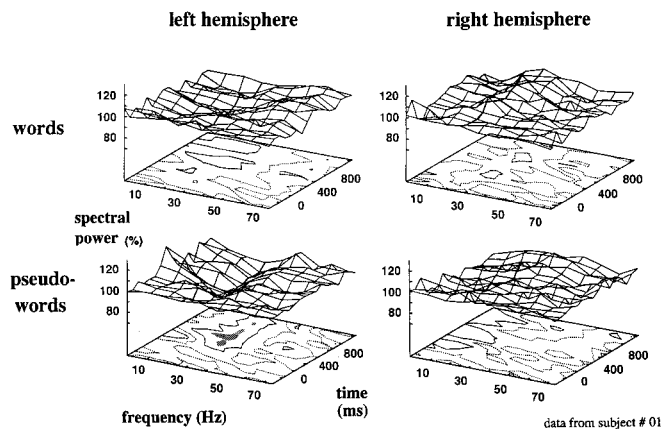


Figure 8. Spectral power calculated from biomagnetic signals elicited by words (upper diagrams) and pseudowords (lower diagrams) recorded over the left and right hemispheres. Normalized spectral power is plotted as a function of time and frequency. Word/pseudoword differences in high-frequency spectral responses are present around 30 Hz in recordings from the left hemisphere. (From Pulvermüller et al. 1996a.)

be different for words of different types. In the studies on word/pseudoword differences discussed above, stimuli of a narrow frequency range of word frequencies were used, but no restrictions regarding semantic properties were applied. The large variety in linguistic properties of stimulus words and, hence, in the additional cortical areas possibly involved may explain why there were no consistent word/pseudoword differences outside left-perisylvian areas.<sup>25</sup> This view is in need of further support, however. Namely, evidence should be collected concerning differential involvement of extra-perisylvian areas during processing of word categories using high-frequency spectral responses as the dependent measure.

As emphasized earlier, the Hebbian approach not only implies that words have cell assemblies and that pseudowords lack such cortical representations, it also leads to the postulate of different cortical distributions of cell assemblies representing action and vision words (Fig. 5). This predicts that topographies of high-frequency cortical responses vary as a function of word category. Words inducing visual associations should evoke stronger high-frequency responses over visual cortices, whereas action words with strong motor associations should lead to stronger high-frequency responses over motor cortices.

Nouns inducing strong visual associations and verbs evoking strong movement associations were presented in a lexical decision task while electrocortical responses were recorded in the EEG. Stimuli were matched for word frequency, length, arousal, and valence. EEG data were submitted to current source density analysis and to analysis of spectral power. A significant word category by recording site interaction was obtained in the analysis of spectral responses in the 30 Hz range 500 msec after stimulus onset and later (Pulvermüller et al. 1996b). There was a double dissociation of word categories and loci where stronger high-frequency signals were recorded. Differences were most pronounced at central and posterior recording sites. High-frequency responses to nouns at 30 Hz were stronger over visual cortices (recording sites O1 and O2 of the international 10-20 system; Jasper 1958), whereas 30 Hz responses to verbs were stronger close to motor cortices (recording sites C3 and C4). The diagram on the right in Figure 9 displays cortical topography of the difference in 30 Hz power elicited by action and vision words. No statistically significant differences between stimulus classes were seen in any of the other frequency bands analyzed. Because the topographies of the differences between action and vision words obtained with event-related potentials and with high-frequency responses as dependent measures are very similar (although not identical; see Fig. 9), it is possible that similar cortical generators underlie these electrocortical differences between the two word classes.

If a cell assembly is conceptualized as a neuron network that generates well-timed and fast-changing neuronal activity in many neurons, the data reported in this section provide evidence:

1. that cortical assemblies in left perisylvian cortex are being activated when meaningful words are processed but fail to ignite when meaningless pseudowords are being presented, and
2. that action and vision words activate cortical assemblies with different topographies.

