



## Review article

## Neural reuse of action perception circuits for language, concepts and communication

Friedemann Pulvermüller<sup>a,b,c,\*</sup><sup>a</sup> Brain Language Laboratory, Department of Philosophy & Humanities, WE4, Freie Universität Berlin, 14195 Berlin, Germany<sup>b</sup> Berlin School of Mind and Brain, Humboldt Universität zu Berlin, 10099 Berlin, Germany<sup>c</sup> Einstein Center for Neurosciences, Berlin 10117 Berlin, Germany

## ARTICLE INFO

## Article history:

Received 20 January 2017

Received in revised form 12 May 2017

Accepted 13 July 2017

Available online 20 July 2017

## Keywords:

Language

Action

Perception

Grounded cognition

Symbolic system

Motor system

Perisylvian cortex

Neurocomputation

Neuroanatomy

Neuroimaging

Neuropsychology

Phonemes

Words

Semantics

Pragmatics

Communication

Social interaction

## ABSTRACT

Neurocognitive and neurolinguistics theories make explicit statements relating specialized cognitive and linguistic processes to specific brain loci. These linking hypotheses are in need of neurobiological justification and explanation. Recent mathematical models of human language mechanisms constrained by fundamental neuroscience principles and established knowledge about comparative neuroanatomy offer explanations for *where*, *when* and *how* language is processed in the human brain. In these models, network structure and connectivity along with action- and perception-induced correlation of neuronal activity co-determine neurocognitive mechanisms. **Language learning leads to the formation of action perception circuits (APCs) with specific distributions across cortical areas. Cognitive and linguistic processes such as speech production, comprehension, verbal working memory and prediction are modelled by activity dynamics in these APCs, and combinatorial and communicative-interactive knowledge is organized in the dynamics within, and connections between APCs. The network models and, in particular, the concept of distributionally-specific circuits, can account for some previously not well understood facts about the cortical 'hubs' for semantic processing and the motor system's role in language understanding and speech sound recognition.** A review of experimental data evaluates predictions of the APC model and alternative theories, also providing detailed discussion of some seemingly contradictory findings. Throughout, recent disputes about the role of mirror neurons and grounded cognition in language and communication are assessed critically.

© 2017 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Contents

1. The need for neurobiological explanation	2
2. Neuroscience foundations of language	3
2.1. Neuroanatomical structure and connectivity	3
2.2. Learning and neuronal circuit function	5
2.3. What is special about human infants?	8
2.4. Toward building a language brain: neural reuse of action and perception mechanisms	8
2.4.1. Mirror neurons, action perception circuits and language	11
2.5. Attention and flexibility of activation	12

**Abbreviations:** A1(C), primary auditory cortex, model area; AB(C), auditory belt cortex, model area; AF, arcuate fascicle; APC, action perception circuit; BA, Brodmann area; **CNA, combinatorial neuronal assembly**; DTI/DWI, diffusion tensor/weighted imaging; EEG, electroencephalography; fMRI, functional magnetic resonance imaging; IFC/G, inferior frontal cortex/gyrus; IPC/G, inferior parietal cortex/gyrus; IPS, intraparietal sulcus; M1(C), primary motor cortex, model area; MEG, magnetoencephalography; MTC/G, middle temporal cortex/gyrus; PB(C), auditory parabelt cortex, model area; PF(C), prefrontal cortex, model area; PM(C), premotor cortex, model area; ROI, region of interest; SD, sequence detector; SMC, sensorimotor cortex; SMG, supramarginal gyrus; STC/G/S, superior temporal cortex/gyrus/sulcus; tDCS, transcranial direct current stimulation; TMS, transcranial magnetic stimulation.

\* Correspondence to: Brain Language Laboratory, Department of Philosophy & Humanities, WE4, Freie Universität Berlin, Habelschwerdter Allee 45, 14195, Berlin, Germany.

E-mail address: [friedemann.pulvermuller@fu-berlin.de](mailto:friedemann.pulvermuller@fu-berlin.de) (F. Pulvermüller).

<http://dx.doi.org/10.1016/j.pneurobio.2017.07.001>

0301-0082/© 2017 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

2.6.	Multiple cognitive demands: memory, intention, prediction	13
2.7.	Combinatorial learning and generalization	15
2.7.1.	Formation of combinatorial neuronal assemblies	16
2.7.2.	Anti-Hebbian dissociation learning and statistical preemption in language acquisition	18
2.8.	Communicative interaction	19
2.9.	Interim summary	20
3.	Action-perception circuits (APCs) for language: experimental tests	20
3.1.	Topographically specific APCs for speech sounds in frontotemporal perisylvian cortex	21
3.1.1.	Is there a causal influence of sensorimotor cortex on speech perception and comprehension?	24
3.1.2.	Overcritical approaches to action perception integration	25
3.1.3.	Action perception circuits for speech and the motor theory of speech perception	26
3.2.	Topographically specific semantic circuits	26
3.2.1.	Abstract culture-specific concepts as a test case	30
3.2.2.	Semantic grounding and the role of modality-preferential areas in language processing	31
3.2.3.	Neurobiologically constrained modelling of semantic grounding	32
3.2.4.	Objections on empirical grounds	32
3.3.	Circuits for communicative actions and interactions	36
4.	Summary and outlook	36
	Acknowledgements	37
	References	37

## 1. The need for neurobiological explanation

Humans are special. They have a special kind of language and cognize particularly efficiently in cooperative and meaningful social interaction (Dunbar, 1993, 2016; Tomasello, 2008, 2014). Animal communication systems seem more restricted, although some species show surprising abilities to process structurally rich strings (Abe and Watanabe, 2011; Gentner et al., 2006). Animal ‘languages’ tend to lack rich repertoires of vocal and manual gestures, large vocabularies, complex syntax, arbitrary symbol-meaning mappings and flexible schemas of action sequences that guide the interactions between individuals. There is no need to explain why humans, but not other closely related species, develop these communicative and social skills. Fortunately, recent neuroscience research has provided important insights into the specific features of human brain anatomy and function, which open new perspectives on answering the big question about the specificity of human cognition by mechanisms rooted in human neurobiology. However, to achieve this, it is necessary to spell out and understand the *mechanistic relationship* between language and communication and their basis in neurobiological structure and function.

Most current attempts to connect cognition and language to their neurobiological substrate present ‘linking hypotheses’ of the form “brain area B houses cognitive process C”. **These hypotheses are descriptive without offering explanations.** For example, a recent effort towards a ‘real neuroscience of language’ proposes that one important aspect of cognition, *conceptual processing*, crucially involves the posterior middle temporal gyrus and/or superior temporal sulcus (p. 178f, Hickok, 2014). The questions *why* these brain sites should become most important for the cognitive mechanism of interest, and *why* other areas should stand by, remain unanswered.<sup>1</sup> However, for scientific explanation it is necessary to address the questions *why* and *how* the brain substrates of abstract cognitive processes might come about, and which functional properties of these processes might be relevant for their implementation in neuron circuits. Such explanations can be based on a rich reservoir of established neurobiological facts and principles. These include information about processing properties of nerve cells, their functional

connections to other close-by neural elements and to neurons in distant regions of the brain, about principles according to which these connections are being modified as a consequence of previous activations, and about a multitude of other facts.

In the present paper, I will argue that an understanding of human brain-cognition relationships requires that *why* questions about the links between cognitive and neuronal mechanisms be addressed explicitly and systematically. I will also argue that neuroscience evidence currently available enables us to provide such explanations, i.e. there is no need to wait until more evidence is available (although new data are always welcome to improve current theory). Neurobiological explanations answering *why* questions require foundation in established neuroscience wisdom and principles. As many of these principles refer to neurons as the fundamental units of brain function, explanations cannot be restricted to the coarse level of cortical areas or fiber pathways but need to deal with neurons and their connections and interaction in neuronal circuits. **In the case of human-specific language and communication, explanation requires a move beyond the level of area labelling, towards spelling out language in terms of neurons and neuron circuits with specific distributions and topographies.**

This review will discuss established models in search of possible explanations for where and how language and communication is organized in the human brain, when linguistic comprehension and production processes emerge and how different components of distributed language circuits interact. Questions such as the following will be addressed: Why would language areas develop in humans but not in non-human primates? How would the nerve cell circuits carrying language be wired up and distributed across cortical areas? How would these circuits activate in service of language understanding and use, and across different tasks requiring memory or prediction? Which circuit structures and distributions emerge in semantic learning and how would such circuits process aspects of the meaning of words and sentences, and the communicative function of speech acts?

A basic theoretical foundation for brain language theories comes from neuroanatomical structure and neurophysiological principles, such as Hebbian learning mechanisms. These principles determine the formation of distributed cell assemblies, which link information about motor movements and actions with information about perceptions. **A crucial postulate, which I explored earlier, is that these distributed neuronal circuits are reused for language, for example for binding articulatory and acoustic**

<sup>1</sup> To be fair, the author mentions ‘ingredients’ of an explanation of why these areas are involved, which include ‘the hierarchical organization’ of conceptual and brain systems (p. 171). There is still a long way to go from such ‘ingredients’ to an explanation why a specific area becomes relevant for concepts.

phonological information in the processing of speech sounds and syllables, and for binding linguistic form and semantic knowledge in meaning processing (Pulvermüller, 1999). Although, originally, the proposed framework had narrow scope and a limited data basis, substantial evidence has meanwhile accumulated and some of its critical predictions have received direct experimental support (see Kiefer and Pulvermüller, 2012; Pulvermüller, 2005, 2013a; Pulvermüller and Fadiga, 2010). The present paper now advances the proposal of neural reuse of action perception circuits for language (1) by systematically applying new insights from recent cognitive and neuroscience research, particularly in our knowledge about human neuroanatomy, (2) by extending the proposal to cover communication and social interaction, and (3) by founding it in mathematically precise neurocomputational models that imitate cortical structure and function. The latter step is especially essential for the explanatory main purpose of this account, as the model simulations can indeed provide proofs-of-concept that the envisaged circuits emerge and function in a biologically-constrained architecture and that the same circuits can provide the mechanistic basis of cognitive processes as diverse as prediction, decision, memory, attention, combination and generalization. A further aim of this contribution is to address current debates in the cognitive and neurosciences in light of recent data, in particular about mirror neurons and grounded conceptual processing. As language mechanisms are a central focus of these disputes, the present review will critically evaluate the arguments brought forward and work towards an integrative framework for brain language research that fruitfully applies major neurocognitive achievements.

## 2. Neuroscience foundations of language

### 2.1. Neuroanatomical structure and connectivity

Detailed structural and functional knowledge is available about the brain. Some of this knowledge applies to brains of various species and can be summarized by general principles, for example the structural principle of *topographical projections*, which states that long-distance connections preserve neighborhood relationships. For language mechanisms, the cerebral cortex is the most important brain structure, although it is clear that its functionality crucially depends on subcortical input, in particular from the reticular formation, and functional interaction with basal ganglia, thalamus and limbic structures (see, for example, Nadeau and Crosson, 1997). The majority of cortical neurons are excitatory pyramidal cells, which typically have local axon branches reaching adjacent nerve cells (within about 1 mm), but also a long axon, which can be several centimeters long and reach distant locations (Braitenberg and Schüz, 1998).

An important neuroanatomical principle is that of rich local but more selective and specific long-distance connectivity. Neighboring neurons up to a distance of a few hundred micrometer have a relatively high probability of being connected with each other (Braitenberg and Schüz, 1998), which falls off with local distance (Perin et al., 2011). Adjacent areas are still reasonably likely to exhibit direct next-neighbor projections, but areas far apart are linked in a more specific and selective fashion (Braitenberg and Schüz, 1998; Schüz and Braitenberg, 2002; Sporns et al., 2007; Young et al., 1994). The connectivity principle of 'local richness and long-range sparseness' is manifest quantitatively in the number of fibers of local cortical connections (in the order of  $10^{11}$  in human cortex), semi-local so-called U-fibers ( $10^{10}$ ) and true long-distance tracts ( $10^9$  fibers of length  $>3$  cm, Schüz and Braitenberg, 2002). This implies a 'small world' network structure and allows for interlinking neural units via relatively small numbers of connection steps (small average topological 'path length', van den Heuvel

and Sporns, 2013). Topological analysis of the cortical between-area connection structure of different mammal species (cat, macaque, human) suggests highly connected local systems or 'modules' for processing information from one modality (visual, auditory, motor etc.) interlinked by sparse and specific long-distance fiber tracts. Of particular relevance for interlinking cortical systems are areas strategically placed at the interface between sensory and motor systems, sometimes called 'connector hubs' (Sporns et al., 2007; van den Heuvel and Sporns, 2013). As neuronal activity from sensory and/or motor areas converges on neurons in these anatomically-defined 'hub' areas, Antonio Damasio labelled them 'convergence zones' (Damasio, 1989).<sup>2</sup> The summarized features of statistical neuroanatomy and connectivity suggest that it is a main function of the cortex to efficiently 'merge and mix' information carried by neurons in sensory, motor and convergence/connector hub areas (Braitenberg and Schüz, 1998).

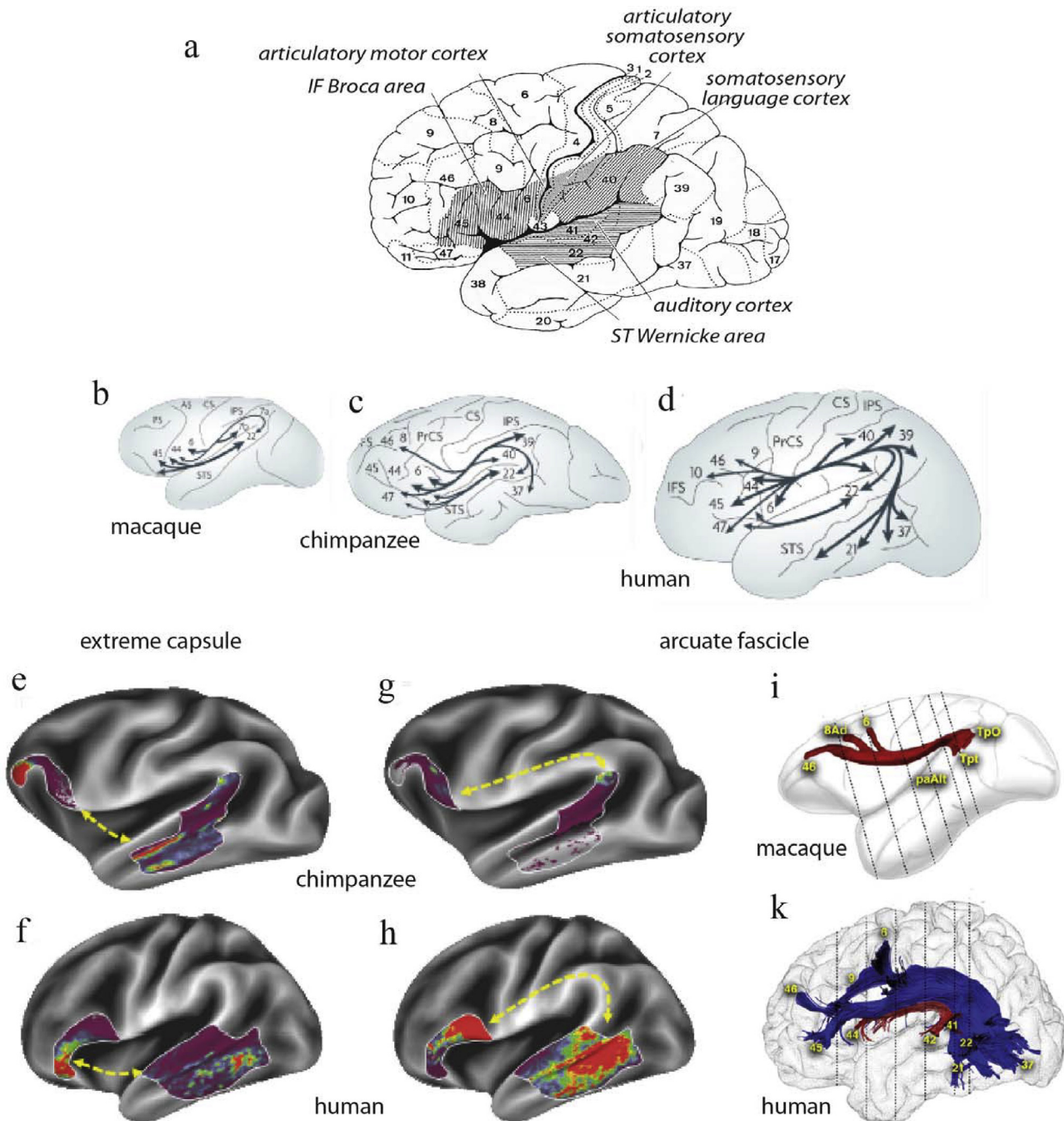
Cortical connectivity has functional implications. Due to the cortex' small world topology and the availability of connector hubs, a neuron in motor cortex may receive and process sensory information in addition to its role as a motor unit. Given the connections between sensory and motor areas via connector hubs are reciprocal, the motor neuron may not only process sensory information but even influence, or 'have a say', in perceptual processing. Vice versa, a neuron in sensory cortex may contribute to motor processing because of a short 'path length' and strong connections to motor systems. Between-area connectivity structure of the cortex can be used to explain important aspects of cortical function and dysfunction (Deco et al., 2013; Fornito et al., 2015; van den Heuvel and Sporns, 2013) and may therefore contribute to a better understanding of language mechanisms.

When searching for mechanisms supporting uniquely human abilities such as language, specific structural features of the human brain are of special interest. Already gross anatomical measures, such as relative brain size (Jerison, 1975) and the total number of neurons (Herculano-Houzel, 2009; Schüz and Sultan, 2009), seem to set apart humans from other primates, although it is not clear how an exceptionally large or linearly scaled-up version of a primate brain might bring about the crucial mechanisms for human cognition and language. Laterality of anatomical structures has long been assumed to be crucial for language (e.g., Jacobs et al., 1993; Steinmetz et al., 1991) although, again, it is unclear how, for example, a relatively larger planum temporale on the left may contribute to, or result from, language use. Human cerebral cortex includes  $1.5\text{--}3.2 \times 10^{10}$  neurons (Haug, 1987; Pelvig et al., 2008), most (ca. 85%) of which are excitatory pyramidal cells, each carrying some  $10^4$  synapses (Braitenberg and Schüz, 1998). Therefore, assuming random connectivity, a given neuron reaches ca.  $10^4$  other neurons with one synaptic step, ca.  $10^8$  with two, and a path length of three steps should be sufficient for connecting any two cortical neurons (Palm, 1982; Palm et al., 2014). However, as network topology is biased towards local connections within sensory and motor systems (or 'modules') and long-distance links between these 'modules' via connector hubs are relatively sparse, the shortest sensorimotor path lengths may require more than three synaptic steps, thus making sensorimotor information flow less direct and possibly more complex.

Fig. 1 illustrates a recently reported specific feature of human cortical anatomy, which seems to be of special relevance to language. Studies using diffusion tensor and diffusion weighted imaging (DTI, DWI) along with non-invasive probabilistic tractography indicate that the arcuate fascicle or AF, a dorsal fiber bundle

<sup>2</sup> The classic term '(multimodal) association areas' will also be used occasionally in the same sense.





**Fig. 1.** Neuroanatomical connectivity of the language cortex. a. Left perisylvian language cortex. b–d. Schematic connectivity structure of the perisylvian cortex in macaque, chimpanzee and human (modified from, Rilling et al., 2008). e–k. Tractography results from macaque (i), chimpanzee (e, g) and human (f, h, k). Projections of the ventral extreme capsule (e, f) and the dorsal arcuate fascicle (g–i) are separated (adopted from Rilling et al., 2012; Thiebaut de Schotten et al., 2012).

connecting temporal and inferior frontal cortex, is developed particularly strongly in humans (Fig. 1h, k). Whereas the ventral connections between temporal and frontal lobes, including the extreme capsule, do not seem to have changed massively in primate evolution (Fig. 1e–f), this dorsal bundle is weakly developed in macaques, more prominent in chimpanzees, but rich and strong in humans (Fig. 1b–d, g–k), where it is estimated to contain ca.  $10^7$  fibers (Schüz and Braitenberg, 2002). It needs to be noted that tracer studies in macaques have demonstrated AF connections between frontal and superior temporal cortex in macaques (Petrides and Pandya, 2009); the comparative DTI work

therefore suggests that rather than being newly established, this link has been modified and extended during evolution so that it is now much more powerful in humans than in monkeys or apes. Crucially, much of the human AF connection is available already early in ontogeny, shortly after birth (Dubois et al., 2014, 2009). As it is strongly lateralized to the left hemisphere – the language dominant hemisphere in most of us – the arcuate may provide a prime substrate for specifically human language mechanisms (Rilling, 2014; Rilling et al., 2008; Thiebaut de Schotten et al., 2012). The AF's importance for language is further bolstered by its role as a main connection highway between those areas of human

cortex that are most relevant for language and are therefore sometimes called the 'language cortex' or 'language network'. Note that the language cortex/network and its sub-areas are defined quite differently by different researchers (see, for example, Bogen and Bogen, 1976). Still, as a common denominator, it is normally taken to include the first convolution surrounding the Sylvian (or lateral) fissure, which is called the *perisylvian language cortex* (Bogen and Bogen, 1976). In this sense, the perisylvian language cortex includes the posterior inferior frontal area named after Broca (Brodmann areas, BA, 44, 45) along with the superior temporal cortex (BA 42, 22) sometimes considered Wernicke's region, plus further adjacent sites in inferior frontal, parietal and temporal cortex. Substantial lesions in the left perisylvian language cortex typically cause aphasia affecting both production and comprehension of language (Bates et al., 2003; Rosenbek et al., 1995); this general rule holds for most right- and the majority of left-handers, although exceptions from it are well-known (Basso et al., 1985).

Although a wide range of cortical areas in both hemispheres along with subcortical sites are relevant for language, and especially for semantic-conceptual processing (see especially Section 3.2), the structure and function of a core part of the perisylvian areas in frontal and temporal cortex may be of special importance to the understanding of language mechanisms. These core areas can be subdivided into inferior motor (M1: BA 4), premotor (PM: BA 6 and 44) and prefrontal cortex (PF: BA 45 and adjacent sites) in the frontal lobe and auditory core (A1), adjacent auditory belt (AB) and parabelt (PB) cortex in superior temporal lobe gyrus and sulcus (Fig. 2a, see Garagnani et al., 2008; Romanski et al., 1999). A similar subdivision applies to the parietal perisylvian areas. The AF connects some of these areas by curving dorsally around the posterior end of the Sylvian fissure. The other main long-distance fiber bundles interlinking perisylvian language areas are the already mentioned ventral extreme capsule, which connects inferior prefrontal cortex and superior temporal sulcus, and the inferior branch of the superior longitudinal fascicle, which connects inferior frontal and parietal areas. In close vicinity of perisylvian cortex, the uncinated fascicle bridges between temporal pole and orbitofrontal cortex. Connection bundles between perisylvian and extrasylvian sites include the frontooccipital fascicle interlinking inferior temporal and occipital areas with prefrontal areas and the aslant tract between inferior frontal and dorsomedial frontal sites (for review, see Dick et al., 2014). From a comparative neuroanatomical perspective, the arcuate stands out against these pathways for the massive changes it underwent in primate evolution (Thiebaut de Schotten et al., 2012).

The dorsal frontotemporal connection by way of the AF may not only have become more powerful across evolution, it may also have changed the topology of the core language network. This suggestion comes from DTI/DWI data and tractography performed in parallel in humans and in monkeys or apes. DTI/DWI and tractography data indicate that the weak dorsal link in macaques primarily connects temporal parabelt with prefrontal areas (including the homologue of BA 46, Thiebaut de Schotten et al., 2012). Likewise, the ventral and dorsal connections revealed by DTI in the chimpanzee seem to interconnect superior temporal cortex with anterior prefrontal sites, but not the more posterior and premotor areas homologue to human BA 44 (Rilling et al., 2012). These connections result in a topological structure of core language areas with next neighbor connections within frontal and temporal perisylvian areas (M1-PM-PF and A1-AB-PB) plus a long distance link between prefrontal and auditory belt areas (PF-PB). The latter 'higher' areas would thus serve as connector hubs (green arrows in Fig. 2c). In contrast, tractography studies in humans indicate a direct connection not only between inferior prefrontal and temporal parabelt areas, but, in addition, equally expressed

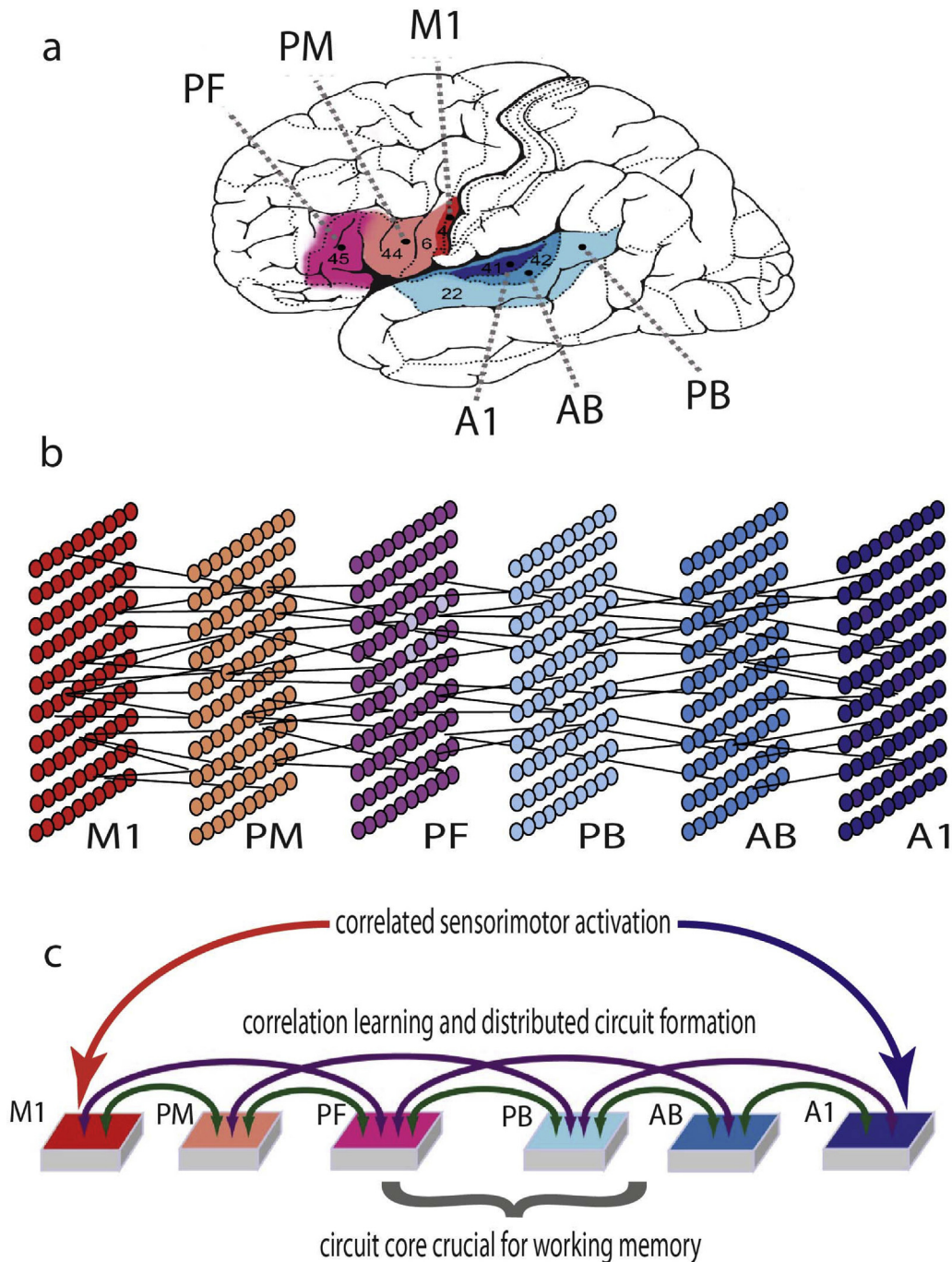
shortcuts that jump adjacent areas of the next-neighbor perisylvian area topology and reach second-next neighbors. 'Jumping' connections would thus provide direct connections between prefrontal and auditory belt and between premotor and auditory parabelt areas (see Fig. 1d, h, k and purple arrows in Fig. 2c, Rilling et al., 2012; Thiebaut de Schotten et al., 2012). Fig. 2c attempts to integrate these results by contrasting the tractography-based perisylvian topology in macaques (green arrows) with that in humans (green arrows plus the dorsal-specific 'jumping links' in purple, for discussion, see Garagnani and Pulvermüller, 2013; Pulvermüller and Garagnani, 2014; Schomers et al., 2017).