Whereas most event-related potential studies summarized in section 4 indicate that word class-differences occur

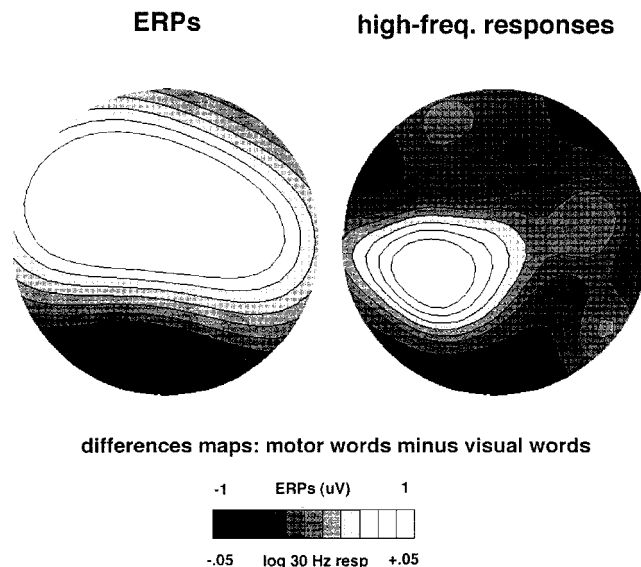


Figure 9. Processing of action verbs and nouns with strong visual associations is accompanied by significantly different electrocortical responses. Difference maps (nouns minus verbs) are shown. Large circles represent the head seen from the top (anterior is up, left is left). Differential topographies of event-related potentials (left diagram) are compared to evoked spectral responses in the frequency range 25–35 Hz. The two circles represent the head seen from above. The nose is up and left is left (right diagram). (Adapted from Pulvermüller et al. 1996b.)

quite early (150–300 msec) after onset of visually presented words, high-frequency spectral dynamics related to word properties were visible only at 300 msec and later. This is consistent with predictions 5 and 6 derived in section 4.1. Two processes, one early and the other late, may be reflected in these physiological measures. The Hebbian perspective offers the view that whereas the early process is the ignition of stimulus-specific assemblies, the late process is reverberation of neuronal activity in these assemblies. The fact that the early and late physiological signs exhibit the same double dissociation with regard to their cortical topography during processing of action and vision words (and even yielded comparable difference maps) may be taken as additional evidence that similar neuronal machineries – but different activity states of these machines – underlie the differences recorded in the EEG and MEG. One possibility is that these activity states reflect the psychological processes of word perception and of active memory of particular words, respectively.

## 7. Summary, implications, open questions

Neuroimaging studies were used to evaluate the hypothesis that words are cortically represented in distributed cell assemblies with defined topographies that vary with semantic word properties. Electrophysiological and metabolic imaging studies provide evidence that not only the language cortices in the left hemispheres, but additional cortical areas outside the left perisylvian areas play a role in word processing. Comparison of ERP responses to content and function words indicates that both hemispheres are strongly involved in processing concrete content words, whereas predominantly left-lateralized activity in or close

to perisylvian regions appears to be related to processing of highly abstract function words. PET, fMRI, and ERP studies revealed that cortical areas devoted to motor programming or visual perception are activated when words with strong motor or visual associations are being processed.

Based on the studies summarized, it appears likely that motor, premotor, and/or prefrontal cortices and possibly additional areas in middle temporal gyrus contribute to the processing of action words, whereas inferior temporal and/or occipital areas close to the primary visual cortex can be involved in processing words with visual associations. The observed physiological double dissociations provide additional support for the idea that semantic characteristics of words determine the loci of their cortical processing. It is not yet clear, however, to which degree primary, secondary, and higher-order association cortices, respectively, participate in word processing. The results (ERPs and PET) summarized here would suggest that both lower- and higher-order sensory and motor cortices, as well as multimodal association areas, can play a role. A further open question concerns the contribution of the right hemisphere to word class-specific processes. Most ERP and fMRI investigations advocate such right-hemispheric contributions, whereas most PET studies available at present do not.

The summarized data support a Hebbian model of word representation (cf. predictions 1–4 in sect. 4.1). Differential laterality of cortical processing of concrete content and abstract function words can be explained by bihemispheric cell assemblies with different degrees of laterality (Fig. 4). Differential involvement of motor and visual cortices in processing of action and vision words may indicate that representations of these words extend toward these extra-perisylvian areas, respectively (Fig. 5).

Although these results on differential cortical localization of word classes support the Hebbian model, they do not prove it to be correct. Obviously, the data cannot ultimately answer the question of whether there are word-specific cell assemblies. According to Hebb, each cognitive entity – each concept, word, or gestalt – has a separate assembly. An alternative appears to be that words of a particular semantic category are represented and processed by the same neurons. Processing of different words could, for example, be realized by different activity landscapes in the same neuron population, or, alternatively, by different spatiotemporal activity patterns occurring in the same neurons. Nevertheless, it may appear likely that words with similar but different meanings are organized in overlapping neuron sets, as proposed above. Some neurons of these sets may contribute to the processing of word 1 but not of word 2. If word 1 is “crocodile” and word 2 is “alligator,” for example, different color-sensitive neurons in visual cortices may be included in the respective assemblies. Furthermore, there is physiological evidence indicating that neurons can be found in human cortex that respond rather specifically to acoustic presentation of particular low-frequency words or phoneme sequences (Creutzfeldt et al. 1989a; 1989b). Whereas this observation fits into the Hebbian framework, it would be difficult to explain if different activity patterns of the same neuron set were the physiological counterparts of similar words.

It must again be emphasized that recent large-scale neuronal theories and their applications to language mechanisms (Braitenberg 1980; Braitenberg & Pulvermüller 1992; Damasio 1989a; Mesulam 1990; 1994) agree on the

postulate that cell assemblies with distinct topographies are the neuronal counterparts of words, and that perceptual properties of meaning-related stimuli determine cortical distributions of these assemblies. There are, however, diverging assumptions that distinguish proposals in the Hebbian tradition. For example, a contribution of the right hemisphere to language representation and processing is not assumed (or, at least, not explicitly postulated) by most of these proposals, but it is a necessary postulate according to the radical Hebbian perspective detailed in sections 2 and 3. Bihemispheric contributions to word processing are not only suggested by split-brain research (Zaidel 1976) and behavioral experiments in healthy individuals (Banich & Karol 1992; Mohr et al. 1994b), they are also consistent with recent metabolic imaging studies (Fiez et al. 1995; Mazoyer et al. 1993; Zatorre et al. 1992) and electrophysiological depth recordings (Creutzfeldt et al. 1989a; 1989b; Ojemann et al. 1988; Schwartz et al. 1996). In addition, differential laterality of electrocortical responses to words of different kinds and word-class processing-differences recorded from the right hemisphere appear to further support a radical Hebbian approach. At this point, however, one may consider it premature to reject any of the competing large-scale neuronal theories of word processing postulating that words are processed exclusively in the left hemisphere. More research is necessary to decide the controversial issue of right-hemispheric contribution to language and, in particular, to word class-specific processes.

Further open questions address the inner structure of cell assemblies. According to the proposal in section 3, most word representations consist of two parts, a perisylvian part related to the word form and a part located mainly outside the perisylvian areas representing semantic word properties. Similar proposals have earlier been formulated (Warrington & McCarthy 1987; Warrington & Shallice 1984). However, the assumption that word representations only include semantic and phonological parts may be questioned. It is evident that not only the semantic characteristics of a word and its phonological form need to be represented in the brain, but that information about its grammatical or functional properties needs to be stored as well (Garrett 1988; Levelt 1989). This information would, for example, include the knowledge about the lexical category of a word, about whether it is masculine or feminine, or transitive or intransitive, about its possible thematic roles in sentences, about whether it takes regular or irregular inflection, and, more generally, about the complements it requires. It has been argued that cortical representations of words may include a third part where such knowledge is laid down (Damasio et al. 1996). Although this is possible, it is difficult to see how formation of a separate third assembly part storing grammatical knowledge about a word can be explained by associative learning or by other biological principles. Based on Hebb’s ideas it appears more likely that grammatical knowledge is represented (1) in the connections between individual cell assemblies, (2) in the connections between overlap regions of several assemblies (Pulvermüller 1995b), and (3) in the activation dynamics that cell assemblies exhibit after their ignition. For example, it can be shown that between-assembly connections and activity dynamics are a possible basis of grammatical phenomena, such as center-embedding (see sect. 8). The representation of grammatical properties of words does not, therefore, require separate cortical neurons or areas

exclusively devoted to the storage of grammatical information (see Caramazza 1996 for further discussion). Nevertheless, the possibility that there are separate neuronal representations of grammatical features of lexical items is not ruled out by the data available at present.

The Hebbian framework has several implications regarding the loci of word representations in the brain (that is, regarding the where question). Only four of them have been evaluated here (see postulates 1–4, sect. 4.1). Additional predictions not evaluated here concern, for example, words for which meaning-related stimulus information enters the brain through channels other than the motor and visual modalities, and more fine-grained contributions of different parts of frontal or occipital cortex to processing of action and vision words (see sect. 3.3.3). With regard to action and vision words, however, at least an elementary answer to the where question appears to be possible based on the present data.