Notably, these suggested topological differences are somewhat in contrast with the already mentioned results of tracer studies in macaques, which show that AF and extreme capsule are both present in macaques and do in fact interlink the homologues of both inferior frontal language areas, BA 44 and 45, with superior temporal parabelt areas (Petrides and Pandya, 2009). It is therefore necessary to consider the possibility that the results about different topologies of the primate cortices are an artifact, for example of smaller brain size in the less 'advanced' species and the resultant smaller tracts and difficulty to map the curvature of the AF using tractography. However, if both tracer and tractography results are reliable, the 'jumping' links via both ventral and dorsal frontotemporal connections would be present but too sparse to yield a clear tractography result in the non-human species. In this case, the stronger and richer 'jumping' connections within the perisylvian core language network may represent a functionally-relevant difference between human and non-human primates and therefore are of potential relevance for the evolution of human language.

The suggested change in the topology of functional connectivity implies 1) a shorter minimal 'path length' from motor to auditory cortex (the schematic in Fig. 2c suggests 5 steps in macaques vs. 3 in humans) and 2) a greater 'degree' of connectivity especially of connector hub areas in the perisylvian network (2 in macaques, 4 in humans). The evolutionary step would be the topological change from two separate systems, the auditory and articulatory 'modules', bridged by a single connection, to a more substantially interlinked system in which frontotemporal links connect not only next neighboring areas but, in addition, second-next neighbors. In sum, the generally stronger connectivity and sensorimotor 'shortcuts' in humans imply more opportunity to 'mix, merge' and integrate auditory and motor information. In which way could such quantitative and qualitative topological features become relevant for human language?

## 2.2. Learning and neuronal circuit function

Without any doubt, language mechanisms crucially depend on neuronal function and learning. To yield a language with its specific repertoire of speech sounds, its unique vocabulary and peculiar set of syntactic rules and dialogue types, any network structure must be imprinted with novel information. Over and above human-neuroanatomical features, the functional principles underlying neuronal function and learning are therefore relevant to language theory. Language mechanisms are built from neurons, that is, functional elements characterized by activity states. In any given neuron, input from other neurons, or from sensors, changes the dendritic membrane potential, which is converted probabilistically into discrete action potentials, so that, at any point in time, neuron A either fires an action potential ( $f(A)$ ) or not ( $\neg f(A)$ ), so that it can therefore be seen as active or inactive. Now, crucially, connected neurons that 'fire together, wire together' and strengthen their mutual connections (Hebbian learning, Gustafsson et al., 1987; Hebb, 1949). This 'Hebbian rule' is one of the principles underlying neuronal plasticity. There is also the reverse effect that 'neurons



**Fig. 2.** Neurocomputational model of a central part of the language cortex. a. Frontotemporal perisylvian areas implemented in the model. b. Schematic illustration of the network structure. c. Connectivity structure between the six model areas, taking into account tractography results. In addition to next-neighbor links between areas (green arrow), which are manifest in macaques and in ventral and dorsal connections in humans, the 'jumping links' between second next neighbors of the human dorsal arcuate fascicle are implemented (purple).

out of sync delink' (Artola and Singer, 1993; Bienenstock et al., 1982; O'Reilly, 1998); uncorrelated and anti-correlated firing leads to weakening of any links. Put less colloquially, the change in weight  $\Delta w_{AB}$  of a particular synaptic connection from neurons A to neuron B increases with the probability of joint firing  $p(f(A), f(B))$ , but it also decreases with the probabilities of each neuron firing while the other is silent ( $p(\neg f(A), f(B))$  and  $p(f(A), \neg f(B))$ ). In this sense, correlation learning for a connection between two neurons A and B can be defined as the weight change

$$\Delta w_{AB} = c_1 p(f(A), f(B)) - c_2 p(\neg f(A), f(B)) - c_3 p(f(A), \neg f(B))$$

(where  $c_1, c_2$  and  $c_3$  are constants, with  $c_1 > c_2, c_3$ ). In addition, the timing of cortical firings can be important: whereas a presynaptic spike followed by the postsynaptic one typically leads to strengthening, the reverse order leads to weakening of the link (Bi and Poo, 1998; Caporale and Dan, 2008; Markram et al., 1997). Such spike-timing-dependent synaptic plasticity (STDP) entails the mapping of temporal information on neuronal connectivity. Therefore, correlated activity with appropriate timing leads to stronger connections, given the links necessary for learning are available in the first place.



**Table 1**

Hebb's original proposal of co-occurrence learning (top panel) is contrasted with correlation learning incorporating both Hebbian and 'anti-Hebbian' learning (bottom panel). Both proposals cover that 'neurons that fire together wire together' (long-term potentiation, LTP). Only the correlation approach takes into account 'that neurons out of sync delink' (long-term depression LTD). The latter may be necessary for modelling higher cognitive functions such as language (Garagnani et al., 2009b; Pulvermüller, 1999).

CO-OCCURRENCE RULE		Postsynaptic neuron B	
		active	inactive
Presynaptic neuron A	active	Synaptic strengthening	No synaptic change
	inactive	No synaptic change	No synaptic change
CORRELATION RULE		Postsynaptic neuron B	
		active	inactive
Presynaptic neuron A	active	Synaptic strengthening	Synaptic weakening
	inactive	Synaptic weakening	No synaptic change

As correlation mapping is one of the main neurobiological mechanisms believed to underlie learning, it provides a likely basis for language learning and acquisition too. Neuroanatomical connectivity sets the stage for such learning, whereby cortical connectivity structure not only enables, but also limits it.

Note that the cumulative synaptic modulation effects of (Hebbian) co-occurrence and ('Anti-Hebbian') one-sided-only firing in correlation mapping are different from classic associationist co-occurrence learning only (see Table 1); the former but not the latter type of learning makes it possible to model important aspects of language (Garagnani et al., 2009b; Keysers and Perrett, 2002; Pulvermüller, 1999). Please consider, for a moment, the contrast between the mapping of correlated neuronal activity with traditional associationist accounts. In the behaviorist tradition, the co-occurrence of stimuli, or of stimuli and responses, together with a reinforcing signal, has been proposed as basis for language learning (Skinner, 1957; Watson, 1925). This type of approach is problematic, not only because it does not take into account cognitive processes such as memory and prediction and fails to sufficiently treat forms of learning lacking reinforcement (for in-depth discussion in the context of semantics and action, see Alston, 1964). From a neurobiological perspective, a most problematic feature of associationist accounts in the behaviorist tradition is their focus on co-occurrence learning. In language use, many words occur together with many other words, and words and sentences co-occur with a wide variety of non-linguistic stimuli and responses, so that any co-occurrence mapping approach would inevitably associate everything with everything. Clearly, such an associationist account of language is not feasible and researchers emphasizing its inadequacy therefore make a valid point (Chomsky, 1959; Hickok, 2010; Leshinskaya and Caramazza, 2016). Still, such justified skepticism cannot lead to ignorance toward neurobiologically realistic learning mechanisms, which, doubtlessly, have great relevance for language. It is therefore important to keep in mind the differences between neurobiologically founded and outdated associationist concepts of learning. These differences lie not so much in the 'association term' of the above equation, expressing that 'what fires together, wires together'. An important difference lies in its 'out-of-sync-delink'-part, which is absent from traditional associationism: Connections weaken when pre- and post-synaptic activity states differ, or if post-synaptic precedes pre-synaptic activity. This implies dissociation learning in case of unrelated or anti-correlated activation. Dissociation learning prevents associations of "everything-with-everything", and thus provides the specificity needed for various facets of language mechanisms. Therefore, association conjoined with this type of dissociation learning is needed for any neurobiological implementation of language, already when storing

different word forms sharing articulatory, phonetic and phonological features (Garagnani et al., 2009b). This article will come back to this important point in different contexts, addressing, for example, word form and grammar learning along with semantic mapping.

Individual pyramidal neurons exert a small influence on other such neurons in cortex (Abeles, 1991); therefore, the joint action of neuron sets is necessary to cause relevant cortical effects. Above the level of the individual neuron and below the level of whole areas or brain systems, coherently acting groups of cortical neurons are therefore considered as functionally relevant in cortex (Abeles, 1991; Hebb, 1949). Although the need for description at the level of neuronal circuits is generally acknowledged, different researchers and schools use different terms and emphasize different aspects of these circuits (e.g., cell/neuronal assembly, cognit, action perception circuit, synfire chain, neuronal avalanche, see Braitenberg, 1978; Buzsáki, 2010; Fuster and Bressler, 2012; Garagnani et al., 2009b; Harris, 2005; Huyck and Passmore, 2013; Kaplan et al., 1991; Palm et al., 2014; Plenz and Thiagarajan, 2007; Wennekers et al., 2006). Structural definitions of cell assemblies focus on their strong reciprocal internal connection strength (Braitenberg, 1978), whereas functional definitions emphasize their correlated activation (Buzsáki, 2010; Plenz and Thiagarajan, 2007). Structural definitions of neuronal circuits can be related to graph theoretical descriptions of area connectivity, although there are two important differences: neuronal circuits operate at finer scale (neuronal ensembles instead of areas) and their structure strongly depends on functional correlation and on learning.

The interplay of genetic and stochastic processes in ontogenesis leads to some cortical neurons being more strongly connected to each other than to other neurons (Perin et al., 2011), and therefore it is reasonable to see some (local) neuronal assemblies as being 'built into' the brain. However, the genetic code with its ca.  $6 \times 10^9$  base pairs does by far not include enough information to specify the more than  $10^{14}$  connections between the more than  $10^{10}$  cortical neurons (see Section 2.1 and Haug, 1987). Therefore, much critical information must be acquired through experience. The epigenetic formation of cell assemblies as a consequence of both pre-existing connections and correlation learning provides an established mechanism for such information storage (Pulvermüller, 1999). If a pool of cortical cells are connected and activated together, their mutual connections strengthen and the pool becomes strongly linked functionally. With increasingly stronger connections, a functional nonlinearity is reached so that partial activation of the pool leads to the activation of the majority of the assembly members, a process sometimes called 'ignition' (Braitenberg, 1978; Kaplan et al., 1991). Associative correlation learning implies the formation of such functional circuits in the larger network of the cortex and simulation studies confirm their emergence in networks with different degrees of structural similarity to cortex (Huyck and Passmore, 2013; Kaplan et al., 1991; Knoblauch and Palm, 2001; Palm, 1982; Willshaw et al., 1969; Zipser et al., 1993). Neurons within circuits can be connected in sequences, thus forming synfire chains or avalanches, which generate specific spatio-temporal activation patterns and high-frequency rhythms (see von der Malsburg, 1999; Wennekers and Palm, 2009). If more and more cell assemblies are 'learned into' an auto-associative network, there is danger of explosion like spreading of activation across the network and 'epileptic' activation (Milner, 1996); on the other side, there is the possibility that activity ceases if average weights are too low. Therefore, regulation mechanisms are required to keep the network at criticality, in-between the ceasing and overshooting of activation (Braitenberg, 1978). Local cortical inhibition along with subcortical competition mechanisms involving basal ganglia and thalamus contribute to such regulation of activity in cortex (Palm et al., 2014). Activity regulation and control keeps network activity at criticality and

thereby optimizes information processing in networks containing strongly connected circuits (Shew and Plenz, 2013). The formation, dynamics and functional interaction of 'mini-circuits' in the larger network (cortex or entire brain) may be crucial for language mechanisms, for storing speech sounds, syllables, words and larger constructions, and possibly for mapping further types of linguistic information. Before discussing possibilities for progressing towards a more explicit mechanistic model of language, known key features of language at the behavioral level will be in focus.

### 2.3. What is special about human infants?

When discussing specific features of human language, aspects of syntax and grammar usually come up first. However, in human ontogeny, the earliest behavioral and perceptual patterns related to language and communication by far precede syntax development. These features require specific neurobiological explanation.

Already shortly after birth, infants are able to discriminate speech sound types along phonetic categories (Eimas et al., 1971). However, the ability of categorical perception is shared with animals, for example monkeys and chinchillas (Kuhl, 2004; Kuhl and Miller, 1978), and thus may rely on neural mechanisms not specific to human language processing (Steinschneider et al., 2013, 1994). As mechanisms for imitation are crucial for language (Rizzolatti and Arbib, 1998), reports about newborns' imitation of tongue, face, head and finger movements (Marshall and Meltzoff, 2014; Meltzoff and Moore, 1983; Nagy et al., 2005, 2013) are of great relevance. In parallel, infants show surprising sensitivity to immediate syllable repetition (Gervain et al., 2008). A human predisposition to repeat appears to contrast with reduced imitation abilities of other higher primates ('apes do not ape', Call and Tomasello, 2007). However, some arguments have recently been raised against a strictly innate human ability to repeat body movements (Cook et al., 2014). In addition, repetition abilities have not only been reported in human infants, but also in some monkey species (Ferrari et al., 2013), thus indicating that, similar to phoneme perception, this feature may not provide a unique key to human language. A third early feature of human development is the spontaneous production of speech-like utterances, single vowels within the first months of life and later whole syllables (Locke, 1993). The so-called 'babbling phase' in the second half-year of human life starts with repetitive articulations of different syllables. Surprisingly, initial 'babbles' are similar in babies raised in different language environments, and even appear in deaf infants raised by deaf parents, which argues in favor of a strictly inborn mechanism (Lenneberg, 1967; Locke, 1993). Babbling is not restricted to the articulatory domain, but includes repetitive hand and body movements, too ('manual babbling'). Interestingly, it is towards the end of this phase that the infant starts to repeat gestures, syllables and words produced by others and also produces learned gestures and words spontaneously (see also Section 2.4). This sequence suggests (but doesn't of course in itself prove) that babbling is key to the emergence of repetition abilities in human infants and that both are essential for language development (see Section 2.4, Locke, 1993; Locke and Pearson, 1990).

There is no evidence for similarly rich early repetitive activity in non-human apes, although simplified babbling-like behavior characterizes few monkey species with elaborate, almost human-like social structures (Elowson et al., 1998); the 'sub-song' stage of songbird ontogeny (Aronov et al., 2008) also offers similarity with human babbling. Although each of the putatively 'specifically human' features is shared at least to a degree with at least some species, the triplet of categorical phonetic perception, rich babbling and profound repetition ability seems to be unique to humans' early ontogeny.

Languages differ in their speech sound inventories, form-meaning mappings, combinatorial conventions (Dunn et al., 2011) and pragmatic functions (Wierzbicka, 2003, 1985). In the phonological domain, the exquisite production skills, perceptual sensitivities and discriminatory abilities the infant is equipped with are subsequently tuned by the speech sound types established in the language of the environment, the *phonemes* of that language. Other sounds get lost as they are not used in the language of the environment (Werker and Tees, 1999). This leads to the observation that, for example, English babies initially babble sounds falling into the phonemic category /y/, but then lose this sound during the babbling phase, so that adult monolingual English speakers are unable to produce it. In contrast, babies raised in French, German and Finnish speaking environments, where the /y/ phoneme is established, continue to use it. At about the same time, phoneme production is paralleled on the perceptual side. Instances of the same speech sound type (e.g. different versions of an /u/) become more difficult to discriminate from prototypical variants of the same phoneme, a phenomenon called the 'perceptual magnet effect' (Kuhl, 2004; Kuhl et al., 1992). Even this perceptual prominence of prototypical phonemes has a correlate in production, as, at later stages of babbling, babbles are morphed towards that language's prototype phonemes (de Boysson-Bardies and Vihman, 1991). This language-specific parallelism between production and perception is a core feature of early human communication (Locke, 1993; Werker and Tees, 1999) and its explanation and relationship to the mechanisms underlying human babbling, repetition ability and phonemic discrimination appear essential for understanding language mechanisms.

### 2.4. Toward building a language brain: neural reuse of action and perception mechanisms

For better understanding the mechanisms underlying the formation of neuronal circuits for language, it is advantageous to build mathematically precise neurocomputational models that resemble the anatomy and function of the brain (Arbib, 2016; Arbib et al., 2000; Chersi et al., 2010; Garagnani et al., 2008; Giese and Rizzolatti, 2015; Guenther and Vladusich, 2012; Keyser and Gazzola, 2014; Kilner et al., 2016; Perlovsky, 2011; Weber et al., 2006; Westermann and Reck Miranda, 2004). The epigenetic emergence of language mechanisms can be modelled using artificial neuronal networks, which adopt the structure of relevant parts of the 'language cortex' and operate according to established neurophysiological principles. This modelling enterprise guided by brain theory may show and explain how and why mechanisms that primarily carry motor and sensory functions in animals are altered and enhanced in humans to allow for their neural reuse in the service of language and other higher cognitive functions (Anderson, 2010, 2015).

A first step towards understanding language mechanisms may be taken by modelling the six frontotemporal perisylvian areas in Fig. 2, along with features of within- and between-area cortical connectivity, nonlinear activity summation, habituation, correlation learning, and inhibition and competition processes. To simulate the earliest language-like activities, the production of babbles and early words, and their brain-internal consequences, specific neuron sets in the articulatory 'area' can be activated; these activations can be seen as a model equivalent of the processing of information about the articulatory-phonetic features of one specific syllable or word form. Because articulations lead to auditory signals, which stimulate the auditory system, the model simulation also includes the activation of a second set of neurons in the primary auditory area, which codes for the acoustic phonetic features of the self-produced and perceived acoustic sounds.



Note that, in the real brain, motor commands in the primary motor cortex and consequent acoustic feedback activation follow each other with slight delay; however, as most syllables and monosyllabic words last 200–400 ms, there is reason to assume substantially overlapping activation times in frontal and temporal areas when syllables and word forms are mapped. Studies using transcranial magnetic stimulation show rapid signal transmission from motor cortex to the muscles (e.g., ca. 20 ms delay between cortical magnetic stimulation and hand muscle activity, [Rossini et al., 2015](#)) and rapid auditory activation after sound onset (latency ca. 20–30 ms, [Eldredge and Miller, 1971](#); [Rupp et al., 2002](#)). In addition, cortico-cortical axonal conduction delays between frontal and temporal cortex are in the range of tens of milliseconds ([Matsumoto et al., 2004](#); [Pulvermüller et al., 2003](#)), with the fastest axons bridging this distance in only few milliseconds and the slowest taking hundreds (for review, see [Miller, 1996](#)). Together, this rapid activation succession and the variability of latencies imply that, when a syllable is spoken or a word uttered, there is substantial overlap of motor and adjacent frontal activation dynamics with auditory and adjacent temporal neural activity, thus offering ample opportunity for mapping simultaneous co-activation as well as near-simultaneous spatiotemporal patterns of activity.

The implication of these observations is that near-simultaneously active neurons distributed across frontal and temporal cortex become strongly linked to each other and build circuits of strongly connected neurons, exactly the type of circuit Hebb envisaged and called “cell assemblies” ([Hebb, 1949](#)). Because correlated action-related and perceptual information, manifest as activity in motor and sensory neurons, drives the formation of these circuits, they will be called ‘*action perception circuits*’ or APCs. APCs may not only provide a basis for syllables and word form processing but for the processing of object and action related information more generally, and, in particular for that of meaningful symbols and communication (Sections 2.8, 3.2, 3.3; [Fuster, 1995](#); [Fuster and Bressler, 2012](#); [Pulvermüller and Fadiga, 2010](#); [Lansner, 2009](#); [Strijkers & Costa, 2016](#)).

In the neurobiologically-constrained simulation studies, motor and auditory activations at the opposite ends of the network led to travelling waves across perisylvian areas and to a complex but specific pattern of neuronal co-activations, which strengthened the connections between partaking neurons within and across areas. Correlation mapping based on biologically realistic synaptic learning rules (Section 2.2) linked together the frequently co-activated artificial neurons in different network areas, which thus were bound into one strongly interlinked cell assembly or APC. After learning a repertoire (‘lexicon’) of articulatory-motor patterns and their correlated auditory-sensory patterns, the network was able to reconstruct the activation of each learned ‘lexical’ circuit, so that part of specific articulatory patterns were retrieved upon presentation of the corresponding auditory inputs ([Garagnani and Pulvermüller, 2013](#); [Garagnani et al., 2008](#)).

The summarized simulation studies focused on the role of frontal and temporal areas in processing action-related and acoustic information. The role of parietal cortex in processing somatosensory information about speech gestures was ignored in the interest of keeping simulations simple and manageable. However, to obtain a fuller picture of phonological and language processing, it will be necessary to include somatosensory and other parietal cortices into the models (see, for example, [Ostry and Gribble, 2016](#)). Parietal components of APCs may be relevant at the phonological and semantic levels.

The formation of APCs has a range of functional consequences. (1) *Information mixing*: the formation of distributed APCs implies the emergence of neurons that are active during both production and perception processes, because of the long-distance

transcortical links. Although these neurons may have been unimodal before learning, APC formation makes them carriers of multimodal information. (2) *Spatiotemporal specificity*: the net effect of circuit activation is correlated and well-timed activity across several areas, including sensory, motor and multimodal connection hub areas. (3) *Circuit level of explanation*: The occurrence of such conjoined and correlated ‘activation of areas’ becomes explainable as an epiphenomenon of distributed circuit activity ([Garagnani et al., 2017](#)). These issues will be discussed in more detail below.

Whereas before learning, auditory stimulation had minimal effect in the network, the same stimulation led to substantial network excitation, specific circuit activation and motor pattern retrieval after sensorimotor learning had taken place ([Garagnani et al., 2007](#)). This model of sensorimotor learning suggests that trans-area associative learning and binding of acoustic phonetic and articulatory phonetic information may explain why it is toward the end of the babbling phase that infants become able to repeat gestures, syllables, and eventually words and larger constructions spoken by others. The repetition ability requires that action-perception connections exist in the first place, and that a lexicon of syllables and word form representations, each including specific articulatory-acoustic links, has been set up. These links are realized in the model as discrete sensorimotor circuits each of which binds information about one specific type of articulatory action sequence and concordant auditory perception.

It is noteworthy that pure associative co-occurrence learning of articulatory and auditory patterns was not successful in producing distinct and discrete cell assemblies for different word forms when these overlapped to a degree, that is, when they shared some of their neurons. Such ‘neuronal overlap’ can be used to simulate the sharing of phonemic features between word forms or other common features of cognitive representations. When overlapping neuronal sets were to be learned by the networks, co-occurrence learning alone led to merging of these circuits into one. In order to guarantee that overlapping cell assemblies do not ‘clump together’ in this way, it is essential to implement the kind of ‘anti-Hebbian’ *dissociation learning* revealed by neurobiological research, in addition to association mechanisms (see Section 2.2, [Garagnani et al., 2009b](#)).

The human language ability may crucially depend on the ability to build large numbers of discrete APCs for gestures, word forms and whole-form-stored constructions<sup>3</sup>; and the formation of large numbers of articulatory-auditory APCs may require large numbers of fibers and a short path length between inferior frontal and superior temporal areas. Therefore, the human-specific enhancement of left frontotemporal connectivity, especially by way of the left-lateralized AF, may be of crucial importance for human language. Empirical evidence consistent with this hypothesis comes from recent studies showing that structural features of the arcuate correlate with the ability to learn new words ([Lopez-Barroso et al., 2013](#)) and the ability of subjects to repeat and memorize novel word-like elements is reflected in the degree of their motor and premotor activation during speech perception ([Szenkovits et al., 2012](#)).

A recent neural modelling study by Schomers and colleagues explored the functional implications of the evolutionary neuroanatomical change in AF connectivity between monkeys and humans. To this end, the same perisylvian language areas were modeled with different connectivity structure, thereby imitating the ‘monkey’ and ‘human architectures’ of frontotemporal

<sup>3</sup> Other mechanisms are necessary for neuronal processing of abstract combinatorial knowledge (see Section 2.7).

connectivity. The key distinguishing feature was the absence or presence of ‘jumping links’ and the related changes in sensorimotor path length (see Fig. 2, Section 2.1). Results show that associative learning was possible in both cases, but only the human architecture allowed the formation of circuits in which activity could specifically reverberate and therefore persisted over longer periods of time. As maintenance of activity is essential for working memory (Verduzco-Flores et al., 2009; Zipser et al., 1993), which in turn is a requirement for language, this study offers a direct explanatory pathway from the human-specific change in perisylvian network connectivity to circuit dynamics and hence to verbal working memory and language (Schomers et al., 2017).

As babbling is equally manifest in the manual mode, an analogous argument for action perception mapping can be made for the learning of hand gestures and other actions. Interestingly, a prominent branch of the AF reaches posterior inferior temporal cortex (see Fig. 1) of the ventral visual stream (Rilling et al., 2012). In parallel to the amplification of auditory-motor links, an elaborate machinery for visuomotor association therefore becomes available in humans. Manual babbling in deaf children follows the same time course as articulatory babbling in hearing children, with babbles becoming gradually more similar to words or signs used in communication and social interaction (Locke, 1993; Petitto and Marentette, 1991).

Although it appears to be useful to approximate action perception mapping by neurobiologically constrained simulation studies, the currently available studies can only be considered a start in this endeavor. Much more detail and precision of predictions could be added by working out further structural and functional features of such networks. For example, the precise emergent spatiotemporal activation patterns and oscillatory dynamics of APCs can be targeted (Garagnani et al., 2017; Tomasello et al., 2017).

Considering spike-timing-dependent plasticity (STDP), Keyzers and colleagues argued that, in sensorimotor learning, where fronto-central action-related activity and posterior/sensory excitation follow each other with slight delay, only frontal-to-posterior connections would strengthen, whereas the reverse link between sensory and motor information related to the same motor act would typically weaken (Keyzers and Perrett, 2002; Keyzers & Gazzola, 2014). This postulate rests on the assumption of long motor-to-sensory activation delays (in the order of 200 ms and longer) and lacks consideration of axonal conduction delays and multiple corticocortical fronto-posterior pathways (through ventral and dorsal and “jumping” links). Considering long syllable duration (200–400 ms) compared with short motor output and auditory input delays (ca. 50 ms) along with the short and variable frontotemporal axonal conduction delays, there is ample opportunity to map pre-then-post synaptic activity in both frontal and temporal areas during syllable and word learning. Still, Keyzers and colleagues make a very important point in drawing attention to STDP in context of action perception learning. The frontal-to-temporal activation delay of action and action-perception related information processing, even though it may be minimal and variable, suggests that different long distance fibers and pathways have gradually different roles in fronto-temporal vs. temporo-frontal connections of language circuits. Presynaptic-frontal immediately followed by postsynaptic-temporal activity is preferentially carried by fast frontotemporal connections, which likely take advantage of fast-conducting thick myelinated axons and the short path length of the dorsal AF. In contrast, for strengthening the reverse temporal-to-frontal synapses by STDP, where a slight delay is necessary to achieve millisecond simultaneity or the pre-then-post pattern, the slower-conducting axons with smaller axonal calibers and the multiple-step next-neighbor-area connections through the ventral extreme capsule may be relatively

more important. Furthermore, as these authors state, predictive-sequential links between a motor act representation to its successor action circuit may well rely relatively more on temporal-to-frontal connections (see also Sections 2.6 and 2.7). Considering the interplay of STDP, axonal conduction times and network topologies opens new exciting perspectives for advancing biologically-constrained neuronal models of language and action.

Sensorimotor representations (APCs) for speech units, syllables and word forms offer an account of key features of early human development. The action perception model explains why the ability to repeat not only follows upon, but apparently requires babbling. For example, infants unable to babble – due, for example, to temporary tracheostomy – show abnormal articulations subsequently (Bleile et al., 1993; Locke, 1993; Locke and Pearson, 1990), children with severe neurological motor deficits that affect articulation show reduced auditory vocabularies (Bishop et al., 1990), and a reduced rate of babbling is seen in infants who later develop autism spectrum condition, a deficit characterized by severe communication disorder (Patten et al., 2014), along with abnormalities in frontotemporal connectivity in the arcuate fascicle (Catani et al., 2016; Fletcher et al., 2010; Joseph et al., 2013; Moseley et al., 2016). These observations further bolster the importance of functional interaction between linguistic action and perception information. Still, this evidence should not lead one to deny the importance of sensory, purely perceptual learning in language acquisition (Saffran et al., 1996; Werker et al., 2007), which, however, appears equally powerful in monkeys or starlings (Gentner et al., 2006; Hauser et al., 2001). Evidence for different functional consequences of perceptual and action-based learning comes from neuroimaging. In neuroscience experiments, perceptual learning alone and articulatory-perceptual learning based on babbling-like repetition lead to different neurophysiological correlates of novel ‘words’ in the human brain (Pulvermüller et al., 2012b). While being consistent with the APC model, this finding challenges views denying a role of articulatory mechanisms in speech perception (see Section 2.4.1).

The gradual coupling of action and perception mechanisms during the babbling period provides a neurobiological foundation for the well-known developmental parallelism in speech production and perception noted in the previous section (Werker and Tees, 1999). In particular, the narrowing and tuning of phonological articulations to the phoneme repertoire of the environmental language (de Boysson-Bardies and Vihman, 1991) is addressed by the postulated mapping of arbitrary articulations to their corresponding self-produced auditory percepts, followed by selective activation of specific articulatory variants by the auditory speech sounds most common in the environment. Evidently, such explanation requires specific mappings between acoustic phonetic sequences spoken in the environment and the articulations of the infant, so that frequently heard syllables and speech sounds activate their corresponding motor programs and therefore become established and tuned, whereas non-perceived units leave their babbling-induced circuits unused so that they can fall victim to correlation-based disintegration. A remarkable neuronal model of such neuronal selection and action perception tuning explains, and imitates in detail, the gradual adjustment of infant articulations to different language-specific phoneme systems (Westermann and Reck Miranda, 2004). Furthermore, it has been shown that perceptual classification of sounds into vowel categories can be learned by Hebbian mechanisms based on language-specific perceptual input alone (Vallabha et al., 2007) and that the emerging difficulty in discriminating between variants of a phoneme (magnet effect) is likewise open to an explanation in terms of perceptual learning (Guenther and Gjaja, 1996). Such perceptual mechanisms contributing to the

environmental tuning of infant language are equally relevant for models of APCs interlinking perceptual and motor representations.