This answer may be considered preliminary, because many of the studies summarized are subject to methodological problems discussed in great detail in sections 4, 5, and 6. In some cases, tasks (such as verb generation or naming depicted objects) were chosen that may allow conclusions on word processing, but may be accounted for as well by attributing physiological changes to other cognitive processes (such as memory search or perceptual analysis). Furthermore, confounding factors related to psychological and physical properties of stimuli or responses (word length, frequency, arousal, valence, repetition, and context) may influence the results to different degrees. However, consistency of results obtained over different paradigms and with different recording methods appears to be a strong argument for their acceptability.

If the where-in-the-brain question can be answered in the case of content, function, action, and visually-related words, the Hebbian approach also provides a tentative answer to the question of why their different localizations develop and why they involve particular cortical lobes and gyri. It is unclear, however, from the localization studies dealt with in section 5 how the cortical representations are organized and how they function. To obtain information about such functional characteristics, fast changes of cortical activity need to be monitored. Cell assemblies widely distributed over distant cortical regions are probably difficult to observe through electrophysiological recordings from local neuron clusters or small areas. If large-scale neuronal theories of cognitive functions are correct, fast, large-scale recording techniques, such as EEG and MEG, are necessary to visualize activity changes in distributed assemblies, and for investigating their cortical topographies.

High-frequency spectral responses revealed processing differences between words and matched meaningless pseudowords. EEG and MEG experiments indicated that words elicit stronger high-frequency responses than pseudowords. This can be explained by assuming that although cell assemblies generating well-timed high-frequency activity are activated when words are being processed, no assembly becomes fully active when pseudowords are being perceived. A double dissociation in high-frequency responses could be observed over motor and visual cortices when action and vision words were processed. This provides additional evidence that cortical topographies of assemblies representing action and vision words differ. Furthermore, these findings suggest that the inves-

tigation of high-frequency cortical responses is a relevant tool for addressing important questions in cognitive neuroscience.

Experiments on electrocortical counterparts of words eliciting motor and visual associations showed some agreement between event-related potential and spectral response data. These measures revealed similar topographic differences between word types (see difference maps in Fig. 9). It may therefore be speculated that these measures provide information about different processes occurring in the same cell assemblies. The Hebbian framework suggests that initial activation (ignition) of cell assemblies and subsequent reverberation of neuronal activity in the assembly may be related to the two measures (see postulates 5 and 6 in sect. 4.1). After its ignition, the assembly may reverberate so that large numbers of neurons repeatedly become coherently active. Consistent with this view, peaks in the event-related potential distinguishing between word classes were seen as early as 150–200 msec after the onset of stimuli, but differences in spectral responses between words and pseudowords or between action and vision words were significant only around 300 msec or later. Early word class-specific components of the event-related potential and later dynamics in high-frequency spectral responses may therefore be related to early activation (ignition) and subsequent reverberatory activity of cell assemblies.

There is another aspect with regard to which the Hebbian approach sharply differs from many current theories of word processing. Many models assume that different aspects of a word – its phonological, grammatical, and semantic information – are processed in successive steps, for example, starting with phonology and finally arriving at semantics (see, e.g., Caplan 1992). In contrast, Hebb's view would imply that sufficiently strong partial stimulation of an assembly leads to its full ignition (Braitenberg 1978b). Ignition is assumed to be an instantaneous process, thus activating all parts of an assembly at almost the same time. According to the present proposal, phonological information is stored in perisylvian cortices, whereas at least some aspects of a word's meaning are laid down in assembly parts housed outside perisylvian space. The assumption of a one-shot activation would imply that the earliest signs of activation occur near-simultaneously in “semantic” areas and in “phonological” perisylvian cortices.

In lexical decisions tasks, electrocortical differences between words and pronounceable and orthographically regular pseudowords are usually not found before 200 msec after the onset of visual stimuli (Pulvermüller et al. 1995a), although differences between words and letter strings can occur earlier, that is, around 150 to 200 msec (Compton et al. 1991). This may be interpreted as evidence that the phonological or orthographic word form is being accessed around  $\frac{1}{5}$  of a second after the stimulus is being displayed. It is important to note that word class-differences in event-related potentials summarized above were also present around 200 msec post stimulus onset, or even earlier.<sup>26</sup> Some of these differences were present close to visual and motor cortices and far from left-perisylvian areas. This can be explained by assuming that they relate to semantic word properties.<sup>27</sup> If this interpretation is correct, the summarized data would support the idea of a one-shot activation of cell assemblies in which phonological and semantic information of words are bound together and simultaneously accessed around 150–200 msec after the onset of visually

presented word stimuli.<sup>28</sup> It is clear that simultaneous activation of left-perisylvian “phonological” and extra-perisylvian “semantic” areas would be inconsistent with many serial models of word processing, provided that activations of processing stages are assumed to follow each other with a measurable delay.

Although the Hebbian framework may represent an alternative to current modular theories of word processing, the two approaches are not entirely incompatible. Strictly speaking, a double dissociation in neuropsychology implies that distinct brain parts are necessary for performing two tasks. It does not prove that the processes necessary for each of the two tasks are independent of each other (although this is sometimes a convenient assumption). The Hebbian framework highlights how processes that are subject to neuropsychological double dissociation can nevertheless be tightly linked and occur simultaneously in the intact brain.<sup>29</sup> The double dissociation between nouns and verbs seen in certain aphasic patients does not prove that these word categories are processed in independent modules. Rather, it appears that they share most of their neuroanatomical “processing space,” but that each involves neurons in additional specific areas. Likewise, lesions affecting these “additional” areas can explain category-specific impairments. The double dissociation between agrammatism and anomia with regard to the processing of content and function words (agrammatics have difficulty with function words, anomics with content words) can also be explained based on the assumption of overlapping but distinct neuroanatomical “processing spaces” (see sect. 3, and Pulvermüller 1995a for further discussion). In some cases, the neuropsychological double dissociations correspond to interactions of the stimulus and topography variables in psychophysiological data. Processing loci may be distinct, although, functionally, neurons at distant loci interact. These neuropsychological data are entirely compatible with the view that phonological and semantic information is bound together in functional units and accessed almost simultaneously when words are being processed.

## 8. A word on syntax

For many language scientists, the question of how words are represented and processed in the brain is only a very basic one. Even substantial improvements in our understanding of the mechanisms underlying word processing may therefore be acclaimed only if the theoretical framework they are based on can offer perspectives on the solution of more complex problems known to be crucial for language. The question of how serial order is achieved in syntactic word strings has long been considered to be at the heart of our language faculty, and, from a theoretical point of view, it appears important to make clear whether a neurobiological model of word processing can offer perspectives on the biological reality of grammar (Lashley 1951).

The claim held by many linguists that language mechanisms cannot be explained by associative learning principles<sup>30</sup> may be the basis for a premature rejection of an approach postulating that so much can actually be explained based on Hebbian learning. However, one is well advised not to throw out the baby with the bathwater (i.e., a perspective on the neurobiology of language with the principles of associative learning). As pointed out in much detail

in section 3, even a Hebbian explanation can only function if inborn neuroanatomical and neurophysiological prerequisites are assumed to be present. More to the point, perisylvian cortices need strong intrinsic connections (Deacon 1992a; 1992b) to allow for the formation of assemblies representing word forms, and the statistics of neuroanatomical connectivity must be such that an assembly exhibiting the functional states of ignition and reverberation can form. It is possible that inborn properties of the brain have functional consequences that are crucial for processing sentences. This would be entirely compatible with a Hebbian approach: Although associative learning is an important factor for the formation of assemblies, some of their functional properties may be genetically determined and hard-wired in cortico-cortical connectivity. It would be nice to show that universal principles underlying activity dynamics of cell assemblies can be relevant for achieving serial order of words in sentences.

Looking at what is known about activity dynamics in stimulus-specific neuronal populations, Abeles’s well-timed activity patterns discussed in section 3 come to mind. In addition, Fuster’s (1989; 1995) finding that cells activated by specific stimulus properties can stay active for a period of several seconds after stimulus presentation may be of utmost importance here. It is certainly possible that these “memory cells” retain their activity because they are part of stimulus-specific cell assemblies in which neuronal activity reverberates. If so, these cells reveal important information about activity dynamics of the cell assemblies to which they belong (Fuster 1994).

Many memory cells exhibit well-defined activity dynamics. They do not show constantly enhanced activity after presentation of the stimulus they are specialized for, but, instead, they are strongly activated initially and lose activity almost exponentially thereafter (Fig. 10, left). What role could such neuronal units with exponential activity loss play in processing syntactic information?