#### 2.4.1. Mirror neurons, action perception circuits and language

As recent discussion (for example Caramazza et al., 2014; Hickok, 2014) has questioned the theoretical relevance of mirror neurons (Rizzolatti et al., 1996) and related sensorimotor cells (Bruce and Goldberg, 1985) in monkeys and humans (Mukamel et al., 2010) for cognitive processes, it seems essential to briefly address their significance for brain theory in general and for language mechanisms in particular.

That neurons in the motor system become active during the perception of actions (including both mere motor movements and goal-directed actions, Rizzolatti et al., 2014; Rizzolatti and Craighero, 2004) is a crucial prediction of models postulating functional links between action and perception mechanisms. It is therefore justified to claim that the experimental proof of mirror neurons provides important support for the existence of cortical circuits interlinking the processing of information about actions and perceptions. From this perspective, the discovery of mirror neurons in the premotor cortex shows that these APCs reach into the motor system. Explicit neuronal network models explain the functional coupling, and therefore ‘mixing’, of information processed by sensory and motor neurons. Thus, information mixing would be ‘done’ by the formation – and later-on activation – of APCs, which mirror neurons are part of (Chersi et al., 2010; Garagnani et al., 2008; Giese and Rizzolatti, 2015). Therefore, what had been a ‘modality-specific’ motor, or sensory, neuron before learning (because of its afferent in- or efferent output connection) will become a neuron with ‘amodal’ – or rather multi-modal – sensorimotor processing capabilities after circuit formation.<sup>4</sup> Correlation learning across pre-established and partly genetically determined connection pathways offers an epigenetic mechanism for sensorimotor and mirror neuron formation for elaborate repetition (already in manual and verbal babbling) and for acquiring a rich set of language- and culture-dependent signs (Barsalou, 2013; Keysers and Perrett, 2004; Pulvermüller, 1996; Pulvermüller and Fadiga, 2010).

Note that even if selected mirror neuron links were fully determined genetically and established already shortly after birth, these alone cannot suffice for building the flexible repertoire of actions specific to different languages (see Section 2.3). Not only do the specific articulatory-phonological actions in part differ between languages, but, crucially, the language-specific and largely arbitrary binding between symbolic form and meaning requires a mechanism for learning (Section 3.2). Genetic and experience-independent mirror mechanisms are therefore not sufficient as an account of language. Likewise, associative learning accounts not considering the relevance of pre-established genetic information (see, for example, Cook et al., 2014) fail to explain the specificity of language and repetition mechanisms to the human brain (for discussion, see Kemmerer, 2015b; Pulvermüller et al., 2014b). Therefore, an epigenetic perspective is necessary taking into account the relevance of both the general neurophysiological principles of associative (and dissociative) learning along with

genetically determined human-specific connectivity for mapping the correlational structure of linguistic actions and perceptions (Ferrari et al., 2013; Keysers and Gazzola, 2014; Pulvermüller and Fadiga, 2010). In this perspective, both features, associative/dissociative learning and pre-established pathways, are required. Correlation learning within pre-established connection pathways provides the epigenetic mechanism for mirroring and action perception mapping for language and communication. Put in slightly different words, a “generic structure for matching action observation on action execution” (Rizzolatti and Arbib, 1998) is necessary, one that is specifically tuned to the auditory articulatory domain of spoken language but also covers the visuomotor link required for sign language. This link is provided by the human-specific rich dorsal connectivity between inferior frontal and superior and inferior temporal areas immanent to the human arcuate fascicle. By associative learning of self-produced and perceived actions and by perceptual learning of actions performed by others, action perception links are selectively being strengthened and flexibly adjusted to specific languages and interaction schemas.

Some criticisms of mirror neurons culminate in the statement that their behaviors can be explained solely by associative learning (e.g., chapter 6 in Hickok, 2014). Is this a valid criticism and, if so, what would be a feasible theoretical alternative? The human cortex can be seen as a device evolved and optimized for associative learning and mapping of correlations (Braitenberg and Schüz, 1998), which, in turn, is helpful in predicting the future (Friston, 2010; Fuster and Bressler, 2015; Kilner et al., 2007; Pickering and Clark, 2014). Language learning is most efficient in humans and it appears that specific neurobiological features contribute to this capacity (Sections 2.1–2). It appears that the capacity to associatively learn a large vocabulary of syllables and words is a unique feature of humans and the ability to associatively but specifically link tens of thousands of symbols and their specific meanings is one more unique human feature. Stating that perceptual motor system activity is ‘just’ a consequence of associative learning seems to miss this important point. The relevant insight is that, if not only ‘associations’ are mapped but instead correlations, the ‘just’ requires replacement by an ‘even’. Any explanatory cognitive theory needs to take into account the brain’s most powerful learning mechanism and its implications for the mechanisms of cognition. That learning of correlations seemingly works efficiently during sensorimotor activity (babbling) and equally in social interaction (the laughing infant being laughed back at by caretakers, Heyes, 2010) is of putative importance for communication and sets apart humans. The question of evolutionary interest is why associative learning was brought to such fruit in the domains of language and communication in humans. The tentative answer arising from this review and related neurocomputational work is that an evolutionary change in neuroanatomical connectivity structure answers this important question at least in part (Schomers et al., 2017).

In sum, the specific activation of mirror neurons during action perception confirms a new prediction of accounts postulating that action and perception mechanisms are tightly interwoven with each other at the functional level. Proposals in the cognitive modular tradition (Caramazza and Coltheart, 2006; Fodor, 1983), according to which perception and motor mechanisms reside in different functionally separate processing components, do not make this prediction (for discussion, see Pulvermüller et al., 2014b). Therefore, these latter cognitivist, or even ‘modularist’, traditions have a hard time accounting for mirror neuron activity during action perception and some ‘cognitive’ activations in the motor system more generally (see Section 3). A model with modules for action and language perception would not predict motor system activations during perception and suggests to

<sup>4</sup> Even after action perception learning, one may want to consider the neurons ‘close to’ the motor output (and those to the sensory input) to remain “essentially” motor (sensory) (see, for example, Barsalou, 2013). An entirely different framework suggests that, irrespective of the mixed-and-merged information they process, each neuron retains its individual “representational format” (here motor or sensory). (Mahon, 2015; Mahon and Caramazza, 2008). However, given these neurons activate indeed similarly in motor and sensory tasks, a motivation for such a position could only come from ontogeny or from the possibility of ‘cutting’ and splitting the circuit parts again – although even then, the motor feedback to ‘sensory’ neurons may be functionally relevant.



‘explain’ them post hoc as ‘epiphenomenal’ or ‘spill-over’ activation of no functional value. Within distributed circuits interlinking action and perception information, sensorimotor neurons have functional roles similar to those of other circuit member neurons (see Section 3).

A justified criticism of some previous statements is that the motor system and its mirror neurons are not *the only, or most important*, locus of perception and understanding processes.<sup>5</sup> As the link between sensory and motor neurons is mediated by neurons in connector hubs, it cannot be claimed that the resultant distributed system is functionally dominated by one type of neurons in only one area (for example model simulations, see Garagnani et al., 2017, 2008). Thus, mirror neurons and motor cortices would not provide the only or the most important locus for action or language perception or understanding. Within distributed neuronal circuits, each neuron makes a functional contribution to the whole. This functional contribution depends on circuit size, the connectivity strength between the neuron and other circuit members, and the strategic placement of the neuron within the circuit. Due to redundancies, dysfunction may only appear with substantial lesions. Lesions in the topological center of a circuit (its ‘kernel’) will generally cause more pronounced dysfunction than lesions in its periphery (its ‘halo’, see Section 2.6). Lesions at the motor end will impact on motor mechanisms more than on the circuit-internal feedback activation waves by which motor neurons contribute to perceptual activations reaching the sensory end of the circuit. These well-known arguments from old literature on aphasia (e.g., Freud, 1891) were re-evaluated in neuropsychology (Shallice, 1988) and backed by neural network simulations (Plaut and Shallice, 1993), which can explain both associations and dissociations of deficits caused by focal brain lesion, for example aphasia symptoms (Pulvermüller and Preissl, 1991). Crucially, they show that focal lesions in distributed neuronal circuits interlinking action and perception mechanisms can explain the correlation of abilities to perceive and act in large patient groups, and equally account for double dissociations between predominant perceptual and action impairment following lesions in sensory and motor areas. Therefore, one of the strongest inference schemes proposed by mirror neuron critics, that lesion-induced double dissociations between action and perception falsify the action perception perspective, is in fact invalid (Pulvermüller and Fadiga, 2010).

The main discussion of the role of mirror neurons and action perception circuits has moved away from the strong statement that motor activity is ‘epiphenomenal’ to perception and understanding, to a more moderate claim that motor systems serve a “modulatory, not primary role” (Hickok, 2009, 2014). The problem with this type of proposal is that it lacks clear criteria. In Hickok’s hands, a small impairment in perception caused by frontocentral lesion is seen as evidence for an at best minor and “modulatory” role of the mirror neuron system, whereas a gradual perception deficit resulting from left posterior temporal lesion is seen as compatible with a role of this brain part as a “gateway to understanding” (Hickok, 2014; p. 181). The action perception perspective provides the less biased account: In both cases, dysfunction in relevant parts of action perception circuits may be the cause of a measurable and significant deficit (for further discussion, see Section 3).

<sup>5</sup> Such claims have sometimes been made in the mirror neuron context and in related fields, for example related to the motor theory of speech perception (Section 3.1.3).

## 2.5. Attention and flexibility of activation

Distributed action perception circuits are unitary entities but their activation needs to account for great flexibility. Infants may occasionally repeat words they just heard, but in adults, specific mechanisms seemingly hinder the automatic activation of motor mechanisms in response to sensory input (Jacob and Jeannerod, 2005; Jeannerod, 2006). Likewise, when speaking, less attention is paid to acoustic input compared with a listening condition. Words may be articulated in different ways dependent on previous input and context and likewise the meaning of utterances is influenced by communicative context. At the neurobiological level, influences of context, task and attention can be captured by level and type of pre-existing activity within a circuit (see Sections 2.6 and 2.7) and by general regulation determining the degree of activation and inhibition in the entire network or specific parts thereof.

Cognitive theories of attention, in particular the well-known biased competition model (Duncan, 2006), see attention to objects, words or loci in space as a result of two main factors, the degree to which there is a bias towards them and the degree to which different putative targets of cognitive processing compete with each other. In cell assembly networks, the degree of pre-activation or priming of a circuit (that is, its level of pre-existing activity) provides the neuromechanistic correlate for the bias, and the network’s degree of regulatory excitation or inhibition provides the mechanism for competition (for discussion, see Garagnani et al., 2008).

There is independent motivation for modulatory mechanisms within neuronal circuit models. As discussed in Section 2.4., any strongly interconnected neuronal architecture is in danger of catastrophic over-activation and, on the other hand, extinction of excitation (Braitenberg, 1978; Braitenberg and Schüz, 1998; Milner, 1996). For this reason, activity regulation and control is necessary for processing information in brain-like systems. This can be realized by what has been called a ‘threshold regulation mechanism’ (Braitenberg, 1978), a feedback loop which controls the general level of excitation/inhibition within the whole cortex or in a specific area (Bienenstock, 1994; Elbert and Rockstroh, 1987; Knoblauch and Palm, 2002; Palm et al., 2014; Wennekers et al., 2006). A range of neurocomputational simulations combined background activity (noise), which prevents extinguishing of network activity, with area-specific inhibitory feedback regulation to prevent several ignitions at a time and to control the degree of competition between partly active reverberating cell assemblies (Bibbig et al., 1995; Garagnani et al., 2008; Knoblauch and Palm, 2001; Palm and Sommer, 1995; Wennekers et al., 2006). The gain or amplification factor of the regulation function can be changed per area thus allowing adjustment to task specific requirements.

Adjusting regulation gains differently across areas allows for modelling task-specific processing differences. For example, to model a perceptual or cognitive task, the gain of inhibitory feedback regulation in motor cortex (M1) is adjusted to a high value, to reduce motor activity for preventing overt action. Likewise, the inhibitory gain is specifically increased in perceptual areas (including A1) for modelling motor tasks, to limit perceptual input. The difference between a language task and one in which attention is directed away from language is implemented by high vs. low gains (weak vs. strong mutual local inhibition) within the six areas of the frontocentral perisylvian language cortex (Garagnani et al., 2008). For implementing semantic tasks, feedback-inhibitory gains are set low in some or all semantic areas (see Section 3.2). More generally, any specific task can be modeled as a gain function across model areas.

To illustrate the flexibility of brain activity elicited by meaningful words and matched meaningless pseudowords, please consider a concrete simulation study of attention to language,

which uses the architecture introduced in Section 2.4. This model, proposed by Garagnani and colleagues (Garagnani et al., 2008), simulates the recognition of a word as ignition of its respective circuit upon auditory input. As ‘pseudowords’, that is, novel meaningless spoken syllable sequences, would not be represented in the brain and thus would not have specific corresponding APCs, their perception was modelled as partial stimulation of auditory representations of several words, which typically failed to elicit full ignitions. Attention-to-language was modeled, as indicated above, by reduced vs. increased inhibitory gain in perisylvian model areas, thereby realizing more vs. less attentional resources. Interestingly, strong attention-related variability of brain responses to pseudowords was seen: Strong brain responses were elicited with ample attentional resources (mild inhibition), whereas weak responses occurred when only limited attention resources (strong inhibition) were present. This big difference between strong vs. weak activations in attend and non-attend conditions was not seen for words, because the process of ignition and subsequent back-regulation of activity was shared across attention conditions. The shared ignition-inhibition pattern stabilized the network response to words, so that attention-related modulation of excitability had only a reduced effect. These simulation results led to the experimental prediction of a double dissociation of the word- and pseudoword-elicited brain responses in attend and non-attend conditions (Garagnani et al., 2008). Garagnani’s model predictions were in good agreement with subsequent EEG and MEG studies (Garagnani et al., 2009a; Shtyrov et al., 2010). Garagnani’s model explains the differential task-related effects on word and pseudoword processing as a result of attentional gain and the ignition process implied by cell assembly dynamics, which, due to its constancy and stabilizing role, counteracts attention-related changes in activation (Garagnani et al., 2008). This network implementation illustrates how task and attention effects are modeled in the APC framework.<sup>6</sup>

## 2.6. Multiple cognitive demands: memory, intention, prediction

APCs associate sensory and motor information and provide a mechanism for perceptually-tuned babbling, repetition and categorical speech perception. Could these same APCs explain the neural reuse of sensorimotor mechanisms for higher cognitive functions such as memory, decisions, attention and intelligent behavior (Gallese and Lakoff, 2005)? In apparent contrast with this suggestion, areas far removed from sensory or motor fields, in prefrontal, parietal and temporal association cortex, are crucial for a range of cognitive capacities. Therefore, a *multiple demand network* of multimodal association cortices is postulated, which supports these broad cognitive capacities (Duncan, 2010). Consistent with this position, memory cells, the ultimate neurobiological substrates of working memory, which remain active and store specific content for several seconds, are most commonly found in prefrontal cortex and higher multisensory areas (Fuster, 1995, 2009; Goldman-Rakic, 1995). Likewise, predictive neurophysiological activity preceding speaking and acting seemingly involves prefrontal areas first, before motor cortex becomes active (Edwards et al., 2010). Thus, similar to working memory, intention and decision processes are associated with prefrontal cortex and other multimodal areas (Deco et al., 2013). The great relevance of

multimodal cortices for higher cognition may raise doubts in the relevance of action perception circuits for these functions and has, in fact, led some researchers to argue against models of cognition that emphasize the reuse of action and perception mechanisms (Caramazza et al., 2014; Hickok, 2014).

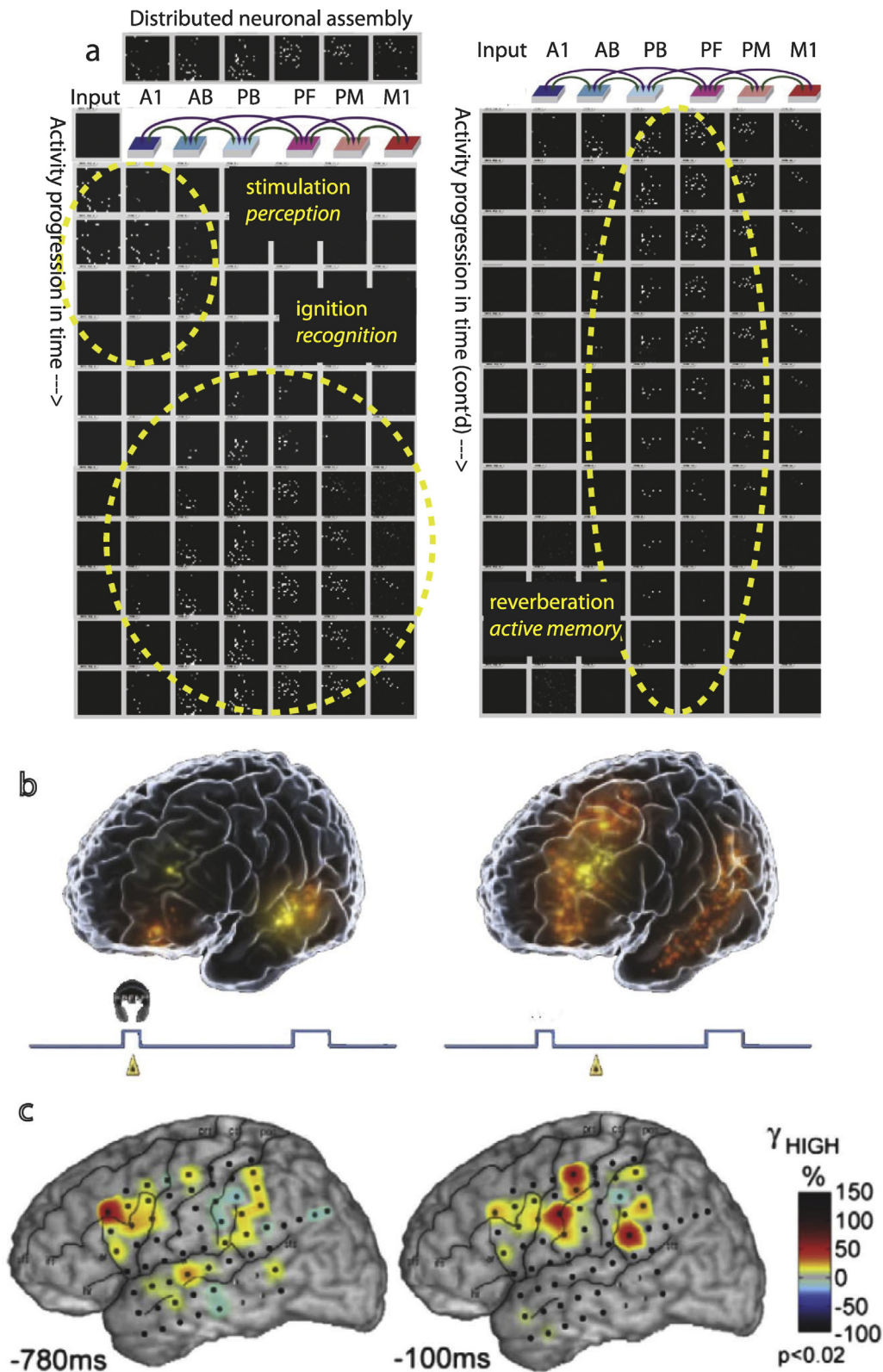
However, from a stringent perspective, the very existence and emergence of a multi demand cognitive system in distributed multimodal areas is in need of neurobiological explanation. This section argues that the APC model is not only compatible with a multi demand cognitive system, but even offers an explanation for its broad function and specific localization.

Neurocomputational simulations of processes in brain-inspired networks offer insights into the mechanisms of cognitive and language mechanisms over and above the level of sensorimotor association (Arbib et al., 2000; Kaplan et al., 1991). They have also been used to model memory circuits and indeed correctly predict the dynamics of memory cells in prefrontal cortex (Verduzco-Flores et al., 2009; Zipser et al., 1993). However, if correlated activity in sensory and motor systems underlies the formation of memory traces, why should memory cells emerge in areas distant from their antecedent activations in sensorimotor areas, thus leading to the ‘disembodiment’ (movement away from sensorimotor systems) of memory mechanisms?

Recent models of the formation of memory circuits (Fig. 2) have addressed this issue, taking advantage of neuroanatomical wisdom about area connectivity in the perisylvian language cortex. Fig. 3a presents the activation dynamics in the ‘areas’, of a model, which had previously ‘learned’ a vocabulary of several ‘words’ and was subsequently stimulated by the auditory code of one specific ‘word’. Note again that, before learning had taken place, the network response to the same stimulation had been minimal. After learning, stimulation-induced activity spread from the sensory area to adjacent secondary and multimodal association areas and, after a little lag, an explosion-like activation across areas occurred, which involved most of the APC, including some of its neurons in articulatory-motor and auditory-sensory areas. This activation illustrates the non-linear ignition process postulated by cell assembly theorists (Braitenberg, 1978; Palm et al., 2014), which can be considered a neural correlate of recognition, that is, word or construction recognition in the domain of language, and action and object recognition otherwise. Following upon ignition, activity reduces across the network, which is, in part, a result of the inhibitory influence of regulation mechanisms (Palm et al., 2014). Notably, reverberating activity persisted in part of the APC, thus providing a network correlate of verbal working memory (object or action memory when other cortical systems were modeled). The majority of these memory neurons were in the topological center of the architecture, in the network parts implementing the cross-modality connection hubs, i.e. the inferior prefrontal and temporal parabelt cortices (Pulvermüller and Garagnani, 2014; Schomers et al., 2017). Therefore, the mathematically implemented action perception model replicates and explains aspects of the neocortical topography of working memory processes, in particular its pronounced manifestation in prefrontal and other multimodal areas (Fuster, 2009).

Why do multimodal areas become important for memory? Clearly, although long-term memory trace formation is driven by correlated neuronal activity in sensory and motor cortex, the lack of direct connections between primary perceptual (auditory, visual) cortices and primary motor cortex makes it necessary for sensorimotor activity to travel through neurons in multimodal connector hubs, which therefore are being incorporated into APCs. Following upon ignition of these circuits, the activity shift away from sensorimotor areas and to multimodal association areas is explained by features of corticocortical connectivity. First, these areas are placed at the interface between sensory and motor

<sup>6</sup> Note that a common misunderstanding of the APC approach had once been that cell assemblies must always activate in the same way (‘automatically’) (see Raposo et al., 2009; Willems and Casasanto, 2011). These views are in contrast with explicit statements to the contrary (Pulvermüller, 1999) and with the results of model simulations illustrating and documenting task and attention effects, such as the study by Garagnani et al., (2008). For further discussion of the flexibility and automaticity issues, see Section 3.2.



**Fig. 3.** Memory and decision in model and brain activation. **a.** Activation dynamics of the model of the language cortex (Fig. 2) after stimulation with a learned 'word'. The sequential processing steps (and their cognitive counterparts) of stimulation (perception), ignition (recognition) and reverberation (active/working memory) are indexed by ellipses. Note that reverberating memory activity retreats to the central connector hubs (prefrontal and parabelt areas, from Pulvermüller and Garagnani, 2014). Cell assembly structure is shown at the upper left. Each small square represents one model 'area', each white dot one active 'neuron'. The leftmost column shows the input to the network and the six squares to the right activity in the six model areas (indicated and labeled at the top). Simulation time runs from top to bottom. **b.** Brain activity during speech stimulation and verbal working memory as summarized in a previous review (Fuster, 2009). Please note the pronounced inferior frontal memory related activity. **c.** Electroencephalography recordings of brain activity when subjects decide on uttering a word to be produced (Canolty et al., 2007; Garagnani and Pulvermüller, 2013). Please note the activity progression from prefrontal to central cortex.



systems and therefore, as mentioned, hold a strategic position as connector hubs within the language network. Second, further topological features of the human language cortex are relevant, especially the multimodal areas' high 'degree' of connectivity (the number of connections of these areas to the other relevant areas) and the shorter 'path lengths' between sensory and motor neurons implied by 'jumping links' (Fig. 2c). As already discussed in Section 2.4, these topological features play a causal role in verbal working memory emergence (Schomers et al., 2017). Now, due to the prefrontal and parabelt cortices' role as connector hubs with high degree of connectivity, their neuron members are more likely to be strongly linked to other members of their APCs and, therefore, to be part of the most strongly connected 'kernels' of APCs. Since kernel neurons are best connected, their reverberation lasts longest after cell assembly ignition. Thus, predominance of memory activity in multimodal connector hub areas is accounted for by circuit topology and sensorimotor associative learning.

A parallel argument applies to the putative mechanistic underpinnings of activity spontaneously emerging in a resting or idling network. Such emergent activity has been related to aspects of cognitive processes as different as prediction, intention, decision or 'memory for the future' (Deco et al., 2013; Fuster, 1995; Fuster and Bressler, 2015; Kilner et al., 2007). When a network including APCs is left alone but noise and regulation parameters keep it close to criticality, ignitions will occasionally take place spontaneously; the spontaneously emerging predictive activity and the subsequent ignitions can, respectively, be seen as neurobiological correlates of intentions and decisions to act (Deco et al., 2013). More precisely, minimal activity accumulation significantly above the noise level in neurons of a specific APC can be interpreted as neurobiological mechanism of an 'intention' to perform the action the circuit codes for, and the ignition of the APC can be seen as correlate of the 'decision' to perform the action. Under downregulation of motor activity, emergent activity in APCs can provide a correlate of perceptual predictions. In a study of intention and decision processes using the language model of Fig. 2, intention-related predictive activity was indeed seen primarily in connector hubs, subsequently spreading to second-order areas (premotor and secondary sensory cortex), whereas ignitions related to decision processes spread across all areas, including sensorimotor ones (Garagnani and Pulvermüller, 2013).

Fig. 3c shows brain activation as recorded through subdural grid electrodes just before subjects generate verbs according to their own decision (Edwards et al., 2010). As can be seen, the activation indeed originates in prefrontal and multimodal temporal cortex and then spreads to sensorimotor regions. Similar activation dynamics spreading from association to motor cortex characterize the cortical generators of the readiness potential for speech (Carota et al., 2010). The intention/decision model shows a dynamic topography, which is the reverse of the memory dynamics previously discussed (Garagnani and Pulvermüller, 2013). The neuromechanistic explanation once again draws upon circuit kernels, whose especially strong links do not only support maintenance of activity but also accumulation of specific inputs from other circuits and unspecific background noise, thus yielding future ignitions. Note that the mirror image behavior of the networks seen in the working memory and action prediction simulations is consistent with, and offers an explanation for, 'multiple demands' carried by connector hubs.

In summary, the action perception model offers an explanation why neurons in specific areas take on multiple cognitive demands as variable as memory and decision/prediction processing: Within APCs, the same neurons of the most strongly interlinked kernel support the processes of reverberation (memory) and spontaneous activity emergence (intention/prediction). In APCs built by sensorimotor correlations in networks with realistic connection

structure, these neurons are most likely to emerge in connector hubs due to network topology. Relevant connector hub areas for spoken language are in inferior prefrontal cortex and auditory parabelt in superior temporal lobe (Fig. 2c); a similar argument is possible for posterior inferior parietal cortex if articulation-related somatosensory activity is also considered. Circuit kernels in connector-hub areas become the basis of memory and intention processing and, as discussed below, combinatorial knowledge about sequences of events, actions and symbols.

## 2.7. Combinatorial learning and generalization

Language is sometimes characterized as a system for variable combination of a finite set of small units into a virtually unlimited set of strings. Combination is manifest at different levels: speech sounds combine into word forms, words into larger constructions, and speech acts into dialogues. Neurobiological mechanisms for sequence-building and combination are well-studied, although the links between these mechanisms and specific linguistic phenomena are still not fully understood. For sequencing phonemes into syllables and spoken word forms, synfire chains or neuronal avalanches provide an obvious mechanism (Abeles, 1991; Braitenberg and Pulvermüller, 1992; Buzsáki, 2010; Plenz and Thiagarajan, 2007). However, the linear nature of avalanches appears as insufficient for capturing the hierarchical and temporally flexible structure of sentences and dialogues (for discussion, see Pulvermüller, 2002b). More sophisticated neuronal mechanisms are required.

Neural networks of different types address different aspects of the 'problem of serial order in behavior' (Lashley, 1951), including putative linguistic morphological and syntactic mechanisms. For example, parallel distributed processing models with three layers of artificial neuron-like elements were used to address the processing of past tense morphology (Rumelhart and McClelland, 1987), and specific syntactic structures, even as complex as center-embedded sentences, were modelled with networks that include an additional 'memory layer' (simple recurrent neural networks, Christiansen and Chater, 1999; Elman, 1990; Elman et al., 1996). However, linguists and cognitive scientists raised a long list of arguments against these approaches (Fodor and Pylyshyn, 1988; Jackendoff, 2002). Critically, these distributed 'neural' approaches typically deny the existence of (neural equivalents of) linguistic entities such as phonemes, morphemes, words, or rules of syntax, whereas linguists emphasize the need for *discrete representations* carrying these entities. In this respect, the neurobiological – or 'neuronal' – framework outlined here differs from classic 'neural' distributed approaches. If neuroanatomical and neurophysiological constraints are applied to network architectures and function, neuronal circuits emerge within the larger network structure of the cortex or brain (Palm et al., 2014; Pulvermüller et al., 2014a). These circuits provide a basis for discrete cognitive representations.

Jackendoff outlines problems of neural language models and states that classic parallel distributed processing models fail to capture these (Jackendoff, 2002). The following items are included in his list: (i) There is massive multidimensional *binding* between the linguistic units included in a sentence. (ii) Representations may need to be used more than once when processing a sentence, thus leading to *multiple instantiation* (see also Sougné, 1998). (iii) *Long-term and working memory* mechanisms need to be distinguished.

How would the present neurobiological framework fair regarding these features? Regarding permanent/long-term mechanisms and momentary processing, the difference between language structure and its use is crucial in linguistics, being manifest, for example, in the distinctions between 'langue' vs.

'parole' (de Saussure, 1916) or 'competence' vs. 'performance' (Chomsky, 1965). In one type of model, representations (including words and rules) are located in a 'store' from where they are 'copied' into a 'processor' when sentences need to be assembled and processed. Separate brain analogues of store and processor are not easily distinguished, and any 'copying' between them appears as a metaphoric approximation to reality (for discussion, see, for example, Searle, 1984). The most well-established neurobiological mechanism for long-term storage is the long-lasting modification of the strength of neuronal connections and the consequent formation and alteration of neuronal circuits and the links between them. Working memory can be understood as activation of such circuits, which emphasize circuit kernels (see Section 2.6, Fuster, 1995). Note again that this mechanistic underpinning of Jackendoff's feature (iii) is incompatible with distributed accounts denying discrete neuronal representations, but sits well with neurobiological theories that include discrete representations at the neuronal circuit (cell assembly) level. Still, each circuit may be claimed to be either active or inactive, so that the multiple use, or activation, of the same representation in a given derivation would remain unexplained (see point (ii)). However, this position depends on a 1:1 relationship between neuronal circuits and their activations, which is not a realistic assumption (Pulvermüller, 2010). In fact, network simulations have shown that it is possible to build circuits in which two or more activity waves reverberate and co-exist for some time (Hayon et al., 2005), thus providing a mechanism for multiple activation of the same representation. This allows for activating the same linguistic representations more than once in the understanding or production of a sentence, therefore suggesting a neurobiological perspective on the multiple instantiation problem (ii) (for a grammar model, see also Chapter 12 in Pulvermüller, 2002b).

In previous sections, neurobiological considerations have been restricted to phonemes and morphologically simple words. This is obviously insufficient for modelling combinations between words and morphemes and the meanings these combinations convey. In a neurobiological grammar framework, word form and meaning are bound together by distributed circuits which can, in turn, be bound together by additional circuits. I have proposed that a special type of discrete circuit called '(discrete) combinatorial neuronal assembly', or CNA, provides a neuromechanistic basis for grammatical binding by interlinking words and sentence parts (constituents) with each other, also establishing syntactic-semantic relationships (Pulvermüller, 2002b; Pulvermüller et al., 2013). CNAs may be used to address the grammatical variant of the binding problem (Jackendoff's point (i)). For example, when hearing the sentence "Bob let Sue in", a CNA would link the subject and the verb, a second one the verb with the object, and a further binding mechanism would be available to join verb and particle (Fig. 4a). The links between verb and its subject, object, and particle complements would therefore be represented in non-linear (hierarchical) flat structures resembling aspects of classic grammar descriptions in the dependency grammar tradition (see, e.g., Fillmore, 1967; Goldberg, 1995; Tesnière, 2015). In sentence processing, the multiple binding between sentence constituents would proceed gradually, morpheme by morpheme, with any new meaningful unit first igniting its corresponding APC and subsequently the attached CNA(s), binding it into its sentence context. As a result of serial ignitions of APCs and CNAs, the partaking circuits are 'synchronized', thus forming a coherent unit in working memory.<sup>7</sup> In this model, grammatical relationships of different types (dependency, agreement) and related semantic links are

captured by the same biological mechanism, the binding between APCs and CNAs (Pulvermüller, 2002b, 2010; Pulvermüller et al., 2013).

### 2.7.1. Formation of combinatorial neuronal assemblies

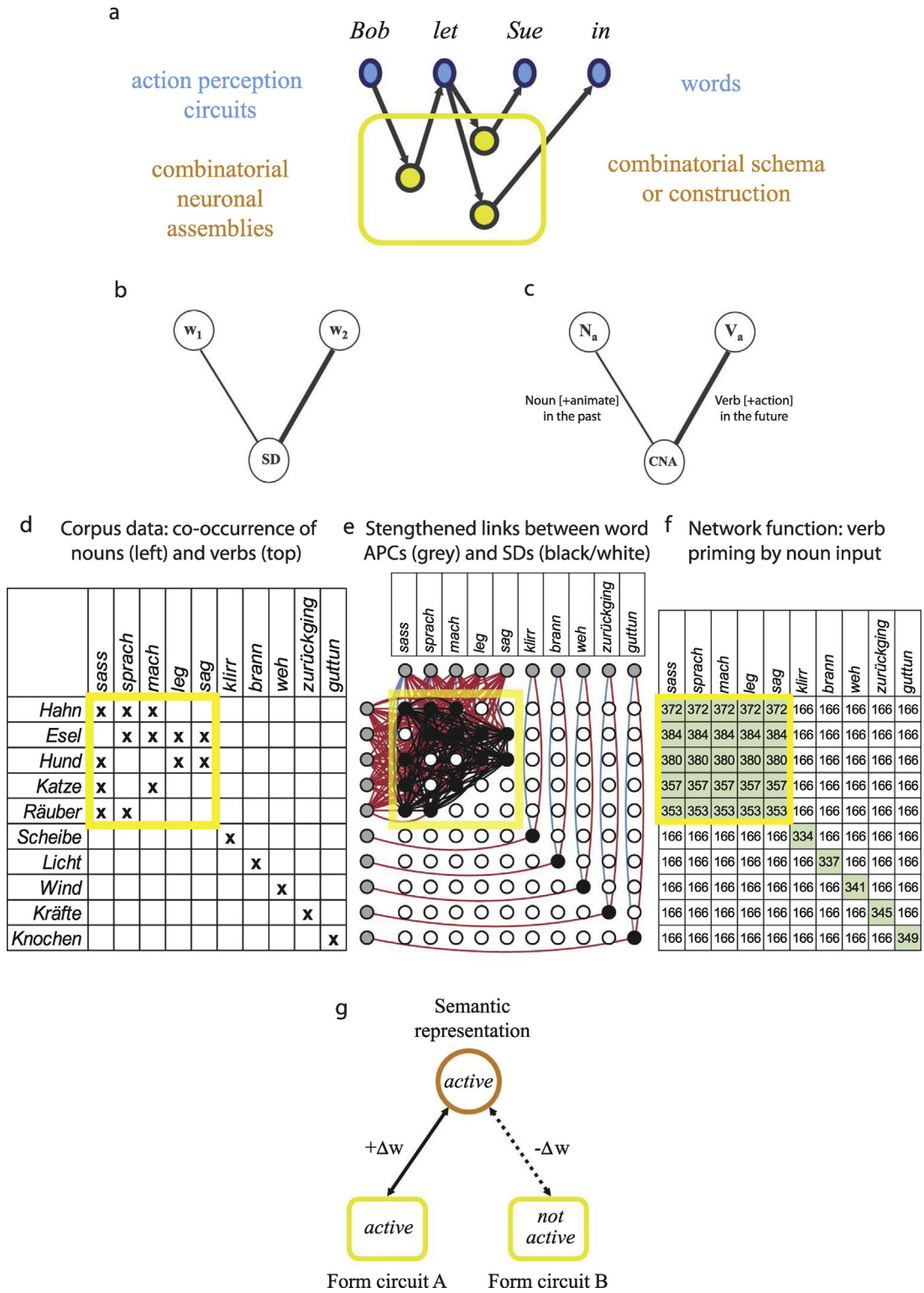
The proposal of a neuronal grammar similar to classic linguistic theories may appear as a trivial restatement of linguistic postulates. However, there are relevant differences demanded by neurobiological constraints. For example, the neurobiological framework does not use tree structures, which prevail in linguistic descriptions of sentence structure. Instead, pairwise combinatorial relationships are implemented between constituents, whereby each item (sentence constituent) may directly connect to several others (as in the case of the verb with subject, object, and particle links, see Fig. 4a). Binary trees would instead allow direct links only between two items (immediately dominated by one mother node), with any further ones connected indirectly. Crucially, a major driving force in the formation of CNAs is correlation learning. This assumption may seem to contrast with theories claiming that major components of grammar knowledge are genetically determined. As grammar learning cannot be explained by standard associative learning, because 'everything may associate with everything' (Section 2.2), it is important to consider in detail how the neurobiological mechanism of correlation learning, which includes both association ('fire-wire-together') and dissociation ('out-of-sync-delink') terms, helps explain the emergence of mechanisms for generalized combination following abstract patterns.

When symbols and words occur together in sentences, their corresponding neuronal circuits become active in temporal vicinity so that their activation functions overlap.<sup>8</sup> APC ignitions follow each other in time and the overlap in activation time is between the persistent memory-activity of the circuit activated first and the ignition of the second. As discussed in Section 2.5, the memory-active kernel parts of linguistic APCs primarily employ perisylvian connector hubs – in inferior prefrontal, auditory parabelt and inferior parietal association cortex. Connections storing combinatorial information about the order of words and phrases in larger constructions and sequences must therefore critically involve such memory active neurons and, therefore, perisylvian hub areas. As, according to the model, both ignitions of APCs of current words and memory activity of kernel-parts of context word APCs are manifest in connector hubs, local between-APC connections in these areas become most relevant for storing combinatorial information.

A typical objection against such 'word-word' associative learning of individual sequences and strengthening of corresponding neuronal connections has been that it is at variance with known linguistic facts and, in particular, cannot explain combinatorial generalization (Bierwisch, 2008; Chomsky, 1959). Why would linguistic-combinatorial principles apply not solely to learned sequences but rather to large word classes and sometimes lead speakers to produce sentences they never encountered before? English speakers know that subject nouns tend to co-occur with predicate verbs, determiners with nouns, adjectives with nouns, and so on, and these 'rules' can be applied to new words. This phenomenon of generalization led linguists to suggest

<sup>7</sup> The multiple use of the same circuit in one derivation is possible in this framework (see Pulvermüller, 2002b, Chapters 10–12).

<sup>8</sup> Although spike-timing-dependent plasticity may suggest 'delinking' of backward connections between the representations of two cell assemblies activated in sequence, known circuit dynamics seem to contradict this. Memory cells and neuronal circuits are active for many seconds (Fuster, 1995), whereas timing-dependent plasticity operates on a time-scale of 50–100 ms (Caporale and Dan, 2008). Therefore, the minimal delays may be relevant for interlinking specific neurons included in partaking cell assemblies, but the near-simultaneity of circuit activations argues for strengthening of links in both directions.



**Fig. 4.** Combinatorial language mechanisms in neurocomputational studies. **a.** Neurobiological binding mechanisms for representing and processing a simple sentence. Action perception circuits (APCs) for words are shown as blue ovals and combinatorial neuronal assembly (CNAs) for word category combinations as yellow circles; the yellow box indicates CNAs which, together, form a combinatorial schema or construction representation. **b.** Illustration of a sequence detector sensitive to one specific sequence of two words and **c.** of a CNA that interconnects lexicosemantic categories (for example animate nouns and action verbs). **d-f.** Combinatorial learning and generalization in a neurocomputational model: **d.** The matrix shows word pair co-occurrences in a text that served as input to the network (verbs in top row, nouns in left column; dots indicate co-occurrences in a text). **e.** Network structure of input words (gray circles) and sequence detectors that had been strengthened by concordant input (black) along with their mutual strengthened connections (black lines) that index the learned CNA. **f.** Priming between nouns and verbs that previously recombined with each other during learning. Numbers give the degree of priming of a given verb-related circuit after activation of a specific (context) noun circuit. High values indicate functional connectivity between the members of a given pair. Note that priming applies also for strings of words including an actor noun and an action verb, which had not previously been experienced in the input. Therefore, the model provides an associative learning account of generalization (adopted from Pulvermüller and Knoblauch, 2009). **g.** Learning of construction semantics by correlation: When the learner understands that a specific semantic content is expressed by using a specific linguistic form (A) or construction, but not an alternative form (B), which would also be available, two processes happen in parallel. The expressed form is entrenched as an instantiation of that meaning (Hebbian associative learning) and the unused form is gradually preempted (anti-Hebbian synaptic weakening and dissociation learning).



that much grammatical knowledge must be stored in the genetic code, as limited experience could not account for the learning of all virtually infinite combinatorial possibilities (sparseness of the input). However, combinatorial structures and word order regularities, which according to linguistic theory, should correlate very well, can vary widely within and across languages (Dunn et al., 2011). On the other hand, ample combinatorial information is immanent to the input any learner receives when being exposed to language use and this input has been shown to lead learners to generalizations, even if arbitrary combinatorial principles are subject to learning (Goldberg, 2006). Modern usage-based approaches to language emphasize such data-driven generalization (Goldberg, 2006; Macwhinney, 2014; Tomasello, 2005). Standard recurrent network models with random connectivity and without built-in neuroanatomical information can extract combinatorial information from language input and produce remarkable generalizations (Elman et al., 1996; Hare et al., 1995), thus demonstrating that associative learning can explain aspects of the learning of what might be called ‘rules’, although (as discussed above) their lack of discrete representations limits their applicability to language.

Most artificial neural networks used to study generalization did not focus on mimicking brain mechanisms. They did apply learning rules that map correlations, but the connection structure implemented in the networks were unrelated to cortical neuroanatomy. However, important clues about combinatorial processing in humans may come from neurobiology. When looking at real nervous systems processing elementary combinatorial information, the concept of a movement detector stands out. **Movement detectors are neuronal units that respond to a sequence  $ab$  of events but not to the reverse sequence  $ba$ .** Thus, they are specific to sequences and can, in fact, ‘detect’ sequences. Such neurons have been found in various animals (insects, amphibians, mammals) and in different anatomical structures (ganglia, retina, cortex) (Barlow and Levick, 1965; Hubel, 1995; Reichardt and Varju, 1959). It seems plausible that humans have similar devices for processing language. Therefore, it appears as reasonable to ask which grammar machinery would fall out naturally if human brain circuits were able to do with words what animal brains can do with elementary stimuli, namely to detect and respond specifically to their sequential arrangement.<sup>9</sup>

Accordingly, there would be neuronal elements that respond specifically to sequences made up by any pair of words. For the sequence of word- $a_i$ -followed-by-word- $b_j$ , there would be a sequence detector circuit  $SD_{ij}$  responding specifically to this sequence of ignitions. If sequence  $a_1$ -followed-by- $b_1$  appears, there is strengthening of connections between the circuits for words  $a_1$  and  $b_1$  by way of the  $SD_{11}$  circuit. Now assume that each of the two words  $a_1$  and  $b_1$  is a member of a group of words and each member of the  $a$  and  $b$  groups, respectively, has a tendency to combine with different members (not necessarily all) of the respective other group. So any  $a_i$  combines with (some, not all) members from category  $B = \{b_1, b_2, \dots, b_n\}$  and, likewise,  $b_j$  with members of class  $A = \{a_1, a_2, \dots, a_m\}$ . As the SD network contains rich auto-associative connections, the learning of any particular symbol pair sequence yields one strengthened connection via one specific SD circuit, and, the more SDs have been associated with a given symbol circuit, the larger the number of SDs that become partly co-activated with this symbol's circuit. Therefore, the net effect of such combinatorial learning is the strengthening of connections not just between the co-activated symbol circuits (by way of the SD specific to that sequence) but, in addition, the co-activation of

increasing numbers of SDs associated with each word representation and therefore formation of an aggregate of SDs. These SD aggregates now become active to any ordered pair sequence  $a_i b_j$  of symbols belonging, respectively, to the two lexical classes  $A$  and  $B$ . As they combine whole sets of symbols with each other, these newly formed neuronal aggregates provide a basis for the CNAs proposed in the neurobiological grammar context. Because whole classes of symbols are combinatorially linked by CNAs, they support retrieval of previously learned combinations as well as generation of new, never encountered ones that follow the generalized  $a_i$ -followed-by- $b_j$  pattern. Correlation-based learning of combinatorial regularities may be an important neurobiological mechanism contributing to what cognitive linguists sometimes call the ‘entrenchment’ of linguistic constructions (Schmid, 2007).

These mechanisms have been documented in simulation studies and support the idea that, given some elementary features of cortical neuroanatomy and function (sequence detectors, massive auto-associative connectivity) are included in the combinatorial network, associative correlation-based learning mechanism support discrete linguistic rule formation at the neuronal level (Knoblauch and Pulvermüller, 2005; Pulvermüller and Knoblauch, 2009). To achieve this, the learning rule applied is, once again, of critical importance. Using a Hebbian association rule alone would lead to undifferentiated clumping-together of large vocabularies; only by using a learning rule with both association and dissociation terms can the necessary specificity be achieved, so that, for example, different subtypes of noun-verb-interlinking CNAs emerge and separate from each other (simulation 2 in Pulvermüller and Knoblauch, 2009). Furthermore, the nature of the emergent ‘rules’ are somewhat different from those conceived by classic grammar theories, because they are specific to word categories defined not only by purely syntactic features but by semantic word properties as well. Combinatorial mechanisms for fine-grained semantic categories are consistent with usage-based approaches to language emphasising the intertwined nature of syntax and semantics (Goldberg, 1995; Tomasello, 2005; see also Pinker, 1989).

### 2.7.2. Anti-Hebbian dissociation learning and statistical preemption in language acquisition

Combinatorial learning in language links together signs but likewise connects string types with content and meaning. When learning sentences such as “Bob bounces the ball” or “Bob walks Ginger”, the learner would normally not be ignorant about the expressed meanings. What Bob is doing might be obvious from seeing him or from previous report, so that, in addition to a string of words, semantic information would be available. Combinatorial learning therefore includes relating a specific linguistic form to meaning (see, for example, Bowerman, 1988; Goldberg, 2006). In this form-meaning linkage, CNAs may play a key role. Given the construction containing an actor (subject) noun, an action verb and a patient (accusative) object is held together by two CNAs (for the subject-verb and verb-object links), these combinatorial units will activate each time the construction is used to express that an actor made a patient move in a certain way (or refrain from an otherwise likely movement, cf. “Bob stops the ball”). As a result, there will be synaptic strengthening between the CNAs and the neurons involved in meaning processing, more precisely in processing semantic features shared by the different strings. Thus, CNAs will load with semantic information. This semantic information will include semantic features shared by verbs typically used in the construction. Therefore, a possible brain basis for constructions are sets of CNAs, which interlink lexicosemantic classes of vocabulary items and relate them to construction-specific meaning (Fig. 4).

This neurobiological mechanism for usage-based acquisition of construction knowledge involves both associative and dissociative

<sup>9</sup> This ignores, for a moment, the semantic aspects characterizing constructions and construction learning. I will return to this issue in Section 2.7.2.

learning. The role of the latter can be illustrated using *statistical preemption*, leading to the exclusion of forms to which children might otherwise generalize established combinatorial patterns (Goldberg 1995; 2006; *in press*, see also Baker, 1979; Bowerman, 1988; Brooks et al., 1999; Clark, 1987). Assume a novel motion verb (V) is introduced in the context of someone bouncing a ball, by using intransitive sentences with only one verb complement, as in the simple intransitive “The ball Vs”, or in the so-called periphrastic causative construction “Bob made the ball V”. In principle, the simple transitive sentence (“Bob Vs the ball”) would be an alternative option to express the same semantic roles of “Bob” and “the ball” as actor and patient complements – instead of the periphrastic causative “Bob made the ball V”, where, again, V comes with only one complement. Given that sentence meaning and the intransitive sentence are processed at the same time, links will strengthen between the activated semantic neurons and the co-activated form representation (CNAs). This Hebbian process, which follows the ‘fire together, wire together’ term of the learning equation, may underlie the entrenchment of words and constructions. Now, crucially, at the same time, any links between the semantic neurons and the not-activated form representation (the CNAs of the transitive sentence) weaken because of the pre-nonpost rule of long-term synaptic depression (see Table 1; equation in Section 2.2). In Fig. 4g, the link between the semantic representation and the form-A circuit would thus strengthen, but those to the alternative circuit of form B weaken. If such weakening happens repeatedly, the form-B link will become ineffective (preempted).

That verbs presented in one construction context indeed ‘preempt’ (or ‘take out of the game’) their use in other possible constructions has been shown experimentally in children. Hearing sentences such as “He makes the ball tam” (where the pseudo-verb has only one complement) makes children less likely to use two-complement versions of the verb, e.g. “He tams the ball” in the same semantic contexts (Brooks and Tomasello, 1999 see also Goldberg, 2006, section 5.1). In this and similar cases, correlation learning, and in particular the dissociation (out-of-sync-delink) term in the equation in Section 2.2, offers a biological explanation why, “if a child has learned to use a particular linguistic form to express a particular meaning, this knowledge will tend to block their usage of another form to express the same meaning” (Brooks et al., 1999, p. 1326). Statistical preemption is of great relevance for language acquisition theories, because it allows learners to avoid over-regularizations in the absence of negative evidence. It applies to a variety of morphological and syntactic phenomena, including noun and verb inflection (\*mouses/mice, \*goed/went), causative constructions (A \*dies/kills B) and locatives (\*rob/steal money from the bank) (Bowerman, 1988; Goldberg, 1995, 2006; *in press*). The case of statistical preemption is one more example where a neurobiological approach to language goes further than so-called ‘associationist’ accounts and offers explanations of known linguistic facts at the neurobiological level.

In summary, the close integration of syntactic and semantic mechanisms implied by the neurobiological combinatorial mechanism sits well with current theories in the construction grammar framework, where syntax and semantics are thought to be tightly intertwined (Goldberg, 2006; *in press*; Hoffmann and Trousdale, 2013; Tomasello, 2005). At the next higher linguistic level, at that of speech act sequences in communicative interactions, the CNA mechanism provides a basis for dynamically linking individual speech acts into dialogic action sequences.

## 2.8. Communicative interaction

Brain research on language has so far mostly dealt with words and sentences, without considering language use and social

interaction. However, the main biological function of language is to make communication possible. The abilities to warn of a specific threat, to inform about an event, or to negotiate and coordinate plans between group members constitute a major selection advantage for the species. Rich vocabularies and/or complex sentences are helpful tools for performing these actions. Although the neuromechanistic basis of communicative action is still poorly understood, recent years have seen increased interest in its neurobiological investigation (e.g., Menenti et al., 2012). So, what are the brain circuits of the communicative actions or speech acts for which linguistic forms are used as tools? In the following paragraphs, an attempt to extend neurobiological models of language to communicative function will be offered. The main proposal is that sequences and complex branching structures made up of elementary communicative actions are stored and that each individual action – each utterance or movement – receives its ‘meaning’ from its place and context in larger action sequence structures (Alston, 1964; Fritz, 2013; Pickering and Garrod, 2004).

The role of action sequence structure is best illustrated by an example. A word or longer utterance can be used for different purposes. For example, the word “water” can be a tool for naming an object in the context of a language exercise, or for requesting and obtaining water from a partner. In these contexts, the same utterance serves different communicative purposes and is associated with context-specific sets of assumptions, intentions and predictions (see, for example, Alston, 1964; Stalnaker, 2002). In the naming context, the utterance comes with the expectation that the label is an established tool to refer to the object. In the request context, the same utterance allows for further social-interactive predictions, e.g. that the other party will hand over the object or that the request is rejected. Note that the single-word utterance can be replaced by more elaborate utterances, “this is water” or “we commonly call this a glass of water” in the naming context or “a glass of water please” or “could you please bring me a glass of mineral water?” in the request context. However, this utterance-centered description does not address the critical difference between contexts. Different speech acts are part of predictable action sequences or communication schemas, from which they receive their meaning: the request from predicting the handing-over of the requested object, the rejection of the request or denial based on an excuse, and the naming action from the prediction on silent approval of, or objection against, the verbal object label. Note also that the term ‘prediction’ may appear too narrow to describe the range of possibilities a dialogue schema opens up, because it may suggest a single action chain, whereas there are normally several alternative possibilities to act at each step of the sequence. Therefore, the ‘prediction space’ of a given communicative action is best described hierarchically, as an action tree characterizing the communication schema (Fritz, 2013).

The brain mechanisms for the postulated action sequence schemas can be conceptualized as a set of interlinked APCs that draw upon connection hubs and reach into sensorimotor systems. The linking may be indirect, via secondary circuits such as the combinatorial neuronal assemblies discussed in the previous section. Clear experimental predictions result from this proposal: If an utterance appears in different communicative contexts, the utterances’ APCs will co-activate and dynamically bind with different communication schema circuits and therefore activate these circuits. In the naming context, the APC of the utterance is activated, including its word form parts and its object-related referential semantic parts. The referential word-object link is known to draw upon inferior-temporal ventral-visual stream circuits. In contrast, when using the same word to request an object, the emphasis is on the link of the utterance with expected actions by the partner. The expectations are that the partner hands

over the requested object or that the partner rejects the request by uttering a sentence such as “Oh sorry, I can’t help you”. Therefore, motor and action sequence circuits in frontocentral cortex need to be primed in addition to utterance-related circuits. By assumption, this results in a relative enhancement of motor and language system activation in the request context. Section 3.3 will highlight experiments performed to directly address these specific predictions. This example may suffice for illustrating how the APC framework offers perspectives on the mechanisms of action sequence structures, which may be critical for processing the dialogic meaning of utterances in social communicative interaction.

### 2.9. Interim summary

Section 2 argued that information in sensory and motor areas drives the formation of distributed neuronal circuits in cortex, which include neurons in primary and secondary modality-preferential areas and neurons in multimodal connector hubs too. The cortical distributions of the circuits are determined by a) the types of information driving circuit formation and b) cortical connectivity structure. Neuroanatomically and neurophysiologically constrained simulation studies show that the undeniable presence of correlated sensorimotor information available in a cortex-like neuronal architecture leads to the formation of action perception circuits, APCs (Sections 2.1–2.4). In humans, APCs can develop for a rich lexicon of spoken language elements, because the newly evolved left-lateralized dorsal AF allows for efficient bridging between auditory and articulatory (and likewise visual and manual-gestural) information. These APCs form the neural basis of human language representation and processing. Whereas APC ignition with open motor gates gives rise to motor movements, its ignition following sensory stimulation with articulatory motor activity down-regulated serves as a mechanism for speech perception (see also Section 2.5). For carrying higher cognitive processes including memory, prediction and combination, the most strongly connected kernel parts of the circuits are most relevant. These lie primarily in the connector hub areas because of their high connectivity degree, where, therefore, a multiple demand network emerges as a by-product of the formation of circuit kernels (Sections 2.6). Combinatorial processing equally draws on these circuit parts and areas (Section 2.7). For circuits neurobiologically underpinning verbal symbols, the left perisylvian cortex is the prime site for circuit formation. Speech act function and the action embedding of symbols into interactive communication can be modelled using branching sequence schemas connecting action representations (Section 2.8).

The proposed framework offers perspectives for language understanding and use carried by the same, or strongly overlapping, neuronal circuits involved in both production and perception. A key statement is that elementary sensorimotor mechanisms are exploited in service of higher cognitive processes including language (see also Anderson, 2010, 2015; Pulvermüller, 1999). Verbal memory, intentions to communicate and linguistic predictions may also be carried by the same circuits. This contrasts with models in the modular spirit, where different processing components are postulated, for example a module for working memory and a different processing component for prediction, which both are seen as separate from perception and production machinery. In the APC framework, no separate mechanism is stipulated for working memory, but ‘active memory’ is a natural consequence of action perception circuit structure and, in particular, of the heavily interlinked kernel part of these cell assemblies. Two different and qualitatively distinct mechanisms are offered for prediction: first, the accumulation of activity and ‘spontaneous’ ignition consequent to strong within-circuit connections (Section 2.6) and, second, the priming between APCs

interlinked by overlap or combinatorial connections (Sections 2.7, 2.8), which allow, for example, for predictions on the next steps of a dialogue. The localization implications of the APC model converge with those of Guenther’s speech production model (Guenther and Vladusich, 2012), where a ‘speech sound map’ localized in prefrontal and premotor cortex interlinked with temporal areas (see their p. 412) provides predictive efference copies and thus the basis for computing articulatory corrections. However, there is the already mentioned important difference: no specific component for comparing predicted vs perceived (articulatory or communicative) actions is included in the APC model, whereas current speech production models emphasize the need for such separate modules (Guenther and Vladusich, 2012; Hickok, 2012).