Assume that several such assemblies have been primed one after the other. Due to the exponential decline of activity, the assembly activated first will later be at the lowest activity level, whereas the neuronal unit that was the last to be activated would still maintain the highest level of activity (Fig. 10, lower diagram). The information about the sequence of activations is thus stored in the hierarchy of activity levels of assemblies. Assuming a read-out mechanism that fully activates and then deactivates only the cell assembly at the highest level of activity, a set of exponentially declining assemblies can be considered equivalent to a pushdown store (Pulvermüller 1993). The unit primed first will be fully activated last, and, vice versa, the last to be primed would be the first to become fully active. If, for example, a speaker intends to say that three different persons have performed three different actions, the speaker could first talk about the actors whereby the neuronal representation of the corresponding nouns would be activated. If there is activity flow from the noun representations A, B, and C to the corresponding verb representations A\*, B\*, and C\*, respectively, the successive activation of noun representations would lead to activity levels of the three verb representations that exhibit the hierarchy shown in Figure 10. Ignition would therefore occur first in C\*, later in B\*, and finally in A\*, leading to a mirror image activation sequence, that is ABCC\*B\*A\*. This mechanism could be crucial for producing center-embedded strings, such as, for

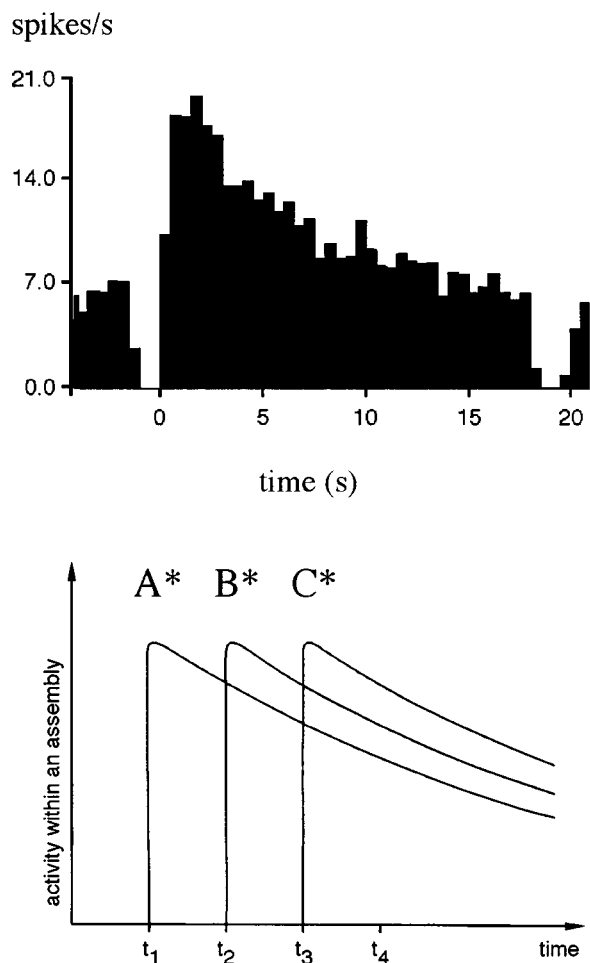


Figure 10. Top: Activity of a memory cell possibly reflecting activity dynamics in a large cell assembly to which it belongs. Note the almost exponential decrease of activity with time. (Adopted from Fuster 1995.) Bottom: If several assemblies of this kind are activated one after the other, the activation sequence is stored in the hierarchy of activity levels of assemblies. This mechanism could be used as a neuronal pushdown store. (Adapted from Pulvermüller 1993.)

example (The man [whom the girl (that the dog bit) met] came).<sup>31</sup> Linguists have argued that a pushdown store is quite helpful in processing such syntactic structures and, more generally, may be an integral part of our language capacity (Chomsky 1963). To derive a more powerful proposal, the assemblies making up the pushdown device may be assumed to correspond not to individual words, but to lexical or syntactic categories (for further elaboration, see Pulvermüller 1993; 1994b; 1996b).