The reason for omitting a prediction component is that no such module is necessary. Competition between activated circuits (see Section 2.5) provides a mechanism for modelling discrepancies between production and perception. For example, in the case of perturbed articulation, the input from auditory (and somatosensory) feedback may indicate a syllable different from the one being produced, so that the sensory-induced activations only partly support the activating APC in its ignition process, whereas additional sensory activity mismatching with the intended syllable adds to the ignition process. Thus, the competing double activation – due to the not-predicted sensory input and the intended APC’s ignition – provides excitatory enhancement serving as an ‘error signal’; and hence no separate module for error detection or input-output matching is necessary (see also Pickering and Clark, 2014). Note again that, within the action perception framework, predictive circuit dynamics can result from connections within and between cell assemblies and the functional dominance of cell assembly kernels explains the ‘shift away’ from primary sensory and motor cortices. Therefore, adding to the neurobiological and neurocomputational motivation of the APC perspective, this framework is more parsimonious than a multi-module proposal.

### 3. Action-perception circuits (APCs) for language: experimental tests

If distributed neuronal circuits driven by correlated action- and perception-related information provide the brain mechanisms for language and communication, a range of predictions can be made for experimental neuroscience research. These address activity spreading across the cortex when human subjects perceive and produce phonemes, words, sentences and communicative actions. Key to the temporal dynamics of brain language mechanisms is the concept of circuit ‘ignition’ (Section 2.2, 2.4), the instantaneous activation of a circuit including its kernel and halo parts, and the prolonged state of reverberation where activity is most prominent in the kernel (Section 2.5). The postulated cortical distribution of the circuits implies specific spatiotemporal activation patterns, which, in the language domain, include primary, secondary and higher association areas. Crucially, circuit distributions and therefore activation patterns may vary considerably across types of linguistic entities defined at different levels. Different circuit topographies have been postulated for different phonemes/syllables reflecting their phonetic and phonological features, for word types reflecting their semantic features and for types of communicative actions reflecting their function in social interactions. Brain theory and converging simulation studies provide a foundation of the relationship between established neuroscience principles, the type of information stored and the postulated circuit topographies. A co-determinant of activation topographies is the level of activity regulation or inhibition in different areas of the cortex, which depends on context, task and attention (Section 2.6). A range of previous reviews highlighted these issues in the context of experimental and computational data (Bak, 2013;



Binder and Desai, 2011; Fadiga et al., 2007; Glenberg and Gallese, 2012; Kemmerer, 2015a; Kemmerer and Gonzalez-Castillo, 2010; Kiefer and Pulvermüller, 2012; Pulvermüller, 2005, 2013a; Pulvermüller and Fadiga, 2010; Wennekers et al., 2006). The data overview below will focus on the spatial-topographical predictions the action perception model makes for different speech sounds (Section 3.1), semantic word types (3.2) and speech acts (3.3), and on recent evidence gained to evaluate these predictions. This focus is chosen, because the APC model makes specific predictions about the cortical mechanisms effective at these levels and a high degree of linguistic specificity may be seen as an advantage of an explicit neurobiological model of language and communication.

According to the model presented in Section 2, APCs distributed over frontal and temporal perisylvian areas take a role in the processing of words and constructions. This claim receives support from studies investigating the neurophysiological differences between words and meaningless but phonotactically regular pseudowords. Physiological differences are manifest in standard event-related potentials and fields (Hauk et al., 2006a; Kutas and Federmeier, 2011; Pulvermüller et al., 2001b) as well as in spectral responses in the delta, theta, alpha, beta and gamma ranges (Canolty et al., 2007; Garagnani et al., 2017; Lutzenberger et al., 1994; Strauss et al., 2014). Activation foci in both temporal and frontal areas, and even the precise timing between these activations, index the word/pseudoword contrast (MacGregor et al., 2012; Pulvermüller and Shtyrov, 2009). Crucially, the neurophysiological differences between meaningful symbols and their matched meaningless counterparts emerge extremely early, already within 100–200 ms after they can be recognized from the acoustic or visual input. This observation is consistent with the idea of an instantaneous ignition elicited by words (but not pseudowords).<sup>10</sup> In addition, upon sensory activation of temporal cortex, recruitment of frontal areas takes place with a further minimal activation delay of 15–25 ms (Pulvermüller and Shtyrov, 2009; Pulvermüller et al., 2003), which is consistent with the signal propagation times between these cortical regions (Matsumoto et al., 2007, 2004). This near-simultaneity therefore provides further support for an instantaneous ignition process involving symbol-related circuits distributed across left frontotemporal perisylvian cortex. These spatio-temporal results would remain unexplained if local circuits in temporal cortex (for example in anterior inferior or posterior middle temporal areas) were the only brain mechanism of symbol comprehension. The possibility that fast frontal activation in language understanding is epiphenomenal to the comprehension process will be discussed in light of neuropsychological studies below. The focus will now be on fine-grained spatial distinctions between the cortical correlates of phonemes, meaningful symbols and linguistic actions.

### 3.1. Topographically specific APCs for speech sounds in frontotemporal perisylvian cortex

Neuroanatomical data show human-specific strong bidirectional functional links within the perisylvian core language network, especially between superior temporal and inferior frontal cortex, by way of both ventral extreme capsule and dorsal arcuate

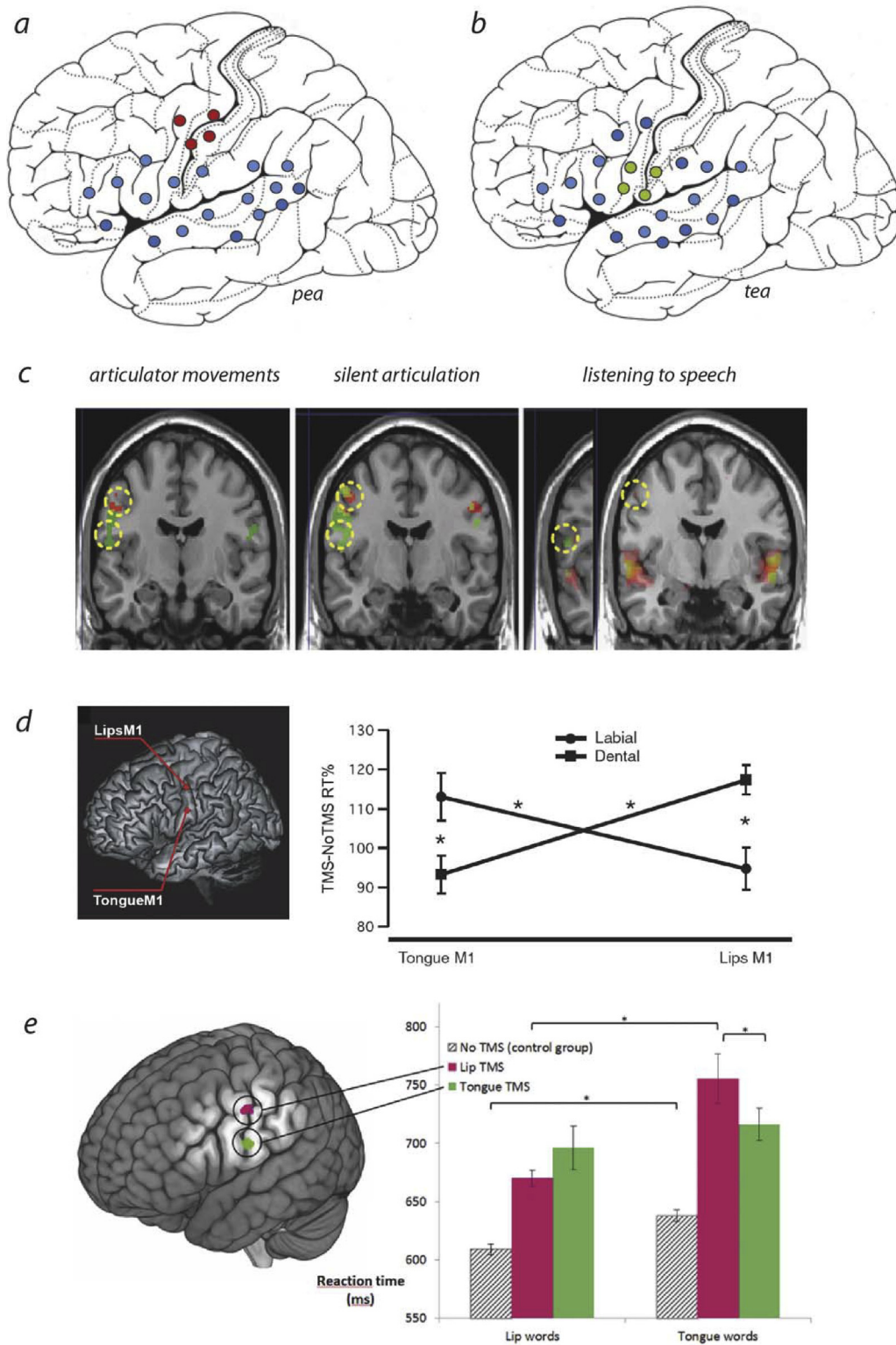
fascicle (Section 2.1, Fig. 1). These bidirectional links imply sensorimotor information flow and functional interaction within perisylvian circuits when language is processed for production, understanding and conceptual thought. In contrast, classic brain language models proposed a more modular organization, according to which separate ‘centers’ or ‘modules’ are respectively dedicated to speech production and perception (see, for example, Geschwind, 1970; Lichtheim, 1885). Although there is now agreement on the general role of temporal cortex in language processing, some historic and current models still consider the inferior frontal and articulatory motor cortex as speech production hubs with little or no role in speech perception and comprehension (Hickok, 2014; Marie, 1906). Similar to the modular models, this suggestion implies local and functional separation of language production from perception/understanding mechanisms. As discussed above (Section 2.4.1), both functional separation and interaction accounts can accommodate the well-known observation from aphasia research that a relatively more prominent speech production deficit arises from inferior frontal lesion, whereas superior temporal lesions frequently lead to a language deficit most prominent in comprehension (for review, see Pulvermüller and Fadiga, 2010). Whereas double dissociations are consistent with both theory types, the association of action and perception abilities present in most cases of aphasia is left unexplained by functionally separate systems.

A current debate surrounds the question of the inferior frontal cortex role in understanding and perception of speech. The alternatives currently under discussion are that the temporal cortex houses the real gateway to understanding”, whereas frontal and sensorimotor cortex under standard conditions play no relevant role, or, alternatively, that distributed circuits reaching into modality-preferential areas, including the sensorimotor system, carry both language perception/comprehension and production as well.

It was argued in Section 2.4 that, when specific phonemes and word forms are being used and learned – thus entailing correlated motor output and concordant acoustic input – neuroplasticity and frontotemporal connectivity imply that reciprocal frontotemporal connections strengthen and APCs specific to phonemes and spoken words develop. According to the interaction account, these APCs provide the substrate for both speech production and perception.

Although all APCs for phonemes and spoken words may be distributed over perisylvian cortex, *distributional differences* between speech sounds and syllables can be predicted. The articulators contribute to different degrees to the production of particular speech sounds and words. Specific motor activity controlling the tongue is necessary to pronounce a tongue-related ‘coronal’ phoneme such as /t/ or a syllable or word such as “tea”, whereas motor programs specifying lip movement trajectories are required for ‘labial’ sounds like /p/ or words such as “pea”. The different articulator-specific motor commands involve different parts of the motor and premotor cortex, because muscles and muscle movements are mapped somatotopically on the motor cortex (Dum and Strick, 2002; Kakei et al., 1999; Penfield and Rasmussen, 1950), where the tongue, lips, hands and the lower body are represented from bottom to top. Apart from neighborhood in body space, an elementary movement’s direction, function and participation in complex sequences (Griffin et al., 2015; Matsuzaka et al., 2007) is represented in the motor system, with similar complex movements being mapped to adjacent loci (Graziano and Afalo, 2007). The somatotopic mapping of the articulators allows for predictions on differences in circuit topographies related to the place of articulation of speech (Pulvermüller et al., 2006). This differential mapping applies not only for motor and premotor cortex but for somatosensory areas too, where a parallel somatotopy applies and tactile

<sup>10</sup> These ‘early’ responses precede the later event-related potentials to language of the N400 brain response (300–500 ms). However, as the first evoked activity in auditory cortex follows stimulus onset by only 20–30 ms (Eldredge and Miller, 1971; Rupp et al. 2002), the question arises why the ‘early’ linguistic brain responses do not even emerge before 100–200 ms. To date, only a small number of recent studies reported such ‘ultra-early’ language-related responses (see, for example, MacGregor et al., 2012; Shtyrov et al., 2014; Shtyrov and Lenzen, 2017). The mechanisms underlying the latencies of the evoked and induced neurophysiological indexes of linguistic processes reported in most studies call for explanations (Pulvermüller and Shtyrov, 2006).



**Fig. 5.** Action perception circuits, APCs, for phonemes and word forms. *a, b.* APCs postulated for spoken words distinguished only by one phoneme carrying either the phonological feature/value [+labial] or [+coronal], as, for example, the minimal pair “tea” vs. “pea”. In the model, the density of local neuron clusters (small circles) in perisylvian areas reflects the differential involvement of tongue (ventral clusters in green) or lip (lateral clusters in red) in processing the word form. *c.* Brain activation during tongue and lip movements, silent articulation of syllables starting with /t/ or /p/, and listening to /t/ and /p/ sounds (green/red for tongue/lip related items, Pulvermüller et al., 2006). *d.* Results obtained with transcranial magnetic stimulation to tongue and lip motor cortex on the perceptual classification of coronal and bilabial stop consonants (D’Ausilio et al., 2009), and *e.* on the comprehension of words starting with labial or coronal stop consonants (Schomers et al., 2015). These results show differential contributions along the sensorimotor cortex to speech sound and spoken word perception and understanding depending on place of articulation.

self-stimulation is present during articulation. Because articulations of 'tongue-related' labial speech sounds and spoken word forms including them draw more heavily on neurons in ventral sensorimotor cortex, whereas for lip-related coronal ones more dorsal sensorimotor areas are relatively more important (Bouchard et al., 2013; Breshears et al., 2015; Pulvermüller et al., 2006), the APCs for tongue- and lip-related speech sounds and word forms are predicted to have different topographies. Gradually different activation topographies and circuit distributions can model this. Accordingly, labial and coronal phonemes would have neuronal ensembles with different local neuron density distributions (Fig. 5a). Similarly, topographical distinctions may also apply to other phonological features and to phonological and word form representations in temporal cortex. However, in this case, the mapping of acoustic-phonological features to specific areas of cortex, and therefore any a priori predictions on their local mapping, is less straightforward than for the place-of-articulation features in the sensorimotor system. In sum, APCs for phonemes and spoken word forms may exhibit local distributional differences related to both articulatory- and acoustic-phonological differences between speech units. Out of these, differences in the place of articulation lead to most straightforward postulates about differential circuit topographies in articulatory sensorimotor cortex and to corresponding experimental predictions.

In agreement with both functional separation and interaction accounts, brain activation related to phoneme perception has been demonstrated in superior temporal regions, especially in and around the superior temporal sulcus and in anterior superior temporal gyrus (Obleser and Eisner, 2009; Uppenkamp et al., 2006). The question whether any of these temporal areas is specific to speech is open; a range of physical and conceptual features of speech and non-speech sounds may be reflected (Griffiths and Warren, 2002; Price et al., 2005). What is crucial from a linguistic perspective is that different phonemes (e.g., /p/ vs. /t/) elicit locally specific activation patterns in superior temporal cortex (Mesgarani et al., 2014; Obleser et al., 2006). A recent study found the phonological features of voicing, manner of articulation, and place of articulation to be reflected in activation patterns distributed throughout bilateral primary, secondary, and association areas of the superior temporal cortex (Arsenault and Buchsbaum, 2015a). One may still argue that some of these works still leave it unclear whether acoustic, physical factors of speech sounds, or rather phonological ones, account for the differences in brain activation. In this context it is crucial to point to classic work reporting that the magnitude of brain responses generated in superior temporal cortex can index that speech sounds are within the phonological repertoire of a speaker's mother tongue (Dehaene-Lambertz, 1997; Näätänen et al., 1997) or that they became familiar in the context of language learning (Winkler et al., 1999). These results underscore the superior temporal cortex' (anterior and posterior superior gyrus and sulcus) role as a phonological processor.

Although most neuroscience work on language focuses on language perception and comprehension, work on speech production is equally important for evaluating neurobiological language models. In this domain, however, most of the available evidence seems to weigh against a concept of simultaneous ignition of frontal and temporal circuit parts in producing words. Instead, the majority of results seem to indicate that, when naming a depicted object, activation proceeds from occipital cortex to middle temporal gyrus, where lexical and phonological information is retrieved, to prefrontal cortex, where syllables are generated, and finally to motor systems, where articulations are controlled (Indefrey and Levelt, 2004). Most reported delays are quite substantial, with frontal and motor areas activating with a lag of 200–600 ms upon presentation of a to-be-named picture. Only

some recent results seem to contradict such slow serial processing, showing, for example, neurophysiological signs of lexical access and even phonological information processing in frontocentral articulatory motor systems, within 150–200 ms upon object picture onset (Miozzo et al., 2015; Strijkers et al., 2010 in press). These latter results argue against serial processing of semantic, lexical and phonological information processing and suggest instead that, similar to the situation in speech perception, ignitions may account for near-simultaneous processing of semantic, lexical and phonological information and activation of perisylvian areas in production too. For an in-depth discussion of predictions of the cell assembly framework in light of recent results on speech production, the reader is referred to a recent review (Strijkers and Costa, 2016). Investigating the different predictions of functional separation and interaction models in speech production is an important topic of future research.

A critical prediction made by the interactive APC perspective, but not by the separation approach, addresses inferior frontal activation in speech perception and understanding. Some studies showed inferior frontal and even motor system activation to perceived speech (Fadiga et al., 2002; Watkins et al., 2003), a property speech sounds seem to share with non-linguistic sounds of self-produced actions (Hauk et al., 2006b; Lahav et al., 2007). However, some of this activity appeared to be specific to phonological type (Fadiga et al., 2002). Fine-grained spatial aspects of motor activity in speech perception were examined further in a task where subjects had to attentively listen to meaningless consonant-vowel (CV) syllables. In this experiment, acoustic stimuli were presented in silence and subjects did not engage in any motor movements – especially no button press response was required in order to avoid general motor system activation, because it could counteract any fine-grained language-related activations (see also Section 3.2.2). In this experiment, inferior frontoparietal activation in and close to the motor system was found to be phoneme specific; it reflected the sensorimotor representation of the articulator most strongly involved in speech sound production (tongue for coronal /t/ and lips for bilabial /p/, Pulvermüller et al., 2006). Crucially, the differential activations found during speech perception were consistent with the motor loci where, in the same individuals, strongest articulator-specific activation was found during production tasks. Concordant results were reported using neurophysiological measures, in tasks distracting subjects from the incoming sounds (see also Section 3.1.1, Möttönen et al., 2013; Rogers et al., 2014). Recent results from multi-voxel pattern classification also indicate that, apart from superior temporal areas, inferior frontal and motor cortices provide phonological information about heard speech sounds (Evans and Davis, 2015). Although a recent publication called into question the existence of such 'phonologically somatotopic' activation (Arsenault and Buchsbaum, 2015b), broader review of the literature indicates that variables specific to the experimental setup – in particular the absence of an active motor (button press) task and the use of 'sparse imaging' techniques allowing to present speech unmasked by fMRI scanner noise – are critical for obtaining motor cortex activation during speech perception in general and these specific findings specifically (Schomers and Pulvermüller, 2016; Skipper et al., 2017).

Motor activity indexing the perception of speech sounds – and even differences between perceived speech sounds – is difficult to account for from a functional separation position, which assumes that inferior frontal and articulatory motor systems are irrelevant for language perception and understanding. Therefore, these results provide direct support for the postulate that different speech sounds activate APCs with distinct frontotemporal topographies that reach into the sensorimotor system (Fig. 5) and



bind phonological-acoustic and phonological-articulatory representations.

### 3.1.1. Is there a causal influence of sensorimotor cortex on speech perception and comprehension?

In neuropsychological research, an area A is considered to be causal for a function F if a lesion or activity change in A changes F. Accordingly, a strict proof of a causal contribution of a given area to language comprehension is possible if activity changes in, or lesions to, the area have measurable effects on comprehension. In the simplest case, functional changes may impair or delay the comprehension process. However, it is well-known that the measurable change does not necessarily imply a deficit or change in processing times. In case of redundancy, with several areas supporting the same function, a partial lesion of one site of the distributed network may be compensated for, so that no overt behavioral changes are present. Instead, a redistribution of relevant neuronal processes may occur within the network (Luria, 1980).

To examine the possibility of a functional change, it is useful to look at patients with language deficits caused by brain lesions. Many studies of aphasia indicate that some difficulties with speech perception and comprehension arise from lesions of the frontal cortex (Bates et al., 2003; Caplan et al., 1995; Moineau et al., 2005), although not all studies agree on this conclusion (e.g., Rogalsky et al., 2011). Note that post-lesion reorganization processes along with the large size of most cortical lesions make it difficult to draw firm conclusions from lesion data on the specific contribution of an area as small as, say, the lip-related articulatory motor cortex. Therefore, TMS studies were performed to address the question about causal influences of focal sensorimotor areas on language perception and comprehension.

When using noise-overlaid meaningless syllables as stimuli, TMS to tongue and lip motor cortex specifically affected phoneme classifications for tongue- and lip-related speech sounds (D'Ausilio et al., 2009; Möttönen and Watkins, 2009). This result was interpreted as evidence for a causal role of articulatory sensorimotor cortex for speech perception. However, noise overlay was criticized as inconsistent with normal language understanding, because it may enforce attention to the stimuli so that any effects might be, in part, related to unnaturally amplified focusing on the critical stimuli (Hickok, 2014).

Although noisy environments are arguably typical for normal language use, an additional experiment was performed to accommodate this concern. A similar experiment with undistorted single word stimuli clearly pronounced by different speakers was performed, thus approximating the variability of natural language within a noise-free environment. The differential influence of tongue and lip stimulation on the perceptual classification of speech sounds was replicated, thus ruling out the attention argument and reconfirming that a phoneme perception process was modulated by focal activity in motor areas (Bartoli et al., 2015). In addition, this study showed that the effect of motor cortex stimulation reflected the perceived perceptual distance between the phonological perception and the speaker's own phonological representations.

A further criticism addressed the task used in D'Ausilio's experiment from 2009, where an explicit classification of speech sounds into phonemic categories was required (Hickok, 2014). Although similar phoneme classification tasks have a long tradition in neuropsychological and psycholinguistic research on speech perception (Basso et al., 1977; Blumstein et al., 1977), their use in the documentation of language action links has been criticized, because such explicit classification is not required in natural language processing. This is indeed correct, although speech sound classification is certainly not the only

psycholinguistic task that differs from language use in everyday communication, and such divergence does not necessarily hinder researchers to obtain valuable information about relevant cognitive and neuronal mechanisms. Still, in the context of speech perception, it would be ideal to reconfirm any results with tasks where subjects passively hear speech and do not attend to it. As such tasks would not offer any behavioral responses indexing speech perception or comprehension, it might be useful, in this case, to take advantage of a more direct measure, ideally a brain response with an established record of mapping attention-independent automatic stimulus-elicited recognition processes.

Such a brain response is the Mismatch Negativity, a brain response considered to index automatic acoustic change detection and speech perception (Näätänen, 1992, 2001; Näätänen et al., 1997). Rapid temporal to frontal activation spreading was indeed revealed by the Mismatch Negativity brain response to speech sounds and words, even when subjects were instructed to ignore the speech stimuli and were distracted from their processing (see beginning of Section 3, Pulvermüller and Shtyrov, 2009; Pulvermüller et al., 2003). This result is consistent with the claim that the activation of speech related APCs does not require focused attention and is, in this very sense, *automatic* (Pulvermüller, 2005; Pulvermüller and Shtyrov, 2006). Further experiments used passive listening conditions to address the influence of articulatory motor cortex TMS (tongue locus) on the brain response to undistorted speech sounds, and found a reduction of the Mismatch Negativity brain responses to tongue-related speech sounds (Möttönen et al. 2013). If the Mismatch Negativity is interpreted as an index of phoneme perception and recognition, this study documents a causal role of lip articulatory cortex on bilabial phoneme perception. In this perspective, the result showed that neither noisy stimuli nor a phonetic task are necessary to document the articulatory motor system's causal contribution to speech perception (Möttönen et al., 2013; Smalle et al., 2015). Furthermore, it was demonstrated that response bias, a feature claimed to underlie some motor system activations (Venezia et al., 2012), is not explained by magnetic stimulation of the motor representations of the articulators (Smalle et al., 2015).

Still, most of these previous results focused on the perception of meaningless phonemes, but not on their main linguistic function: to distinguish between meaningful units of language. The important issue of whether motor cortex activity takes a causal role in the comprehension of meaningful speech *per se* has recently been investigated using a standard psycholinguistic comprehension task, in which subjects heard single spoken words and had to select semantically matching pictures. Among the stimulus words were items that differed only in one speech sound, which was either tongue or lip related, for example the 'minimal pair' "beer" vs. "deer". When an item from a minimal pair was presented (for example "deer"), the pictures to choose from included depictions of the meaning of both minimal-pair alternatives (i.e., a deer and a beer). During stimulus word delivery, magnetic stimulation was applied either to the lip or to the tongue motor cortex. TMS to concordant sensorimotor loci (e.g., 'tongue-word' responses to tongue locus TMS) led to faster comprehension responses than TMS to discordant loci ('tongue-word' responses to lip locus TMS, Schomers et al., 2015). This result – manifest in a significant interaction of articulatory-motor stimulation site and phonological word type – shows that articulatory motor cortex has a causal effect on language understanding even at the basic level of single word processing. This influence depends on stimulation site along the somatotopic sensorimotor map of the articulators and on the place of articulation of critical phonemes distinguishing word stimuli. Concordantly, a study using spoken words with variable degree of noise overlay documented that TMS double stimulation of inferior frontal and premotor cortex influences speech

perception and understanding independent of the level of noise overlay (Murakami et al., 2015).

A range of behavioral studies addressed the influence of sensorimotor changes on the perception of speech, for example by stimulating the face by patterns normally present in speech production (Ito et al., 2009; Nasir and Ostry, 2009; Ostry and Gribble, 2016). The somatosensory change linked to articulations altered speech perception. Interestingly, a causal role of articulatory sensorimotor processing on speech perception has even been revealed by behavioral studies in infants, whose speech sound perception were found to be affected by changes in articulator position (Yeung and Werker, 2013; Bruderer et al., 2015).

These and related observations (for more extensive review, see Schomers and Pulvermüller, 2016) directly support a functional role of posterior inferior frontal cortex and sensorimotor areas in speech perception and understanding. They are consistent with the proposal that phonological APCs, including their member neurons in specific parts of articulatory sensorimotor cortex, play a causal role in language comprehension. The results provide evidence against functional segregation and, in particular, against the claim that motor regions are of no relevance for perception and understanding. Frontotemporal APCs held together by the arcuate fascicle carrying phonemes, syllables and word forms (see Section 2.4) provide a straightforward explanation for this and a range of additional recent findings, including the observations that strength of functional connectivity of the arcuate predicts word learning performance of individual subjects (Lopez-Barroso et al., 2013) and that precentral activation during speech perception reflects verbal short term memory capacity of experimental subjects (Szenkovits et al., 2012). Still, the evidence for a role of frontocentral - including motor - cortex in speech perception and language understanding does not argue against the relevance of temporal areas for the same processes. Within the APC framework, the comprehension process is seen as the result of joint action of distributed neuronal assemblies that include temporal, parietal and frontal neuron members.

### 3.1.2. Overcritical approaches to action perception integration

It may be justified to add some additional specific comments on the current state of the action perception integration debate. These relate to both the language domain and the role of multimodal and mirror neurons in comprehension (see also Section 2.4.1). Many experiments showing motor system involvement in speech perception and understanding have been confronted with unusually persistent criticism. This is understandable, because the action perception integration framework questions many beliefs immanent to classic approaches to brain function (for discussion, see Fuster, 2003), modular cognitive theory included. However, in the context of fundamentally diverging viewpoints, it is particularly important that discussions do not go astray but adhere to standards of scientific dispute. Unfortunately, the discussion of action perception integration is sometimes perturbed by recurring purported 'counterarguments' that cannot appear as particularly convincing. I would like to highlight some of these briefly below.

First, there is overemphasis of spatial precision, especially in the TMS context. For example, possible co-stimulation of postcentral somatosensory with precentral motor cortex has been criticized (Hickok, 2015) in D'Ausilio's and Schomers' TMS studies (D'Ausilio et al., 2009; Schomers et al., 2015). As described above, these researchers targeted loci in articulatory motor cortex in the central sulcus, and these were selected because fMRI experiments showed that they were most active during articulator movements and phoneme production. The criticism seems to address a natural limit of TMS research, as motor and somatosensory cortex are close to each other and functionally

connected. TMS makes it possible to target small cortical loci, and in this sense may be considered a gold standard for noninvasive focal stimulation of the healthy human brain. But there is still a degree of spatial imprecision immanent to this method, especially as the biophysics of magnetic nerve cell stimulation is not fully understood (Bailey et al., 2001; Wassermann et al., 2008). Thus, even if stimulation had been in part to somatosensory cortex, co-stimulation of motor areas would likely be due to both spatial imprecision and to sensorimotor functional interaction. Likewise, a change in somatosensory input, which is also known to change phoneme perception (Ito et al., 2009), may in part be due to motor processes, because even passive somatosensory stimuli may elicit motor responses. In addition, and crucially, the action perception model predicts that both articulator-related motor neurons and somatosensory neurons with correlated activity during phoneme production are embedded in APCs (Pulvermüller and Fadiga, 2010); therefore co-stimulation of motor and somatosensory neurons is appropriate when testing this prediction. When dissociating the functional contributions of inferior frontal and parietal cortices, both were found to influence phonological and semantic comprehension (Hartwigsen et al., 2015).

A further hypercritical strategy is to criticize statistically significant results as insufficient. To demonstrate that two factors make a nonlinear joint contribution to a dependent variable, it is usually sufficient to show that their interaction effect is significant. However, for several studies from the action perception literature, the claim has been made that such interaction results are inconclusive unless additional post-hoc tests show significance (Hickok, 2015; Papeo et al., 2013). For example, in order to show that activity in the sensorimotor cortex influences the processing of different linguistic types (e.g., words starting with a tongue- or lip-engaging sound) in different ways, the *significant interaction* of motor stimulation locus (e.g., tongue/lip sensorimotor cortex) and linguistic type brought about by local TMS is sufficient as statistical support. This interaction shows that 'congruent' and 'incongruent' stimulation (to different motor areas) has distinct effects, which is what had been predicted by theory. Additional significance of post-hoc tests may further strengthen this result and allow for additional, more specific, conclusions, but absence thereof cannot invalidate this result. Let me mention one more example of such statistical greediness: The fact that an effect (of language action interactions) was manifest in reaction times but failed to reach significance in accuracy measures led the author of a recent critical note to conclude that these effects were 'inconsistent' and therefore not conclusive (Hickok, 2015). This once again is not appropriate. Psycholinguists and cognitive scientists are normally happy if they find significant effects on one dependent variable, for example either response times or accuracies. If the other measure(s) do(es) not obviously contradict the significant result, an interpretation is normally unproblematic. Established inference schemes, including the interpretation of either response or accuracy results, should be permitted in the debate about motor cognition as they are elsewhere in cognitive and neural science.

A further problematic strategy applied in the discussion about action perception mechanisms is to play down relevant effects by using unclear verbal labels. Given that it became undeniable that the motor system plays a role in speech perception and understanding (Schomers and Pulvermüller, 2016; Skipper et al., 2017), Hickok claims that the influence of sensorimotor cortex activation on understanding is only 'modulatory' but not 'essential' or 'primary' (Hickok, 2015). As different cortical loci influence language comprehension, there seem to be no scientific criteria for distinguishing 'modulatory' from 'primary' effects. The background idea may be that there is only one 'primary' brain locus for each cognitive task, although 1:1 relationships between

cognitive functions and brain loci appear as unrealistic to many neuroscientists today. More importantly, statistical tests can show whether there is an effect or not, and whether one of two effects is relatively stronger. Clear scientific criteria are necessary for meaningful claims about 'primary' and 'modulatory' roles of brain regions. Relative effect sizes obtained in experiments can be related to quantitative results derived from neuron density functions of circuits distributed across areas as they are provided by simulation studies (Garagnani and Pulvermüller, 2016; Garagnani et al., 2008). And simulation studies of APC activation can also be used to predict and explain the differential effects of focal (frontal or temporal) lesions on speech production and comprehension (see Pulvermüller and Preissl, 1991). However, irrespective of whether the documented causal influence of sensorimotor cortex activation on language perception and comprehension is small or large (and irrespective of whether one prefers to dub them 'primary' or 'modulatory'), it is a causal influence and this is what had been under dispute.<sup>11</sup>

In summary, some critical points raised against the action perception perspective were evaluated, but appeared as not convincing. Whereas critical attitudes toward theories are required for scientific progress, overcritical approaches may well prevent such progress by cutting off fruitful avenues. Therefore, a reasonable and moderately critical strategy is recommended.

### 3.1.3. Action perception circuits for speech and the motor theory of speech perception

A prominent early action-centered approach to language processing is the motor theory of speech perception, which postulates that a module for phoneme production is also most critical for phoneme perception (Lieberman et al., 1967). There is little evidence for such a motor-centered, phoneme-specific system (Galantucci et al., 2006), especially as brain activation patterns for speech sound production and nonlinguistic movements of the articulators seem to activate similar parts of cortex (Pulvermüller et al., 2006). In the tradition of motor theories, the direct realist account (Fowler, 1986; Galantucci et al., 2006) postulates that a mapping between the surface feature of a sound and its inferred antecedents (e.g., in the motor systems) takes place during sound perception and may be crucial for it. This view converges with models postulating action perception integration

<sup>11</sup> Some recent results seem to strongly and generally argue against any relationship between perception and action mechanisms. In particular, Vannuscors and Caramazza (2016) reported that "individuals born with absent or severely shortened upper limbs (upper limb dysplasia), (...) could perceive, anticipate, predict, comprehend, and memorize upper limb actions" just like normally developed individuals. The authors argue that their results "undermine (...) theories that place the locus of action perception and comprehension in the motor system". However, as severely deprived individuals were studied, it is naturally difficult to exclude the possibility of compensatory mechanisms: these subjects' largely unimpaired perception and recognition of hand/arm actions could potentially be related to special perceptual skills compensating for possible contributions from their motor system. Strong conclusions on specific contributions of upper limb motor systems require comparison of the putatively affected processing of upper limb actions with a within-subject control condition – for example the perception of actions involving body parts unaffected by the disease. When Vannuscors and Caramazza compared the recognition and naming of point light animations of upper limb actions to that of non-upper limb actions, 3 out of their 5 arm-dysplastic patients performed substantially worse on arm than on non-arm actions (percent correct for arm vs non-arm actions: ca 20 vs 85, 50 vs 90, and 50 vs 95; numbers read from their Fig. 1b; a comparably large difference was absent in control subjects). Out of the two patients who did not show body part differences, one (ID4) suffered not only from arm-dysplasia, but, in addition, malformation of the leg, so that similar performance on arm and non-arm actions was predicted by action perception accounts. Therefore, in the crucial within-subject comparisons, 4 out of the 5 patients showed patterns of (im)balance between arm and non-arm action recognition/naming consistent with the action perception integration model; these results do not argue against a role of motor mechanisms in perception.

(Hommel et al., 2001; Prinz, 1990), including the action perception perspective on language outlined here, as far as phonemes and action sounds are concerned. A problematic assumption of the motor theory of speech perception had been that speech sound perception generally, or at least the main burden of its processing, is off-loaded into a motor 'module'. The obvious discrepancy between a perceptual process and its modular and (in one reading) exclusive processing outside modality-preferential sensory systems was rejected as implausible (for discussion, see Galantucci et al., 2006; Hickok, 2014; Schomers and Pulvermüller, 2016).

The classic motor theory as a cognitive theory is not a feasible basis for detailed discussion of the neurobiology of language, although it has been used as a straw man in some recent discussions. A realistic alternative to the other extreme – the exclusive or near-exclusive processing of speech information in superior temporal cortex – needs to be considered in order to accommodate the available data. Put in terms of an integrative action perception model, the alternative to modular ideas (regardless of whether they are motor- or sensory-centered) is that motor, sensory and connector hub systems each have a say in perceptual processing, and this claim has gained ample empirical support. The underlying mechanisms are not well described in terms of modules but rather require the concept of distributed circuits that interlink specific action and perception mechanisms. These circuits are causally influenced both by functional changes in the superior temporal gyrus and sulcus and by activity in frontocentral motor areas and adjacent inferior prefrontal cortex. The APC model views interactive frontotemporal circuits as the 'seat' of both performance and perception of linguistic as well as non-linguistic actions and models these action and perception processes as different trajectories of activity in the same circuits (Pulvermüller, 2002b).

### 3.2. Topographically specific semantic circuits

The question of meaning and the brain is one of the most controversially discussed issues in cognitive neuroscience. Proposals about the most important brain loci for semantic processing, sometimes called semantic 'hubs' or 'interfaces', name areas as different as the inferior frontal (Bookheimer, 2002; Devlin et al., 2003), anterior-inferior temporal (Patterson et al., 2007; Ralph et al., 2017), anterior-inferior parietal (Binder and Desai, 2011) and posterior-middle/superior temporal lobe (Hickok, 2014; Price, 2000) (see Fig. 6a,b). Broad reviews of the available data make it likely that all of these areas take their shares in semantic processing and make differential contributions to it (Binder et al., 2009; Price, 2012; Pulvermüller, 2013a; Vigneau et al., 2006). However, the explanation of these findings, and, in particular, of why so many areas might be involved and why they might contribute differentially, requires a neurobiological theory.

When subjects perceive and produce spoken words, APCs in the central frontotemporal perisylvian cortex become active. If frontotemporal circuits carry the form of words and other symbols, it still remains open by which mechanisms these items become carriers of meaning. Distributional or combinatorial theories of meaning postulate that meaning is best defined in terms of the symbol contexts in which individual symbols frequently occur. Therefore, distributional semantic theories use the co-occurrence statistics of words in texts to define their semantic similarity (Landauer and Dumais, 1997), so that words that frequently co-occur with similar sets of context words within standard texts are considered as closely related semantically. Correlation mapping according to Hebbian learning principles provides an obvious pathway to 'brain-implementing' word-word correlations as they are relevant for defining semantic similarity in a distributional



sense (Pulvermüller, 2002a, 2012). In this perspective, the associative connections between word-related APCs provide the mechanism for storing aspects of symbolic meaning. As discussed in Section 2.7, the memory-active parts of APCs in prefrontal and temporal parabelt connector-hub areas are relevant for combinatorial learning. The relevance of these areas for semantics indeed received strong experimental support (Fig. 6a, Bookheimer, 2002; Carota et al., 2017; Devlin et al., 2003; Hillis et al., 2001).

Although the combinatorial definition of meaning allows for fine-grained descriptions of semantic similarity, it runs into difficulty when it comes to explaining the mechanisms that provide the link between word forms and the objects and actions these symbols are used to speak about (Harnad, 1990; Searle, 1980). A mechanism is needed for explaining *semantic grounding*, that is, the semantic links between words and their related actions, referent objects and concepts. Such a mechanism is offered by correlation learning (see also Section 2.7). Word form circuits in perisylvian language cortex co-activate, and thus link up with, conceptual circuits, which include neurons in extrasylvian cortex. In this way, higher-order *semantic circuits* are being built by joining together word form (symbolic) and concept-related circuits. For semantically ambiguous words, the same word form connects with two (or more) perisylvian form circuits, and for synonyms, the same word form representation interlinks with two extrasylvian conceptual circuits. Note that, for this model to operate properly, activity regulation and inhibition between overlapping circuits is absolutely essential, as, otherwise, the overlapping circuits (of synonyms and semantically ambiguous items) could not be kept separate functionally.

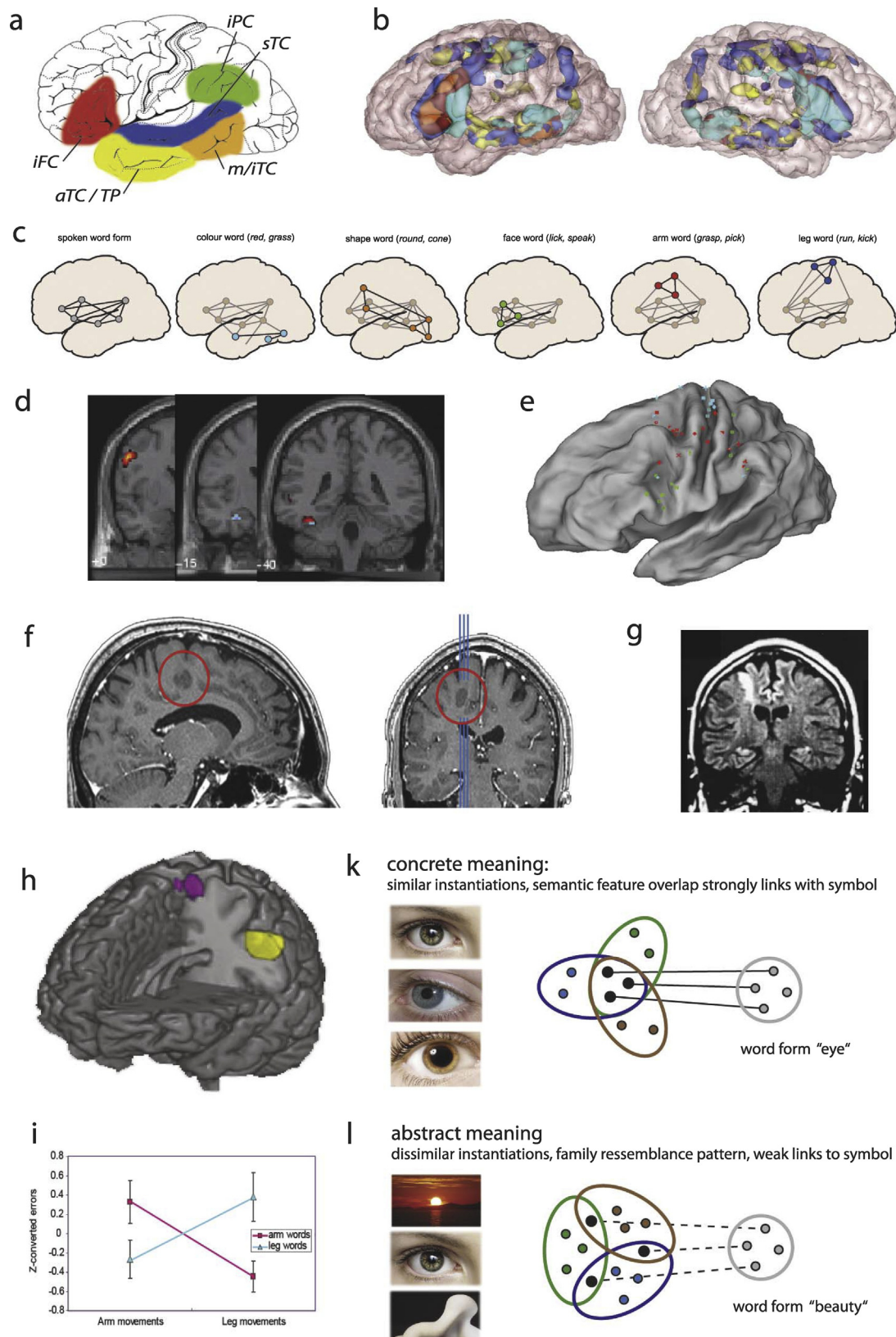
Because different types of information are most relevant in different types of semantic grounding and these may be processed by different parts of the brain, the APC model predicts that semantic circuits have different cortical topographies. Referential semantic links are implemented by connections between APCs for word forms in the central perisylvian system and object representations in the ventrottemporal visual stream (Goodale and Milner, 1992) in temporo-occipital cortex (Pulvermüller, 1999). In contrast, words typically used to speak about actions can be modeled by interlinked perisylvian circuits and action representations extending into the dorsolateral frontoparietal action processing system (Jeannerod et al., 1995) of the human brain (Pulvermüller, 1999). The learning of abstract words must take into account the variability of semantic contexts typical of their usage and their semantic links may therefore ‘detach’ to a degree from sensory and motor neurons (Section 3.2.1, Pulvermüller, 2013a).

These postulates, as most of the other proposals discussed in this article, follow directly from the neurobiological correlation learning principle and cortical area and connectivity structure. In learning situations that provide semantic grounding of words used to speak about specific actions, these action words are typically produced or perceived just before, during or just after performance of a related action (i.e., the word “grasp” just before grasping, Tomasello, 1995; Tomasello and Kruger, 1992). Therefore, specific strengthening of connections between action representations in motor, premotor and adjacent prefrontal cortex and central perisylvian linguistic circuits yields the formation of higher-order action-semantic representations. In addition, as visual motion features of actions are processed in the perception of actions, neurons in posterior-superior temporal areas (in and close to motion processing area MT) likewise become part of the circuits of action word. In semantic grounding of object related words, associative learning between visual representations of word-related objects and word forms yields referential visual-semantic circuits, which are spread out over central perisylvian cortex and areas of the ventral visual stream. Dorsal frontal and parietal action-related systems can be involved too, especially if the words’

referent objects have action affordances, as, for example, in the case of food or tool names. The emerging semantic circuits interlinking information about a word form and aspects of its meaning provide a mechanism for the spontaneous understanding of word meaning upon hearing or reading symbols, a sort of semantic pattern completion (Barsalou, 2013; Pulvermüller, 1999). Note that associative co-occurrence learning is necessary, but not sufficient, for this type of semantic grounding. The fire-wire-together rule alone would falsely predict that a function word (for example the article “the”) or an inflectional affix (third person “-s”), which occur in various kinds of conceptual contexts, associate with all of them, thus leading to a semantically overloaded item. In contrast, the out-of-sync-delink terms of the correlation learning rule (Section 2.2) do not allow these functional items to connect with specific conceptual circuits, because the large majority of contexts would lead to weakening of most individual contextually-related connections (Pulvermüller, 1999).

Semantic grounding can also be indirect when word meaning is learned from context and a ‘grounding kernel’ of items learned in action-perception contexts is available (Cangelosi, 2010; Vincent-Lamarre et al., 2016). In this case, correlated firing of new word form circuits in perisylvian space and context-related semantic activity outside leads to secondary learning of referential semantic links of novel symbols (Pulvermüller, 2013a). Over and above the dorsal (action) vs. ventral (vision) stream subdivision for semantics, the APC-and-correlation model postulates semantic links into all other modality-preferential brain systems for mapping odor, taste, auditory, and somatosensory aspects of meaning. Furthermore, more fine-grained topographical distinctions can be made based on more fine-grained semantic differences.

The general prediction from this model, that the processing of words with different meanings activates the brain in different ways and at partially different sites, has received strong support from a range of imaging studies. It appears that aspects of the meaning of words can be ‘read from’ word-induced activation topographies (Fig. 6a, Huth et al., 2016; Mitchell et al., 2008; Pulvermüller et al., 2009a). Also, and excitingly, semantic similarity between words seems to be reflected in the similarity of word-elicited distributed patterns of brain activation (Carlson et al., 2013; Carota et al., 2017). Likewise, more specific predictions of the semantic APC model received empirical support. Experimental data indeed show that sensorimotor systems are differentially activated by words, phrases and sentences and provide semantic information about category-specific semantic information (Binder and Desai, 2011; Fernandino et al., 2015). Modality-specific sensorimotor areas reflect language-related knowledge about actions (Hauk et al., 2004; Martin et al., 1996) and visual (Damasio et al., 1996; Pulvermüller and Hauk, 2006; Simmons et al., 2007), auditory (Kiefer et al., 2008), somatosensory, olfactory and gustatory features of objects (Barrós-Loscertales et al., 2012; Gonzalez et al., 2006). The processing of emotion related words involves part of the limbic system and, interestingly, motor areas for those body parts with which emotions are normally expressed (i.e., face and upper extremities, Moseley et al., 2012; Vigliocco et al., 2014). It is noteworthy that, in neurophysiological studies, semantically related activations in modality-preferential areas emerged as early as the earliest indexes of meaning processing so far documented in the human brain (Grisoni et al., 2016; Moseley et al., 2013; Pulvermüller et al., 2009b, 2005b; Shtyrov et al., 2014; van Elk et al., 2010), and, in particular, as early as brain indexes of abstract semantics in multimodal connector hubs (Boulenger et al., 2012). Therefore, these modality-preferential area activations, including the motor system, are unlikely to index post-understanding processes. More plausibly, they are part of the earliest semantic access processes.



**Fig. 6.** Semantic grounding in action and perception systems: a. Areas proposed in the literature to process semantic meaning across categories. b. Cluster analysis reveals left-perisylvian activation for all word types (in brown and orange) and, in addition, widespread topographically specific brain activation for different semantic word types (face-related action words in light blue, arm words in purple, leg words in yellow, abstract shape words in dark blue). c. Cell assembly model of cortical circuits for different semantic word types. d. Foci of significantly different brain activity elicited by color (in red) and shape words (in blue, Pulvermüller and Hauk, 2006) and e. by action words related to the face (green), hand (red) and leg (blue) as shown by a range of studies from different labs (from Carota et al., 2012). f. Lesion in the white matter underlying the hand motor cortex in a patient with specific deficit in tool word recognition (from Dreyer et al., 2015) and g. in the motor and premotor cortex of a patient with specific action verb processing deficit (from Neinger and Pulvermüller, 2001). h. Semantic priming in the motor system, as revealed by EEG source localization. Face and leg motor regions,

In functional imaging, the semantic maps of word categories can be quite locally specific and fine-grained as demonstrated by the observation of ‘semantic motor somatotopy’. Words normally used to speak about actions performed with the mouth/face, hands/arms and feet/legs activate motor and premotor cortex in a somatotopic manner (Fig. 6, Aziz-Zadeh et al., 2006; Carota et al., 2012; Hauk et al., 2004; Hauk and Pulvermüller, 2004; Kemmerer et al., 2008; Pulvermüller et al., 2001a, 2009a; Raposo et al., 2009; Shtyrov et al., 2014, 2004; Tettamanti et al., 2005). A degree of variability of the exact loci of semantically-related motor system activations across individuals and studies has been noted (Caramazza et al., 2014), although the well-known use-dependence and plasticity of sensorimotor maps implies such variability, especially in populations with different motor proficiency and language use (see, for example, Elbert et al., 1995). In this context, it is exciting to take a closer look at specialist populations with professional experience with specific actions, such as sportsmen, where neuronal plasticity is also reflected in language-elicited semantic brain responses (Beilock et al., 2008; Lyons et al., 2010). Crucially, results from different labs consistently show somatotopic motor activation indexing the meaning of action-related words (Fig. 6e, see also Carota et al., 2012; Kemmerer, 2015a). To what degree activity in modality-preferential areas is present and contributes to idiomatic and metaphorical language processing is a topic of current debate (Boulenger et al., 2009, 2012; Citron and Goldberg, 2014; Desai et al., 2013; Kemmerer, 2015a).

Are motor and sensory systems of special and critical relevance for understanding specific semantic types of linguistic information?: Causal effects of motor and sensory systems on the recognition and comprehension of specific semantic word types (e.g., action and sound words) have been shown using patient studies (Bak, 2013; Bak et al., 2001; Cappa and Pulvermüller, 2012; Cotelli et al., 2006, 2007; Dreyer et al., 2015; Kemmerer et al., 2012; Tranel et al., 2003; Trumpp et al., 2013), which rule out the suggestion that the respective activations might be ‘epiphenomenal’ or functionally independent of symbolic processing. Some of these lesion results focused on large lexical word categories, such as nouns and verbs, which are semantically different (typical semantic relationship to objects vs. actions) but have different grammatical function and lexical properties too (e.g., Damasio & Tranel, 1993; Neining & Pulvermüller, 2003; Gleichgerrcht et al., 2016). Whereas these results are open to interpretations in terms of semantics or other lexical features (see Bird et al., 2000), other findings enforce a semantic interpretation, because relevant lexical and grammatical variables had been controlled (e.g., Trumpp et al., 2013; Dreyer et al., 2015). The claim that motor system lesions may only modulate action word comprehension in a task-specific manner (Hickok, 2014), does not stand up to scrutiny: the effects across comprehension tasks appear more consistent after inferior frontal and precentral lesions than with posterior temporal ones (Kemmerer, 2015a, 2015b; Kemmerer et al., 2012). Although patients with brain lesions in the inferior frontal and motor cortex frequently suffer from deficits in action word processing (Bak, 2013; Kemmerer et al., 2012), they rarely show fine-grained somatotopic mapping differences (e.g., between arm- vs. face-related words, Arevalo et al., 2012), a possible reason being the large size of most lesions. Still, category-specific semantic deficits restricted to hand-related action verbs and (equally hand-related) tool nouns following lesions in sensorimotor cortex and in white

matter close to hand motor cortex have been documented (e.g., Dreyer et al., 2015; Neining & Pulvermüller, 2001).

In healthy people, focal stimulation of sensorimotor regions with TMS and tDCS (transcranial direct current stimulation) can differentially influence the processing of specific semantic categories, thus showing category-specific causal effects of the motor system on language processing (Gerfo et al., 2008; Kuipers et al., 2013; Liuzzi et al., 2010; Pulvermüller et al., 2005a; Repetto et al., 2013; Willems et al., 2011). Once again, a degree of variability in the patterns of results has been noted (Papeo et al., 2013), in particular, both facilitatory (e.g., Pulvermüller et al., 2005a; Willems et al., 2011) and inhibitory effects (Gerfo et al., 2008; Repetto et al., 2013) of TMS to motor regions on the processing of action-related language were found. Such discrepancies may, in part, be explained by the application of different TMS methods and by other features of the paradigms applied (for discussion, see, for example, Boulenger et al., 2006). Together with the research summarized in Section 3.1.1, these results confirm the role of central and precentral cortex in language and in the processing of symbols semantically related to human actions. Interestingly, one study based the latter conclusion on evidence from healthy subjects engaged in motor training, reporting that such activity improved action language comprehension (Locatelli et al., 2012). Further support for functional interaction between modality-preferential sensorimotor systems and semantics comes from a wealth of behavioral studies in healthy individuals, where, for example, performance of overt motor movements influenced the ability to process action words and sentences (Fischer and Zwaan, 2008; Glenberg and Gallese, 2012; Glenberg et al., 2008; Shebani and Pulvermüller, 2013; Zwaan, 2014).

The involvement of sensorimotor systems in meaning processing is modulated by context (Kiefer and Pulvermüller, 2012; Lebois et al., 2015). In a seminal study, Kiefer and his colleagues investigated the processing of action- and object-related words in semantic feature verification tasks emphasizing respectively their action-related or visually-related features (Hoening et al., 2008). This study showed modulation of word-evoked brain activity depending on whether dominant or non-dominant semantic features of the words were task-relevant. Other work also indicated task-dependence of semantic sensorimotor activation reflecting semantic aspects of words, which were, for example, more pronounced in a semantic task than during letter detection (Tomasino et al., 2007; van Dam et al., 2012). Likewise, behavioral experiments suggest that even the most prominent semantic features of a concept or word are not activated in all contexts, for example when task requirements draw away attention from these semantic features (Kiefer et al., 2012, 2016; Lebois et al., 2015). Grammatical information immanent to a sentence also modulates the semantic brain response as it is manifest, for example, in motor area activation (Papeo et al., 2011; Pulvermüller et al., 2012a). These proofs of the flexibility of sensorimotor activations with different tasks, attention levels and contexts (see also Hauk and Tschentscher, 2013; van Dam et al., 2014; Willems and Casasanto, 2011) require neurobiological models that account for them. The action perception perspective of Section 2 handles these effects by area-specific control mechanisms (e.g., ‘low attention’/strong inhibitory feedback control in areas relevant for semantic processing in non-semantic tasks), cumulative effects of APC

respectively, showed relatively reduced activity when action words were preceded by a body-part congruent action sound (from Grisoni et al., 2016). i. Complex movements of the arms and feet impair the processing of arm- and leg-related words in healthy subjects (from Shebani and Pulvermüller, 2013). k-l. Semantic grounding of concrete and abstract symbols. Whereas a common set of sensorimotor features may be shared by the instantiations of concrete words such as ‘eye’, and the corresponding neuronal elements may therefore strongly bind with the word form circuit (h), a pattern of family resemblance without feature overlap applies to the instantiations of an abstract term, thus leading to weaker correlation-based sensorimotor links but, interestingly, to a new combination of ‘anchor’ features, which may not be present in the environment (for further explanation, see text) (from Pulvermüller, 2013a).



activations and priming interactions between them (Garagnani and Pulvermüller, 2016; Pulvermüller et al., 2014a).

As there is still skepticism about the role of modality-preferential cortices in semantic processing (for discussion, see Section 3.2.2), some recent work focused on a widely accepted and genuine brain index of meaning processing, namely on semantic priming, and asked whether activity in the motor and somatosensory cortex may directly indicate the semantic relationship between meaningful stimuli. Although semantic priming was first established at the behavioral level, the reduced neuronal activation of a symbol presented in the context of a semantically related one (as compared with an unrelated one) can be taken as a neurophysiological index of semantic relatedness and priming. The relevance for sensorimotor cortex for semantic priming was shown by TMS stimulation to the hand motor area, which abolished the semantic priming effect otherwise present in a late component of the event-related potential (Kuipers et al., 2013). Semantic priming effects are established in semantic ‘hub’ areas, such as anterior temporal lobe, as well as in frontal and parietal cortices (Ulrich et al., 2013). However, physiological priming in hand motor cortex could be shown for finger movements and hand action words (Mollo et al., 2016). When face- and leg-related action sounds were used to prime subsequent face and leg words, body-part congruency led to somatotopic priming effects in the face and leg areas respectively (see Figure 6h, Grisoni et al., 2016). These physiological priming studies provide strong evidence that the processing of the semantic relatedness between signs and symbols engages, amongst other areas, the sensorimotor system. In this sense, there seems to be ‘meaning in the motor system’. Similar to other context effects on semantic activation, physiological priming proves the flexibility of the brain indices of semantics.