The proposed neuronal mechanism is but one way a pushdown device could be realized in a neurobiologically plausible network consisting of cell assemblies. For it to operate properly, it must be postulated that several cell assemblies lose activity according to the same slope. It may be that it is a specific feature of the human cortex that it allows for the formation of cell assemblies with similar deactivation slopes. This could be the result of features of cortical connectivity and physiological properties of nerve cells. One possibility is that cortico-cortical links in perisylvian cortex are so numerous that very tightly connected assemblies can form therein. The deactivation slopes of these as-

semblies may then be determined by physiological properties common to many neurons, for example their refractory periods. Admittedly, this is speculative. However, the speculation illustrates how linguistic universals (e.g., center-embedding) can manifest themselves in brain properties, which can be revealed by neuroscientific research.

Although it is not possible to discuss syntactic issues in more detail here, these remarks on center-embedding may suffice to show that neurobiological models of language are not necessarily restricted to the single-word level. In fact, they can offer perspectives on the problem of serial order in behavior that meet linguists' claims that language mechanisms cannot be understood without considering brain mechanisms (see, e.g., Mandelbrot 1954 and Schnelle 1996a; 1996b).

What makes the results summarized in sections 5 and 6 so interesting is that many of them can be systematized and explained based on principles of associative learning evident from neuroscientific research in animals. However, this should not obscure the fact that more than associative learning is necessary to develop a neurobiological perspective on language. This target article was intended to show that biological models of word processing can produce highly specific predictions on brain processes, and that such predictions can produce experiments yielding quite unexpected results that prove the predictions correct. Furthermore, it was argued that there is at least a perspective on further developing the neurobiological approach to account for problems of serial order in behavior. It may appear evident, therefore, that biological research on language – both theoretical and empirical – is fruitful.

#### ACKNOWLEDGMENTS

I am particularly indebted to Valentino Braitenberg, with whom many of the summarized ideas have been developed, and to Werner Lutzenberger, my teacher in psychophysiology and experimental psychology, with whom I had the great pleasure to collaborate in the psychophysiological investigation of word processing. I am also grateful to Ad Aertsen, Robert Miller, Peter M. Milner, and Bettina Mohr, and to the five referees for their important comments, criticisms, and suggestions regarding an earlier version of this manuscript. Furthermore, I should thank my colleagues Siggie Haug, Bettina Mohr, Hans Schleichert, and Ralf Veit for their substantial help. Without this support it would have been impossible for me to spend so many days away from the lab writing this article. This research was made possible by a Heisenberg fellowship and research grants from the Deutsche Forschungsgemeinschaft (Pu 97/2-3, Pu 97/5-1, Pu 97/10-1, and Pu 97/11-1).

#### NOTES

1. The frequency range >20 Hz is called the "high beta band" or the "gamma band."
2. This is based on the assumption (implicit in Hebb's diagram) that loops include only a small number of neuron subgroups. If a loop includes larger numbers of subgroups, the conclusion can only be maintained if shorter connections (and therefore shorter conduction times) are assumed.
3. For the explanation of unimodal deficits such as seen in a few cases of agrammatism (Caramazza & Berndt 1985; Kolk et al. 1985), additional assumptions are necessary. See Pulvermüller (1995a) and Pulvermüller and Preissl (1991; 1994) for discussion.
4. Assembly ignition must be possible without overt articulation of the word. Whereas infants may automatically repeat word forms and use the word form when presented with an object to which the word refers, adults have learned to think of a word with-



out activating motor programs. This may be done by reducing the global level of activity in motor cortices controlling mouth movements. In this case, ignition of an assembly will not cause overt movements.

**5.** The same mechanism may underlie synonymy. Assemblies representing synonyms may share extra-perisylvian neurons but differ in their perisylvian parts (Pulvermüller 1996b).

**6.** I will use the term “semantic categories” although it is clear that the perceptual/motor modalities and stimulus properties involved when learning word meanings are the primary causes of different cortical distribution of representations. Thus, only one aspect of the meaning of words is assumed to play a role. Accepting Frege’s (1980) distinction between “Sinn” and “Bedeutung,” it would be the “Bedeutung” but not the “Sinn” that is relevant.

**7.** See note 6.

**8.** Such activity in the frontal lobe may be related to associations of body movements. As discussed in section 3.3.3, additional activity in posterior temporal and parietal areas may be related to visual perception of movements (being performed by oneself or by others) and to somatosensory self-stimulation during performance of body movements.

**9.** This list is probably not complete. There may be additional properties of word stimuli affecting physiological processes. For the mentioned variables, such effects are well documented.