In spite of their flexibility, it still appears appropriate to call the semantic activations and priming effects in modality-preferential areas ‘automatic’ in the sense that no focusing of attention of semantic aspects of the stimuli is necessary to obtain them (Pulvermüller, 2005). The evidence for this statement comes from studies in which subjects instructed not to attend to the stimulus symbols passively heard words and were even distracted from processing them. Brain responses indexing specific semantic differences, for example between the body-part related action word types, were still manifest under such ‘non-attend’ conditions, thus showing a degree of automaticity of the semantic activations (Grisoni et al., 2016; Pulvermüller et al., 2005b; Shtyrov et al., 2014, 2004). Passive perception of symbols without any overt motor task likewise leads to spontaneous area-specific activations indexing meaning aspects of other semantic word types, for example form- and color-related words (Moscoso Del Prado Martin et al., 2006), or gustation and emotion words (Barrós-Loscertales et al., 2012; Moseley et al., 2012). The automaticity and earliness of semantic circuit activations (Moseley et al., 2013) suggested by these and similar studies do not imply the impossibility to use tasks that suppress or strongly counteract semantic processing when linguistic symbols are presented. For example, if the task is to judge visual features of a letter string or the phonological makeup of a spoken word, semantic activations may be reduced or absent (Kiefer et al., 2016; Kiefer and Pulvermüller, 2012; Lebois et al., 2015). The APC framework explains such modulation by priming and/or gain modulation in specific areas relevant for semantic processing (see Section 3.2). Still, it appears as important to distinguish the use of the term ‘automatic’ in the sense of ‘impossible to suppress’ from the more moderate task-free, or ‘passive’, automaticity of semantic processing shown by non-attend and passive paradigms. It emerges from the latter that semantic processes are, in spite of their suppressability, *passively automatic* pattern completion processes (Pulvermüller and Shtyrov, 2006).

The summarized results are difficult to explain for classic symbolic conceptual theories which consider meanings as stored entries in a symbolic system functionally segregated from sensory and motor systems (for discussion, see Barsalou, 2008; Binder and Desai, 2011; Glenberg and Gallese, 2012; Kiefer and Pulvermüller, 2012). The activations indicating the meaning of utterances can reach into motor and sensory cortex, where magnetic stimulation or lesion functionally influence the processing of specific word types. Classic symbolic system theories did not predict or explain this pattern of results. In essence, semantic processes are not segregated from sensorimotor mechanisms, but, as correlation learning and long-distance cortical connectivity imply, functionally interlinked.

The hub-and-spoke model of semantic processing (Patterson et al., 2007; Ralph et al., 2017) provides an alternative solution for the symbol- and task-specific distributedness of semantic processing. Many aspects of that model resemble the semantic topographies postulated in this section. However, there are three issues which I would like to mention in light of that well-developed theory: First, it allows for only one semantic hub in anterior temporal lobe, where semantic similarity is computed across items; however, there is, as we saw, evidence for a range of areas crucial for semantics, where lesions or TMS seemingly affect semantic processing across meaning types (cf. Pulvermüller, 2013a). Second, it claims general semantic impairments after anterior lobe lesions, whereas several studies indicate some category-specificity of semantic dysfunction arising from such lesions (Shallice, 1988; Shebani et al., 2017; Warrington and Shallice, 1984). The hub-and-spoke model tries to explain this by disease progressing from anterior hub to more posterior temporal category areas. However, the pattern of disease progression speaks against this position: Semantic Dementia, the primary progressive aphasia arising from anterior temporal lobe lesion, tends to include a relatively prominent deficit for color words (compared with other semantic word types) at very early stages of the disease and, with disease progression, such category-specificity vanishes (Pulvermüller et al., 2010). Initial hub involvement and later lesion spreading to category specific spokes predicts the opposite: a general semantic deficit first and category specificity to emerge later. And finally, the model, like many others, does not provide an explanation for the locus of semantics: it does not justify why anterior temporal lobe should be the only hub and frontal, posterior temporal or parietal areas might not house hubs too. Recent simulation studies provide a justification in terms of neuroscience principles for an anterior temporal semantic hub along with other semantic integration sites (see Section 3.2.3, Garagnani and Pulvermüller, 2016; Tomasello et al., 2017).

### 3.2.1. Abstract culture-specific concepts as a test case

Some symbolic theories proposed that concepts have constituted an evolutionary advantage, so that different mechanisms for them have been built into the human brain (for discussion, see, for example, Caramazza and Mahon, 2003). Accordingly, the neural mechanisms, for example, for animal, plant and tool concepts are different irrespective of their sensorimotor attributes. While this evolutionary-conceptual approach might be applicable to concepts shared across human cultures, any genetic programming of concepts appears less promising for concepts that vary across cultures, such as those denoted by the words “beauty”, “game” or “justice”.

A perspective on learning culture-specific abstract concepts is opened by correlation learning. Of special relevance in this context is, once again, the ‘anti-Hebb’ part of the correlation rule (Section 2.2), which implies weakening of links in case of independent or anti-correlated activation. Concrete terms such as “eye” are typically used to speak about quite similar objects.

These objects seem to share perceptual features and these may become part of the semantic representation of the word. The same applies for actions words such as “grasp”, where the grasping movements may be to a degree similar although the to-be-grasped objects vary widely. In contrast, abstract words such as “beauty” are more variable in their usage; they can be used to speak about ‘things’ that look quite differently and lack common semantic-perceptual features (Baker and Hacker, 2009; Wittgenstein, 1953). The perceptual and action features of instantiations of beauty vary in a context-dependent manner, so no common perceptual or motor features are shared across all usages. Features that make a landscape beautiful can be quite different from those that glamourize a face.<sup>12</sup> Still, some beautiful faces and landscapes respectively may share perceptual features, and different perceptual similarities may hold for beautiful eyes and sculptures. Therefore, a pattern of ‘family resemblance’ without general feature overlap, but, instead, partial feature sharing between the instantiations of an abstract concept, results. If sensorimotor features are processed by neurons that become linked to word form circuits by correlation, this variability implies that only those partial overlap features become ‘anchors’ for linguistic connections to the respective symbol or word forms; also, these semantic links are relatively weak due to low correlation. Now, importantly, as the anchor features come from different domains and are dissimilar, their combination may constitute new abstract concepts that do not have any correlate in real life (Fig. 6i). The correlation learning principle therefore offers a mechanism for linking heterogeneous semantic features to the same linguistic expression and therefore for abstract concept formation (Pulvermüller, 2013a). Crucially, the meaning of these abstract terms can be grounded in experiences although the abstract concept itself is not part of reality, or the nervous system, in the first place. It emerges as a result of correlation learning. Thereby, an account is possible of how and why different abstract concepts, such as those of beauty or freedom, are being established at different times and in different social environments.

### 3.2.2. Semantic grounding and the role of modality-preferential areas in language processing

Distributed action perception circuits for binding symbols and their referential meaning provide a neurobiological mechanism for conceptual grounding and ‘embodiment’ as it has been proposed by a range of authors (Barsalou, 2008; Fischer and Zwaan, 2008; Gallese and Lakoff, 2005; Glenberg and Gallese, 2012; Kiefer and Pulvermüller, 2012; Lakoff, 1987; Meteyard et al., 2012). Semantic grounding can be seen as the extension of semantic circuits into sensory and motor areas of the human cortex. The category-specific activation of these areas while processing words, phrases and sentences with specific semantic content along with the category-specific semantic deficits following the lesion in and close to these areas can therefore be seen not only as support for category-specific semantics per se, but, in addition, as consistent with the semantic grounding or ‘embodiment’ perspective. APC formation and referential-semantic linkage between them determined by correlation learning principles effective in an anatomically pre-structured network provides a mechanistic account for both category-specificity and grounding.

As semantic or conceptual grounding theorists and their interpretation of new neurocognitive data have recently been

criticized, I will briefly comment on this discussion. Unfortunately, some critical statements seem to address simplified ideas, which differ from crucial claims immanent to semantic grounding. In particular, it has been argued that this position implies that knowledge of a concept is nothing but “the sum total of (. . .) sensory and motor experiences” (p. 130, Hickok, 2014) and “that conceptual content is reductively construed by information that is represented within the sensory and motor systems” (p. 59, Mahon and Caramazza, 2008).

These statements misrepresent the semantic grounding perspective as it is framed in a neurobiological perspective (see also Barsalou, 2016; Pulvermüller, 2013b). I have therefore proposed to dub the respective positions as “misembodiment” (Pulvermüller, 2013b). In particular, the first, “sum total”, statement ignores the role of correlation learning. When learning a grounding relation between the word “green” and various objects that exhibit the feature of greenness, it is exactly not the sum of all the objects and scenes experienced that are mapped on the symbol. Specifically, correlation learning guarantees that the feature ‘elongated’ of the cucumber, ‘round’ or ‘pear-shaped’ of the avocado and ‘flat’ of the leaf etc. are being disconnected from the semantic circuit of the word “green”. Therefore, it is false to state that a ‘sum’ of experiences is computed when the correlations of neuronal activity indexing semantic features are mapped. This may happen under special circumstances, where information about several typical referents is bound to an expression, or in the learning of proper names. Crucially, ‘parasitic’ semantic learning from context and the family resemblance patterns of abstract terms in real life situations lead to correlation-based selective feature extraction and to their non-linear combination, thus giving rise to previously unexperienced and even unexperienceable feature sets (see Section 3.1.1 and Cangelosi et al., 2002; Meteyard et al., 2012; Pulvermüller, 2012). These mechanisms are inadequately described by a linear summation process, although, in the grounded perspective, there is a significant contribution indeed of sensory and action-related information. In this context, it is also inadequate to speak about association in a behavioristic sense (see Section 2.2 for discussion). The associations and dissociations emerging in auto-associative neural networks as a result of correlation learning emerge between neurons and neuronal assemblies and these can be seen as neuronal correlates of perceptual features, combinatorial regularities or abstract concepts.

The second criticism-attracting statement about “reductive construal” of meaning and concepts in the sensory and motor systems is equally inappropriate, thus misrepresenting dominant grounding views. The grounding (or embodiment) claim is that distributed semantic representations reach into modality-preferential sensory and motor systems, not that they are restricted to these systems (see Barsalou et al., 2003; Pulvermüller, 1999). This claim is motivated by correlation learning and the area structure and connectivity of the relevant cortical areas, and especially by the fact that, in order to link information between sensory and motor areas, a pathway via connector hubs must be taken (Section 2.4, see Garagnani et al., 2016; Tomasello et al., 2017). An important point of the APC framework is that information in the sensory and motor systems drive and influence the formation of concepts and semantic representations, that of abstract concepts included. But there is surprisingly little discussion addressing this relevant point.

The present gridlock between grounded-interaction and symbolic-segregation approaches can be illustrated using ‘compromise solutions’ offered by representatives of the latter. In their 2008 paper, Mahon and Caramazza offer a “grounding by interaction” approach, which, in their view, links the merits of segregated concept processing in a dedicated symbolic system with semantic-conceptual grounding (Mahon and Caramazza,

<sup>12</sup> It would therefore be a rather strong claim to state that the concept they relate to is the same across all usages. Note that some authors indeed stated that the concept a word relates to is the same across all of its usages and contexts (for example Mahon, 2015; Mahon and Caramazza, 2008), a proposition which would need justification.

2008; Mahon, 2015). While symbolic representations are still maintained to be represented at an amodal level normally detached from sensory and motor systems, the grounding problem is claimed to have an easy solution: “A line is drawn from the concept to the corresponding sensory/motor information”, but this line has a “clutch” to be freely removed or added whenever desired (Mahon, 2015). This proposal remains silent about neurobiological mechanisms and offers all possibilities for experimental predictions, thus leaving questions about its testability unanswered. It also has inherent logical flaws. In particular, Mahon and Caramazza admit that, in the hooked-up case, conceptual-semantic and sensorimotor information interact, but they still claim that the unique locus of semantic-conceptual processing is in the amodal conceptual system, and not in the sensory or motor components. This is a logical inconsistency, because, if there is between-systems interaction and exchange of information, including truly semantic information, there must be semantic-conceptual information (representation and processing) also in the sensory/motor system, and therefore sensorimotor systems would be a genuine locus of semantics, a possibility the authors wish to deny (for further discussion of this point, see Pulvermüller, 2013b). The authors’ seem to postulate interaction without information exchange, a logical impossibility.

In defense of their position, Mahon recently claimed that, in spite of information exchange and interaction, the “format” of the representations is different in an amodal conceptual and a (by-line-connected) sensorimotor system (Mahon, 2015; Mahon and Hickok, 2016). However, what format might that be? The language of neurons can be described in a propositional format (Braitenberg and Pulvermüller, 1992). If a cell in visual cortex fires, this may mean “that there is something red at coordinate x/y of the visual field” and a higher visual area neuron may represent “that there is something red somewhere”. With eyes closed, the activation of these neurons would still imply mental reality (‘simulation’) of redness. Now, if a neuron in the anterior temporal semantic hub represents the semantic feature [+RED] in an ‘amodal format’ and connects to the higher visual neuron through a ‘line’ – what might the difference in format between these neurons actually mean? The visual system and hub neurons would be active together and convey the same information. One may even state that, for the anterior temporal hub neuron to represent [+RED], the grounding link is necessary, although threshold regulation mechanisms (Section 2.5) may reduce sensory activations in specific task conditions (cf. Mahon’s clutch). The “grounding by interaction” approach can only account for grounding if the connection and interaction between sensorimotor and semantic mechanisms is seen as fundamental, and if their consequences for information mixing and merging are construed appropriately. Thus, omitting metaphorical descriptions and spelled out in neuronal terms, it may be an action perception model.

### 3.2.3. Neurobiologically constrained modelling of semantic grounding

One may still question whether the statement about the relevance of both semantic hubs and category-preferential semantic areas is firmly rooted in solid neurobiological ground. Therefore, Garagnani, Tomasello and colleagues performed a series of neurocomputational studies to explore the formation of semantic circuits in a neuroanatomically and neurophysiologically constrained network architecture (Garagnani and Pulvermüller, 2016; Tomasello et al., 2017). The extended architecture incorporated the core perisylvian language areas shown in Fig. 2 plus temporo-occipital ventral stream visual areas and dorsal motor, premotor and prefrontal areas. To imitate semantic referential grounding of object and action words, the network was presented with correlated activity patterns in articulatory and auditory ‘cortex’ (as in the previous work discussed in Section 2.4) and, in

addition, either in visual or hand motor ‘cortex’. The simulations of co-occurrences between words and objects vs. actions led to the formation of distributed circuits, which, in the case of the object words, incorporated ‘neurons’ in the visual ‘areas’, or reached into dorsolateral motor ‘cortex’ for action words. These semantic circuits were indeed not restricted to modality-preferential areas. Instead, the distributed circuits were found to have most of their neurons located in those areas of the network that served as between-modality connector hubs. Importantly, the semantic model developed not only one semantic hub, but four different ones, in inferior prefrontal, dorsolateral prefrontal, inferior anterior and superior temporal cortex, which is consistent with the evidence from neuroscience studies of semantics (cf. Fig. 6a; note that parietal cortex was not included in the simulations). The semantic model provides a clear demonstration that, using neurobiological constraints, the claim of semantic representations exclusively in sensory or motor systems, or in both, is not a reasonable one. As this position is unrealistic, it should not be used as a straw man in discussions about semantic grounding (but see Section 3.2.2 and Caramazza et al., 2014; Hickok, 2014).

These simulations of referential semantic grounding using a neurobiologically constrained computational model show that the action perception perspective is compatible with, and can explain aspects of, traditional brain language models of semantics, where semantic centers or ‘hubs’ are postulated to provide the locus of general semantic and conceptual processing. Importantly, the model also offers a formal solution to the current dispute between scientists in the field of neurosemantics, who, as mentioned, diverge in their opinions about the true locus of the semantic ‘hub’ or ‘interface’, arguing in favor of frontal (Bookheimer, 2002; Devlin et al., 2003), anterior-inferior temporal (Patterson et al., 2007; Ralph et al., 2017) or posterior-middle/superior temporal lobe (Hickok, 2014; Price, 2000). In the semantic model of Garagnani, all of these regions take the role of a semantic integration space across the kernel parts of semantic circuits (Garagnani and Pulvermüller, 2016). In addition, this model answers the question of why semantic hubs emerge by referring to neuroanatomical connection structure and correlated activity. Previous simulations did not address this question. For example, work by Ueno and colleagues “implemented the vATL (ventral anterior temporal lobe) semantic system (. . .), (which) was set to generate semantic outputs for comprehension and provided the semantic input for speaking/naming” (Ueno et al., 2011), thus a-priori implementing the semantic system in one of the semantic hubs under discussion. Therefore, quite unsurprisingly, semantic deficits arose from lesions to the pre-defined ‘semantic hub’ of the model. This and similar simulation approaches do not provide explanations of the brain loci of semantics. In contrast, APC model simulations explain the co-emergence of areas for category-specific semantic processing and for general semantics (‘hubs’) in the learning and semantic grounding of words and symbols (Garagnani and Pulvermüller, 2016; Tomasello et al., 2017).<sup>13</sup>

### 3.2.4. Objections on empirical grounds

True objections against the current evidence for the grounded action perception perspective are three fold, addressing, respectively, (1) the replicability of crucial findings, (2) their interpretability and (3) their explanation.

Replicability: It has been claimed that, across studies, activation to action-related language in frontoparietal sensorimotor cortex is not reliable (Watson et al., 2013). However, this review mixed

<sup>13</sup> Similar to Garagnani and Pulvermüller (2016) and Tomasello et al. (2017), Chen et al. (2017) used a learning model guided by cortical connectivity to model category-specificity and semantic hub function of temporal and parietal areas.



**Table 2**

The dispute about semantic brain mechanisms has been perturbed by lack of clarity about the space of relevant hypotheses ('misembodiment'). Amodal symbolic ('disembodied') theories about semantics and concepts state that multimodal brain areas are the only seat of semantic processing (for an example, see [Bedny and Caramazza, 2011](#)). Some authors (including [Mahon and Caramazza, 2008](#), here abbreviated as 'MC') claimed that grounded action perception ('embodied') models place semantics *exclusively* in modality-specific (modality-preferential) systems such as the motor and sensory cortices (top panel), implying that a new proposal might be necessary to acknowledge both types of areas' relevance for concepts and semantics (upper table). However, this argument was flawed. The original claim offered by the grounded action perception and neural reuse perspective was (and still is) that both multimodal and modality-preferential areas take their shares in semantic and conceptual processing (bottom panel; for discussion, see text and [Pulvermüller, 2013a](#)). The latter possibility appears the only one consistent with the experimental evidence (Section 3.2).

MC THEORY SPACE		Semantic processing in modality preferential areas	
		No	Yes
Semantic processing in multimodal areas	No	(not available)	Grounded action perception theories
	Yes	Amodal symbolic theories	?
TRUE THEORY SPACE		Semantic processing in modality preferential areas	
		No	Yes
Semantic processing in multimodal areas	No	(not available)	(not available)
	Yes	Amodal symbolic theories	Grounded action perception theories

linguistic and nonlinguistic studies across a range of experimental paradigms, of which only a fraction presented linguistic stimuli. Those studies from their sample which included conditions with linguistic stimuli or comparisons between linguistic tasks, sensorimotor system activation to action language was present in all experiments applying passive paradigms, in which no overt responses were required (see [Table 3](#)). There were in fact only two linguistic studies in this review that used overt response paradigms and found significant sensorimotor activation to action words (compared with different control conditions). This result shows that an active task requiring overt responses makes it difficult to document fine-grained semantically related modulations of sensorimotor brain activity, possibly due to generally enhanced levels of response related motor activity. It is therefore advisable to use non-respond paradigms in investigation of the brain basis of action semantics. Note the parallel situation observed in the study of phonological effect in motor systems (Section 2.1 and [Schomers and Pulvermüller, 2016](#)).

One specific study used such a passive paradigm and claimed that they could not replicate body-part specific motor area activity for different semantic subtypes of action words ([Postle et al., 2008](#)). In this work, activation of body-part specific areas were analyzed separately for primary motor and premotor cortex, thus raising the question whether the lack of power implied by the subdivision into small areas might be behind the lack of significance. Therefore, results for primary and premotor areas were integrated in one statistical design. Word-category-elicited activity in their 'action observation' localiser ROIs was analysed using a repeated measures ANOVA (with the factors Semantic word category (face, arm, leg), Motor area (primary, secondary) and Somatotopy (face, arm, leg region)). A significant interaction of Semantic word category with Somatotopy emerged ( $F(4,64) = 3.8$ ,  $GG\text{-}eps = 0.64$ ,  $p = 0.022$ ), thus showing differential modulation of brain responses across the motor strip.<sup>14</sup> Therefore, the finding of motor system activity varying with the meaning of

action-related language appears replicable across studies, although not all studies showed significant word category differences in each section of the frontocentral cortex (see also [Carota et al., 2012](#); [Kemmerer and Gonzalez-Castillo, 2010](#)). It is also necessary to emphasize again the modulatory function of task, context and attention, which may enhance, reduce or remove motor system activation ([Hoening et al., 2008](#); [Liljestrom et al., 2008](#); [Pulvermüller et al., 2012a](#); [van Dam et al., 2012](#)), a fact accommodated by neurobiological models (Section 2.6). Nevertheless, motor system activity in processing action-related language seems to exhibit passive automaticity, insofar as it does not require tasks focusing attention on language, concepts or meaning ([Pulvermüller et al., 2005b](#); [Shtyrov et al., 2014, 2004](#); [Grisoni et al., 2016](#)).

(2) Theory space and interpretability: When experiments on motor and sensory cortex activation were first undertaken, the theory space had been subdivided in a straightforward manner. Brain-based abstract symbolic system theories postulated that semantic processes are confined to an amodal conceptual system in one or more semantic hubs, which lie in multimodal cortex exclusively (for an explicit statement of this position, see for example [Bedny and Caramazza, 2011](#)). Therefore, this 'amodal', 'disembodied' or separation approach predicted semantic activation in multimodal cortex, but not in sensory and motor areas. In contrast, the position of grounding and action perception integration states that semantic circuits are distributed across a range of areas and also *reach into* sensory and motor modality-preferential systems; thus, it predicts semantic activation to extend into, but not to be restricted to, the cortices classically seen as modality-specific for motor, visual, auditory etc. information. (see also Section 3.2.2)<sup>15</sup>

After the critical prediction of modality-preferential area activation to specific meaningful stimuli had been confirmed by a large number of studies, an attempt was made to re-divide up the theory space, now incorrectly claiming that action perception accounts would not be compatible with multimodal area activation in semantic processing. After this recasting of theory space, it was also claimed that, within a slight extension of an abstract symbolic system theory, a range of sensory and motor area activations reflecting aspects of semantic meaning might also be explainable. These assumed mechanisms are a) run away

<sup>14</sup> I thank Greig de Zubicaray for allowing me to re-analyze the data. In a recent paper, [de Zubicaray et al. \(2013\)](#) replicated precentral activation to action verbs compared with object nouns. These authors now claimed that their additional finding, that also meaningless pseudowords – which were phonologically similar to action verbs – showed motor system activation, argues against a semantic effect. However, this argument doesn't appear to be convincing. When perceived, pseudowords likely partially activate their 'cohorts' or 'neighbors' of phonologically or orthographically similar real word representations ([Garagnani et al., 2008](#); [Harley, 2008](#); [Marslen-Wilson and Warren, 1994](#)). Therefore, pseudowords with phonological similarity to action verbs may activate motor systems by way of the action semantic links of their phonologically and orthographically related lexical items.

<sup>15</sup> Classic sensorimotor accounts of category specific semantic deficits focused on the effects of lesions including both primary and convergence areas, for example inferior temporal or frontoparietal cortices ([Damasio et al., 1996](#); [Kiefer and Pulvermüller, 2012](#); [Martin, 2007](#); [Shallice, 1988](#)).

'overspill' activity from genuine semantic loci and b) optional semantic 'enrichment' or conceptual 'colouring' (see Mahon and Caramazza, 2008). Overspill activation makes it difficult to account for any specificity of activity indexing semantic features: If unspecific 'overspill' was relevant, why should the word "kick" activate the motor system rather than entail 'overspill' to olfactory cortex or amygdala, and why should odour words produce the opposite specific effect? In order to consider this approach an account, predictive specificity would be needed. Optional semantics or conceptual colouring gives up the original 'disembodied' position whereby semantics and concept processing are restricted to the 'amodal symbolic system', now admitting that some semantic processing is indeed carried by modality-preferential systems. Although there are logical inconsistencies in the 'colouring' argument, this is, in

part, a respectable approach, which might be testable after being worked out in detail in order to assess whether and how it might diverge from the 'embodiment' or grounded action perception position (see Section 3.2.3). In summary, this 'embodiment' debate seems to suffer from inadequate recasting of the theory spaces, because established positions have just been misunderstood or misrepresented as already discussed in Section 3.2.2. The true and distorted hypothesis spaces are displayed in Table 2.

(3) Explanation: Considering that the processing of semantic aspects of a symbol or construction are reflected in both multimodal and modality-preferential areas, some authors chose to claim that the former areas are the 'real gateway to understanding', whereas the involvement of the latter is due to 'mere association' but *not* semantic processing (Hickok, 2014).

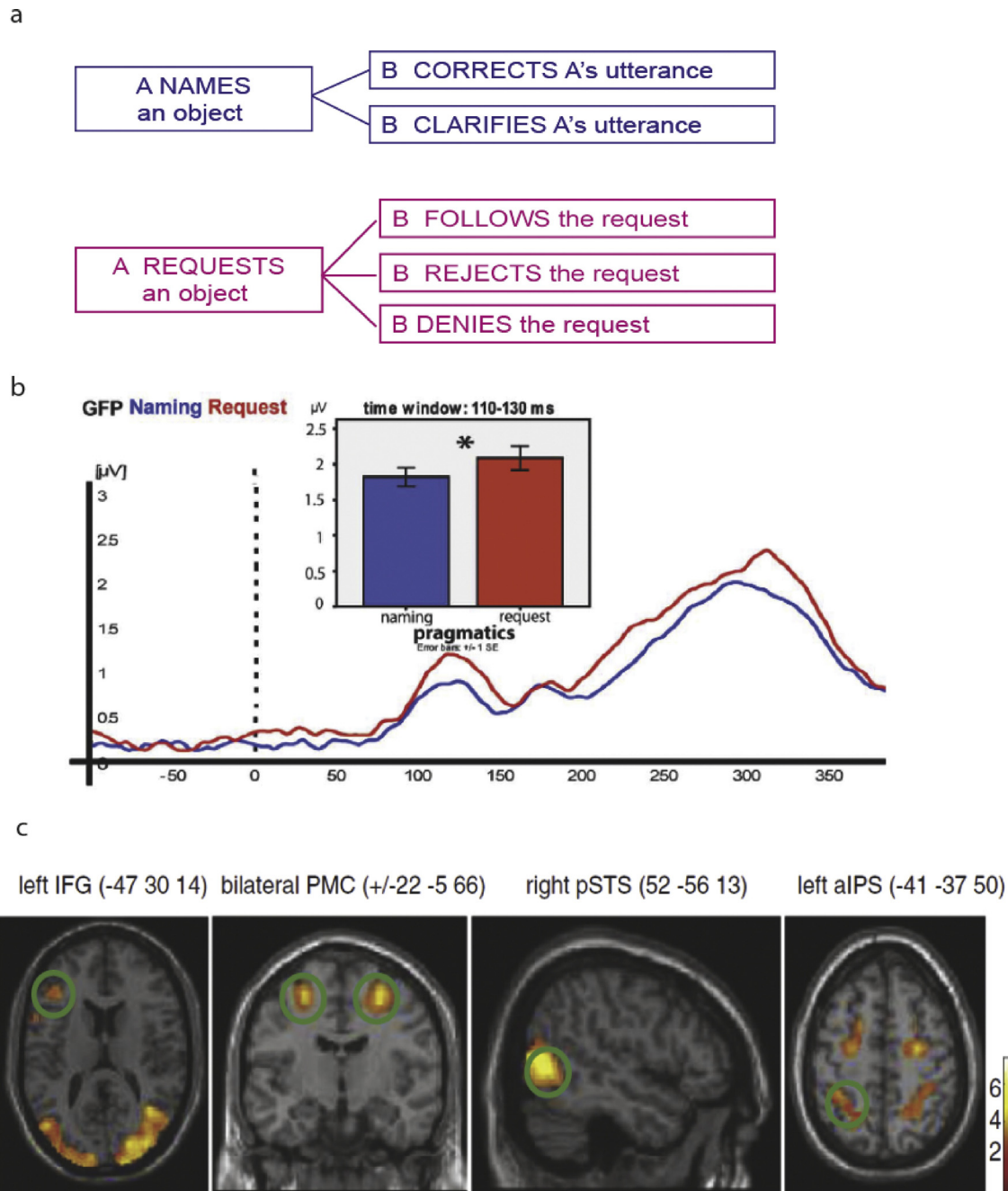
**Table 3**

Overview of results from studies reviewed by Watson and colleagues (Watson et al., 2013). Only those works that used language stimuli or different language tasks are listed. From left to right, the table gives (1) First author and publication year, (2) the number of participants studied (N), (3) the task, (4) the number of stimuli per critical category mentioned in the table, (5) the response requested from subjects, (6) whether responses were required in any catch trials, (7) the critical stimuli, (8) the comparison condition, (9) whether linguistic stimuli were repeated, (10–14) whether activity was found in the following areas in the left hemisphere: pIF – posterior inferior frontal, SMC – sensorimotor cortex, iPC – inferior parietal cortex, pS/MTC – posterior superior/middle cortex, aTC – anterior temporal cortex. The numbers indicate stronger activity to critical stimuli compared with the comparison condition ('1') or the reverse ('-1'), or no difference ('0'), (15) the statistical criterion used (FWE/FDR correction, uncorrected p-value (unc), or regions of interest (ROI) analysis). The majority of studies show activity in both SMC and pS/MTC in the critical comparisons. Please note that all studies not requiring overt responses (see column (5)) revealed activity in sensorimotor cortex during critical stimulus processing. (Results were extracted from the following studies: Bedny et al., 2008, 2012; Berlinger et al., 2008; Hauk et al., 2004; Kemmerer et al., 2008; Liljestrom et al., 2008; Raposo et al., 2009; Rodriguez-Ferreiro et al., 2011; Rüschemeyer et al., 2007; Tyler et al., 2003; van Dam et al., 2010, 2012).