**10.** Additional possible confounds are emotional properties of words, such as arousal and valence (see sect. 4.1). None of the studies checked these variables, and it appears likely that arousal values are higher for content than for function words. Therefore, it cannot be excluded that arousal differences underlie electrocortical differences between these word categories. However, it may be impossible to find high-arousal function words or content words sufficiently low in arousal to allow for stimulus matching for this variable. Differences in emotional stimulus properties – similar to those in concreteness – appear characteristic for the classes of content and function words.

**11.** This latter difference was enlarged when a semantic decision task was used (Kounios & Holcomb 1994).

**12.** It is not always clear from the publications what the exact instructions were. Subjects may have been told to say verbs that describe “what the nouns might be used for or what they might do” (Fiez et al. 1996, p. 1) or they may have been told to “say aloud a use for (. . .) nouns” (Posner et al. 1988, p. 1630). Clearly, the second instruction may bias the response toward the use of action words, whereas there is no such bias in the first instruction.

**13.** There were additional loci of enhanced activity including cerebellum and gyrus cinguli.

**14.** Only numerals and verbs were matched for frequency.

**15.** In fact, stimuli were chosen so that these ratings were likely. In an experiment performed before the actual EEG study, stimuli had been rated by a different set of subjects to allow for selection of action and vision words (Preissl et al. 1995).

**16.** In a recent study (Pulvermüller et al., submitted), we compared nouns from the action word category (tool names and other nouns rated to elicit strong motor associations) to nouns from the vision word category and found electrocortical differences between word categories quite similar to noun/verb-differences. This further supports the interpretation that semantic word properties are relevant for topographies of cortical activity.

**17.** The experiment was carried out twice, with essentially the same results.

**18.** For example, Martin et al. (1996) matched target words of the naming task for word frequency, the PET studies mentioned above did not. Furthermore, in PET studies frequency of stimulus presentation per minute and presentation time are additional factors strongly influencing amplitude and topography of brain responses (Price et al. 1994). Although the influence of these factors is at present not fully understood, it appears necessary to interpret the available data.

**19.** These predictions follow if an associative learning principle is assumed not only for modification of cortico-cortical synapses,

but also for modification of cortico-subcortical connections such as the fibers from the cortex to amygdala and to additional nuclei of the limbic system (see Pulvermüller & Schumann 1994 for discussion).

**20.** It is less likely that increased activity in the temporal pole and amygdala is related to stimulus properties, because faces were also presented in the baseline condition.

**21.** There are exceptions: For example, Damasio et al. (1996) found bihemispheric activation at least during processing of persons’ names, Martin et al. (1995) found bihemispheric differences between verb and color name processing, and McCarthy et al. (1993) obtained bihemispheric activation of inferior prefrontal areas during verb generation.

**22.** There is so far no evidence that processing of action words also activates parietal regions related to the processing of somatosensory self-stimulation during movements – although this appears likely based on the Hebbian approach (see sect. 3.3.3).

**23.** There are meaningful stimuli of other modalities – odors, tastes, nonlinguistic sounds, etc. – that can be assumed to be represented in cortical assemblies.

**24.** Furthermore, current source density analysis makes it possible to obtain reference-free data. Using raw EEG data would make it impossible to determine whether spectral activity was generated at the critical electrode or at the reference, and, in addition, spectral activity at the reference electrode would enter the data as noise.

**25.** The fact that the right perisylvian region did not evidence word/pseudoword differences in high-frequency responses may be the result of a relatively low density of assembly neurons in this areas (Fig. 4). Note that high-frequency signals in EEG and MEG recordings are small even over the left hemisphere (Lutzenberger et al. 1997).

**26.** The variation of these delays may be related to properties of the stimuli, for example their word frequency.

**27.** Further support for early access to semantic information (around 200 msec after stimulus onset) during word processing comes from recent experiments where semantic and visual discrimination tasks were used (see, e.g., Posner & Raichle 1994, p. 143).

**28.** For acoustic presentation, the delay must be longer because the point in time where word recognition is possible is usually several hundred milliseconds after word onset.

**29.** A similar possibility has earlier been discussed by Tim Shallice (1988).

**30.** This claim is somewhat in contrast to the fact that recent attempts to extract linguistic regularities from the language input using hidden Markov models (Charniak 1993) and perceptron-like networks including a memory component (Elman 1990) were quite successful.

**31.** Brackets indicate levels of embedding.