1st author/year	N	Stim /cat	Task	Response	Resp catch trials	critical Stimuli	compared with	stimuli repeated	pIF	SMC	iPC	pS/MTC	aTC	significance
Bedny (2008)	12	100	semantic judgment	button press	.	verbs	nouns	Yes	1	0	0	1	0	corrected
Bedny (2012)	21	50	semantic judgment	button press	.	high motion action verbs	nonliving object nouns	Yes	0	0	0	1	0	corrected
Berlinger (2008)	12	45	verb/noun generation	No	.	nouns(> verbs)	verbs(> nouns)	No	-1	1	1	0	0	corrected
Hauk (2004)	12	45	naming	No	.	actions	objects	No	1	1	1	1	0	corrected
	14	50	reading	No	.	face action verbs	low level baseline	No	1	0	0	1	0	corrected
		50	reading	No	.	arm action verbs	low level baseline	No	1	1	0	1	0	corrected
		50	reading	No	.	leg action verbs	low level baseline	No	1	1	0	1	0	corrected
Kemmerer (2008)	16	24	semantic judgement	button press	.	speaking verbs	low level baseline	Yes	1	0	0	1	0	.001 unc
		24	semantic judgement	button press	.	hitting verbs	low level baseline	Yes	1	1	0	1	0	.001 unc
		24	semantic judgement	button press	.	cutting verbs	low level baseline	Yes	1	1	1	1	1	.001 unc
		24	semantic judgement	button press	.	running verbs	low level baseline	Yes	1	1	0	1	1	.001 unc
		24	semantic judgement	button press	.	state change verbs	low level baseline	Yes	1	0	0	1	1	.001 unc
Liljestrom (2008)	15	100	Action/object naming	No	.	action images	action images	Yes	1	1	-1	1	1	corrected
Raposo (2009)	28		listening	No	.	arm words	nonaction words	Yes	0	1	0	0	0	corrected
	28		listening	No	.	leg words	nonaction words	Yes	0	1	0	0	0	corrected
Rodriguez-Ferreiro (2011)	14	80	Reading	No	Yes	concrete verbs	pseudowords	Yes	1	1	0	1	0	ROI
Rüschemeyer (2007)	20	23	Reading	No	Yes	action verbs	abstract verbs	No	0	1	1	0	0	corrected
Tyler (2003)	12	80	semantic categorisation	button press	.	tool action verbs	low level baseline	Yes	1	0	0	1	1	corrected
			semantic categorisation	button press	.	biological action verbs	low level baseline	Yes	1	0	0	1	1	corrected
			semantic categorisation	button press	.	tool nouns	low level baseline	Yes	1	0	0	1	1	corrected
			semantic categorisation	button press	.	animal verbs	low level baseline	Yes	1	0	0	1	1	corrected
van Dam (2010)	16	54	semantic categorisation	go/nogo button press	.	action verbs	abstract verbs	No	0	1	1	1	0	.001 unc
van Dam (2012)	12	25	semantic judgement	answer re color/action	.	action verbs	abstract verbs	No	0	0	1	1	0	ROI

Such a statement appears as arbitrary. Why should only sensory or motor areas reflect “associations”, whereas multimodal areas may carry concepts?” In Garagnani’s explicit neurocomputational model, both modality-preferential and multimodal area semantic activations result from correlation mapping (Garagnani and Pulvermüller, 2016; Tomasello et al., 2017, 2016). Crucially, Hickok’s critical statements seem to imply that association cannot be the basis of semantics. However, it is uncontroversial that correlation learning offers a mechanism by which the arbitrary relationship between symbols and their typical referents – and therefore one important semantic aspect – can be established in

language learning. In contrast to non-semantic arbitrary association, as it might be produced in a psychological experiment where words and objects are paired randomly, semantic links are based on collective agreement and are shared within a language community (see, for example, de Saussure, 1916). As information about correlations between symbols and their contexts as well as between symbols and objects/actions in the world is richly available to the learner, who is equipped with an exquisite correlation mapping device, an economic theory has to consider the resultant learning processes as one factor – certainly not the only factor – in semantic learning. In the minds of some



**Fig. 7.** Brain basis of communicative actions. a: Action sequence structures for the speech acts of naming (in blue) and requesting (in red). b: Results of EEG and fMRI experiments on the understanding of different speech acts performed with the same utterances. b: In the EEG response, stronger brain activation was seen in request (red line) compared with naming (blue line) contexts from 100 ms. Global field power, GFP, is plotted against time (in ms, adopted from Egorova et al., 2013). c: An fMRI study showed larger local brain activity (circles) for requests as compared with naming actions in left inferior frontal gyrus (IFG), bilateral premotor cortex (PMC), right and left posterior middle temporal gyrus and superior temporal sulcus (STS), and anterior intraparietal sulcus (IPS) (adopted from Egorova et al., 2016).



cognitivism, the term ‘association’ seems to be tied inseparably to behaviourist positions, so that its relevance is still neglected in some areas of cognitive theorizing.<sup>16</sup> However, what the cognitivist-vs-behaviourist debate showed is that language learning cannot be exhaustively explained in terms of a simple concept of association, as it is manifest in some work on classical and operant conditioning (see, for example, Skinner, 1957). That the mapping of information immanent to real world experiences and language use is irrelevant to semantic learning has not been claimed, and current language acquisition models even emphasise its power (see Sections 2.3, 2.7 Goldberg, 2006; MacWhinney, 2014; Tomasello, 2005). The removal of associative and dissociative aspects of correlation learning from language theory would leave a theoretical vacuum making it difficult to explain how the use of words such as “green” or “grasp” could be acquired in typical learners. The hypothesis spelt out here is that semantic brain processes emerge from brain structure and function, with learning of correlations being important. However, equally important are other established brain mechanisms and in particular the specifically human neuroanatomical connectivity structure, which determines that connector hub areas take on eminent conceptual and semantic roles.

### 3.3. Circuits for communicative actions and interactions

As mentioned in Section 2.8, the main function of language as a tool for communication somewhat contrasts with the relatively sparse neuroscience research addressing this function. Meanwhile, there is some interest in such ‘neuropragmatic’ research (Gisladottir et al., 2015; Hagoort and Levinson, 2014; Menenti et al., 2012; Noordzij et al., 2009). Recent highlights include the investigation of the brain basis of direct vs. indirect speech acts (Bašnáková et al., 2014) and the neurophysiological reflection of common ground (Rueschemeyer et al., 2015). A still largely open question is whether basic types of communicative actions are neurobiologically distinct in the sense that different brain activity patterns emerge when different actions are performed or understood.

In linguistic-pragmatic research, a general agreement is to distinguish major types of speech acts. Two important speech act categories are communicative actions similar to statements and requests-like acts; these are called ‘assertive’ vs. ‘directive’ speech acts, respectively (Searle, 1979). Taking on predictions developed in the APC framework (see Section 2.8), the next paragraphs will briefly explain some differences in brain activation consistently observed when human subjects understand naming and request actions, which can be seen as special cases of assertion and direction.

Methodologically, it is difficult to disentangle brain activity related to communicative function from that caused by linguistic structures. As words and sentences of different types may spark different brain networks (see Section 3.2), it is of crucial importance to balance utterance forms between communicative actions under investigation. An optimal way to achieve this is to use the same utterances and place them in different contexts characteristic of different types of communicative action, so that the action sequence structure and associated knowledge and prediction constellations are the critical distinguishing factors. For example, it is possible to use the same one-word holophrases (for example the word “water”) in different contexts to perform naming and request speech acts. By hypothesis, the naming

context activates the APC of the utterance, including its word form parts and its object-related referential semantic parts, thus activating inferior temporal ventral-visual stream circuits. In contrast, when using the same word to request an object, there are substantial expectations of future actions to be performed by the communication partner (Fig. 7a), which, in the APC framework, implies activation of circuits with frontocentral components. A possible brain manifestation of such action predictions is the activation of inferior and lateral prefrontal and precentral areas.

Stronger cortical activation in motor and premotor cortex was indeed found to utterances presented in request contexts than to the same utterances used for naming. A range of experiments performed with different imaging techniques (fMRI, MEG, EEG) confirmed this finding and provided complementary information about the time course of differential activity (Fig. 7b, c). Already ~100 ms after the word critical for understanding the speech acts could be recognized, brain activation distinguished between communicative actions performed with the same utterances (Egorova et al., 2013). In MEG, there was relatively stronger activation for requests and some of this additional activity was localized in frontocentral motor systems (Egorova et al., 2014). A recent fMRI study confirmed that the same utterances elicited stronger brain activity in sensorimotor, left inferior frontal and posterior temporal areas when they were used for requesting compared with when they served to name objects (Egorova et al., 2016). Other data indicate a complex interplay between sensorimotor, linguistic and theory-of-mind-related activations during speech act understanding (van Ackeren et al., 2016). Although research addressing this topic is still in its infancy, the brain basis of social-communicative actions provides a further important target of future experimental research and a challenge for explicit mechanistic brain models of language and communication (Bogels et al., 2014; Engel et al., 2013; Noveck and Reboul, 2008; van Ackeren et al., 2012; Van Berkum, 2009).

## 4. Summary and outlook

Language is sometimes thought of as an entirely novel invention in evolution. The alternative is that language emerged from action and perception mechanisms due to gradual evolution of the neuroanatomical substrate. Based on general principles of neuronal function and neuroanatomical features that, according to recent results, distinguish humans from their closest relatives, this article proposes a model of neural reuse of action and perception mechanisms for language, concepts and communication. Neurobiologically constrained computational networks are used to model and illustrate the emergence and function of action perception circuits for phonological and semantic processing (Sections 2.1–2.5). Action perception circuits provide a basis for higher cognitive mechanisms, including attention, memory, prediction and combination, and how they interact in linguistic and semantic/pragmatic processing (2.6–2.8). In Section 3, model predictions were evaluated in light of recent neuroscience data. At the semantic and conceptual levels, category-specific and category-general mechanisms were in focus, and in the domain of social-communicative interaction, new proposals and data were highlighted, which open rich perspectives for future research. Throughout, theories about brain language mechanisms were put in the context of current neuroscience research and discussions about mirror neurons, semantic grounding and abstract symbol systems.

Structural and functional principles and knowledge from neuroscience, in particular about cortical connectivity and correlation learning, are important for understanding and modelling language mechanisms at the level of neuronal circuits. This neurobiological information implies that brain mechanisms of

<sup>16</sup> A motivation for this position is that, within behaviourist theory, whole stimuli and responses are associated with each other, whereas in neural theories, neurons and circuits sometimes carrying abstract information are being interlinked.

phoneme, word and construction processing are neuronal assemblies distributed across modality-preferential areas related to action and perception and across associated connector hubs. For the emergent distributed but functionally intertwined neuron groups, the term ‘action perception circuits’ or APCs is proposed, because their formation is driven by correlation of sensorimotor information, although their function goes far beyond the processing of motor and sensory information. One mechanism contributing to their cognitive ‘reuse’ is strengthening of direct and indirect connections between APCs following their co-activation, thereby providing a brain basis for combination thus capturing syntactic, semantic and pragmatic information. Symbol-related APCs also link with object and action representations in the brain, which yields a mechanism for semantic grounding. Combinatorial semantic learning is relevant at the level of communicative actions, where sequence schemas provide the basis for action predictions specific to speech act types. Already at the level of speech sounds, the correlated motor and sensory activations during speaking yield correlation and, through (especially dorsal) frontotemporal connections, efficient action perception mapping, which explains the multimodal function of anterior and posterior perisylvian cortices along with sensorimotor parallelisms already in the first year of life.

In this perspective, language and communication mechanisms are built from information about actions and perceptions, and action and perception mechanisms are shaped and ‘reused’ for language, conceptual thought and communication. Therefore, the latter are intrinsically interlinked computationally with perception and action knowledge. Although body movements and input from the environment contribute to the formation of a network of APCs, this system is not restricted to mapping ‘the world’ on neurons: as discussed in detail, it crucially implies ‘creative’ aspects, for example in the mechanisms for combination, abstraction and action sequences. This view contrasts with classic cognitive models, although these approaches can be modified to take into account the important role of motor and sensory mechanisms in conceptual, semantic and pragmatic processing. However, an explanation why grounding in action and perception is relevant for building cognition requires a neurobiological, neuromechanistic theory taking into account both sensorimotor correlation structure and cortical connectivity. The APC framework provides such explanations at the level of detailed verbal description of causal chains and at the level of mathematically precise models constrained by human cortical anatomy and connectivity structure and neurophysiological function.

Key questions for future research address the brain basis of construction semantics and prediction in communication – to name only two most exciting topics. Could we, for example, bring to fruit some of the insights about semantic mappings in finding reliable measures of semantic predictions (see, for example, Grisoni et al., 2017)? Most importantly, however, it appears necessary to gain more direct evidence at the neuron circuit level about the neurobiological mechanisms of language and communication. Neuronal responses need to be tracked at increasingly fine-grained levels to indeed document and distinguish the postulated action perception circuits and their specific cortical distributions, over and above their effects on large-scale brain activity as is possible using current neuroimaging and neuropsychological methods.

## Acknowledgements

For the many valuable comments and suggestions I received while writing, revising and extending this paper, I thank a large number of friends and colleagues and, in particular, Michael Arbib, Veronique Boulenger, Felix Dreyer, Natalia Egorova, Martin Fischer,

Max Garagnani, Art Glenberg, Adele Goldberg, Luigi Grisoni, Olaf Hauk, Thomas Herbst, Valerie Keller, Bettina Mohr, Rachel Moseley, Horst Müller, Michael Pauen, Malte Schomers, Almut Schütz, Zubaida Shebani, Yury Shtyrov, Kristof Strijkers, Rosario Tomasello and Thomas Wennekers. This work was supported by the Freie Universität Berlin, the Engineering and Physical Sciences Research Council and Behavioural and Brain Sciences Research Council, UK (BABEL grant, EP/J004561/1), and the Deutsche Forschungsgemeinschaft, Germany (DFG, Pu 97/15-1, 16-1, 22-1, 23-1). The author declares no conflict of interest related to this work.

## References

- Abe, K., Watanabe, D., 2011. Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nat. Neurosci.* 14, 1067–1074.
- Abeles, M., 1991. *Corticocortical – Neural Circuits of the Cerebral Cortex*. Cambridge University Press, Cambridge.
- Alston, W.P., 1964. *Philosophy of Language*. Prentice-Hall, Englewood Cliffs, NJ.
- Anderson, M.L., 2010. Neural reuse: a fundamental organizational principle of the brain. *Behav. Brain Sci.* 33, 245–266 discussion 266–313.
- Anderson, M.L., 2015. Precipice after phrenology: neural reuse and the interactive brain. *Behav. Brain Sci.* 1–22.
- Arbib, M.A., Billard, A., Iacoboni, M., Oztop, E., 2000. Synthetic brain imaging: grasping: mirror neurons and imitation. *Neural Netw.* 13, 975–997.
- Arbib, M.A., 2016. Primates computation, and the path to language: reply to comments on towards a computational comparative neuroprimateology: framing the language-ready brain. *Phys Life Rev.* 16, 105–122.
- Arevalo, A.L., Baldo, J.V., Dronkers, N.F., 2012. What do brain lesions tell us about theories of embodied semantics and the human mirror neuron system? *Cortex* 48, 242–254.
- Aronov, D., Andalman, A.S., Fee, M.S., 2008. A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science* 320, 630–634.
- Arsenault, J.S., Buchsbaum, B.R., 2015a. Distributed neural representations of phonological features during speech perception. *J. Neurosci.* 35, 634–642.
- Arsenault, J.S., Buchsbaum, B.R., 2015b. No evidence of somatotopic place of articulation feature mapping in motor cortex during passive speech perception. *Psychon. Bull. Rev.* 1–10.
- Artola, A., Singer, W., 1993. Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *Trends Neurosci.* 16, 480–487.
- Aziz-Zadeh, L., Wilson, S.M., Rizzolatti, G., Iacoboni, M., 2006. Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Curr. Biol.* 16, 1818–1823.
- Bašnáková, J., Weber, K., Petersson, K.M., van Berkum, J., Hagoort, P., 2014. Beyond the language given: the neural correlates of inferring speaker meaning. *Cereb. Cortex* 24, 2572–2578.
- Bailey, C.J., Karhu, J., Ilmoniemi, R.J., 2001. Transcranial magnetic stimulation as a tool for cognitive studies. *Scand. J. Psychol.* 42, 297–305.
- Bak, T.H., O’Donovan, D.G., Xuereb, J.H., Boniface, S., Hodges, J.R., 2001. Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the Motor Neurone Disease-Dementia-Aphasia syndrome. *Brain* 124, 103–120.
- Bak, T.H., 2013. The neuroscience of action semantics in neurodegenerative brain diseases. *Curr. Opin. Neurol.* 26, 671–677.
- Baker, G.P., Hacker, P.M.S., 2009. *Wittgenstein: understanding and meaning*. Essays. Wiley-Blackwell, Chichester.
- Baker, C.L., 1979. Syntactic theory and the projection problem. *Ling. Inq.* 10, 533–581.
- Barlow, H., Levick, W.R., 1965. The mechanism of directionally selective units in rabbit’s retina. *J. Physiol.* 178, 477–504.
- Barrós-Loscertales, A., González, J., Pulvermüller, F., Ventura-Campos, N., Bustamante, J.C., Costumero, V., Parcet, M.A., Ávila, C., 2012. Reading salt activates gustatory brain regions: fMRI evidence for semantic grounding in a novel sensory modality. *Cereb. Cortex* 22, 2554–2563.
- Barsalou, L.W., Kyle Simmons, W., Barbey, A.K., Wilson, C.D., 2003. Grounding conceptual knowledge in modality-specific systems. *Trends Cogn. Sci.* 7, 84–91.
- Barsalou, L.W., 2008. Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645.
- Barsalou, L.W., 2013. Mirroring as pattern completion inferences within situated conceptualizations. *Cortex* 49, 2951–2953.
- Barsalou, L.W., 2016. On staying grounded and avoiding quixotic dead ends. *Psychon. Bull. Rev.* 23, 1122–1142.
- Bartoli, E., D’Ausilio, A., Berry, J., Badino, L., Bever, T., Fadiga, L., 2015. Listener-speaker perceived distance predicts the degree of motor contribution to speech perception. *Cereb. Cortex* 25, 281–288.
- Basso, A., Casati, G., Vignolo, L.A., 1977. Phonemic identification defect in aphasia. *Cortex* 13, 85–95.
- Basso, A., Lecours, A.R., Moraschini, S., Vanier, M., 1985. Anatomoclinical correlations of the aphasias as defined through computerized tomography: exceptions. *Brain Lang.* 26, 201–229.
- Bates, E., Wilson, S.M., Saygin, A.P., Dick, F., Sereno, M.I., Knight, R.T., Dronkers, N.F., 2003. Voxel-based lesion-symptom mapping. *Nat. Neurosci.* 6, 448–450.

- Bedny, M., Caramazza, A., 2011. Perception, action and word meanings in the human brain: the case from action verbs. *Ann. N. Y. Acad. Sci.* 1224, 81–95.
- Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., Saxe, R., 2008. Concepts are more than percepts: the case of action verbs. *J. Neurosci.* 28, 11347–11353.
- Bedny, M., Caramazza, A., Pascual-Leone, A., Saxe, R., 2012. Typical neural representations of action verbs develop without vision. *Cereb. Cortex* 22, 286–293.
- Beilock, S.L., Lyons, I.M., Mattarella-Micke, A., Nusbaum, H.C., Small, S.L., 2008. Sports experience changes the neural processing of action language. *Proc. Natl. Acad. Sci. U. S. A.* 105, 13269–13273.
- Berlinger, M., Crepaldi, D., Roberti, R., Scialfa, G., Luzzatti, C., Paulesu, E., 2008. Nouns and verbs in the brain: grammatical class and task specific effects as revealed by fMRI. *Cogn. Neuropsychol.* 1.
- Bi, G.Q., Poo, M.M., 1998. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing synaptic strength, and postsynaptic cell type. *J. Neurosci.* 18, 10464–10472.
- Bibbig, A., Wennekers, T., Palm, G., 1995. A neural network model of the cortico-hippocampal interplay and the representation of contexts. *Behav. Brain Res.* 66, 169–175.
- Bienenstock, E.L., Cooper, L.N., Munro, P.W., 1982. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J. Neurosci.* 2, 32–48.
- Bienenstock, E., 1994. A model of the neocortex. *Network: Comput. Neural Syst.* 6, 179–224.
- Bierwisch, M., 2008. Language and brain – facts problems, mysteries. *Eur. Rev.* 16, 467–481.
- Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. *Trends Cogn. Sci.* 15, 527–536.
- Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796.
- Bishop, D.V., Brown, B.B., Robson, J., 1990. The relationship between phoneme discrimination speech production, and language comprehension in cerebral-palsied individuals. *J. Speech Hear. Res.* 33, 210–219.
- Bleile, K.M., Stark, R.E., McGowan, J.S., 1993. Speech development in a child after decannulation: further evidence that babbling facilitates later speech development. *Clin. Linguist. Phon.* 7, 319–337.
- Blumstein, S.E., Baker, E., Goodglass, H., 1977. Phonological factors in auditory comprehension in aphasia. *Neuropsychologia* 15, 19–30.
- Bogels, S., Barr, D.J., Garrod, S., Kessler, K., 2014. Conversational interaction in the scanner: mentalizing during language processing as revealed by MEG. *Cereb. Cortex* 25, 3219–3234.
- Bogen, J.E., Bogen, G.M., 1976. Wernicke's region – where is it? *Ann. N. Y. Acad. Sci.* 280, 834–843.
- Bookheimer, S., 2002. Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu. Rev. Neurosci.* 25, 151–188.
- Bouchard, K.E., Mesgarani, N., Johnson, K., Chang, E.F., 2013. Functional organization of human sensorimotor cortex for speech articulation. *Nature* 495, 327–332.
- Boulenger, V., Roy, A.C., Paulignan, Y., Deprez, V., Jeannerod, M., Nazir, T.A., 2006. Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *J. Cogn. Neurosci.* 18, 1607–1615.
- Boulenger, V., Hauk, O., Pulvermüller, F., 2009. Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cereb. Cortex* 19, 1905–1914.
- Boulenger, V., Shtyrov, Y., Pulvermüller, F., 2012. When do you grasp the idea?: MEG evidence for instantaneous idiom understanding. *Neuroimage* 59, 3502–3513.
- Bowerman, M., 1988. The no negative evidence problem: how do children avoid constructing an overgeneral grammar? In: Hawkins, J.A. (Ed.), *Explaining Linguistic Universals*. Basil Blackwell, Oxford, pp. 73–101.
- Braitenberg, V., Pulvermüller, F., 1992. Entwurf einer neurologischen theorie der sprache. *Naturwissenschaften* 79, 103–117.
- Braitenberg, V., Schüz, A., 1998. *Cortex: Statistics and Geometry of Neuronal Connectivity*. Springer, Berlin.
- Braitenberg, V., 1978. Cell assemblies in the cerebral cortex. In: Heim, R., Palm, G. (Eds.), *Theoretical Approaches to Complex Systems*. (Lecture Notes in Biomathematics), vol. 21. Springer, Berlin, pp. 171–188.
- Breshears, J.D., Molinaro, A.M., Chang, E.F., 2015. A probabilistic map of the human ventral sensorimotor cortex using electrical stimulation. *J. Neurosurg.* 123, 340–349.
- Brooks, P.J., Tomasello, M., 1999. How children constrain their argument structure constructions. *Language* 75, 720–738.
- Brooks, P.J., Tomasello, M., Dodson, K., Lewis, L.B., 1999. Young children's overgeneralizations with fixed transitivity verbs. *Child Dev.* 70, 1325–1337.
- Bruce, C.J., Goldberg, M.E., 1985. Primate frontal eye fields. I. Single neurons discharging before saccades. *J. Neurophysiol.* 53, 603–635.
- Buzsáki, G., 2010. Neural syntax: cell assemblies synapse, and readers. *Neuron* 68, 362–285.
- Bruderer, A.G., Danielson, D.K., Kandhadai, P., Werker, J.F., 2015. Sensorimotor influences on speech perception in infancy. *Proc Natl Acad Sci U S A* 112, 13531–13536.
- Bird, H., Lambon-Ralph, M.A., Patterson, K., Hodges, J.R., 2000. The rise and fall of frequency and imageability: Noun and verb production in semantic dementia. *Brain Lang* 73, 17–49.
- Call, J., Tomasello, M., 2007. *The Gestural Communication of Apes and Monkeys*. Taylor & Francis Group/Lawrence Erlbaum Associates, New York, NY.
- Cangelosi, A., Greco, A., Harnad, S., 2002. Symbol grounding and the symbolic theft hypothesis. In: Cangelosi, A., Parisi, D. (Eds.), *Simulating the Evolution of Language*. Springer, London, pp. 3–20.
- Cangelosi, A., 2010. Grounding language in action and perception: from cognitive agents to humanoid robots. *Phys. Life Rev.* 7, 139–151.
- Canolty, R.T., Soltani, M., Dalal, S.S., Edwards, E., Dronkers, N.F., Nagarajan, S.S., Kirsch, H.E., Barbaro, N.M., Knight, R.T., 2007. Spatiotemporal dynamics of word processing in the human brain. *Front. Neurosci.* 1, 185–196.
- Caplan, D., Gow, D., Makris, N., 1995. Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. *Neurology* 45, 293–298.
- Caporale, N., Dan, Y., 2008. Spike timing-dependent plasticity: a hebbian learning rule. *Annu. Rev. Neurosci.* 31, 25–46.
- Cappa, S.F., Pulvermüller, F., 2012. Language and the motor system. *Cortex* 48, 785–787.
- Caramazza, A., Coltheart, M., 2006. Cognitive Neuropsychology twenty years on. *Cogn. Neuropsychol.* 23, 3–12.
- Caramazza, A., Mahon, B.Z., 2003. The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends Cogn. Sci.* 7, 354–361.
- Caramazza, A., Anzellotti, S., Strnad, L., Lingnau, A., 2014. Embodied cognition and mirror neurons: a critical assessment. *Annu. Rev. Neurosci.* 37, 1–15.
- Carlson, T.A., Simmons, R.A., Kriegeskorte, N., Slevc, L.R., 2013. The emergence of semantic meaning in the ventral temporal pathway. *J. Cogn. Neurosci.*
- Carota, F., Posada, A., Harquel, S., Delpuech, C., Bertrand, O., Sirigu, A., 2010. Neural dynamics of the intention to speak. *Cereb. Cortex* 20, 1891–1897.
- Carota, F., Moseley, R., Pulvermüller, F., 2012. Body-part-specific representations of semantic noun categories. *J. Cogn. Neurosci.* 24, 1492–1509.
- Carota, F., Kriegeskorte, N., Nili, H., Pulvermüller, F., 2017. Representational similarity mapping of distributional semantics in left inferior frontal middle temporal, and motor cortex. *Cereb. Cortex* 27, 294–309.
- Catani, M., Dell'Acqua, F., Budisavljevic, S., Howells, H., Thiebaut de Schotten, M., Froudust-Walsh, S., D'Anna, L., Thompson, A., Sandrone, S., Bullmore, E.T., Suckling, J., Baron-Cohen, S., Lombardo, M.V., Wheelwright, S.J., Chakrabarti, B., Lai, M.C., Ruigrok, A.N., Leemans, A., Ecker, C., Consortium, M.A., Craig, M.C., Murphy, D.G., 2016. Frontal networks in adults with autism spectrum disorder. *Brain* 139, 616–630.
- Chen, L., Ralph, M.A.L., Rogers, T.T., 2017. A unified model of human semantic knowledge and its disorders. *Nat. Hum. Behav.* 1, 0039.
- Chersi, F., Thill, S., Ziemke, T., Borghi, A.M., 2010. Sentence processing: linking language to motor chains. *Front. Neurobot.* 4.
- Chomsky, N., 1959. Review of verbal behavior by B.F. Skinner. *Language* 35, 26–58.
- Chomsky, N., 1965. *Aspects of the Theory of Syntax*. MIT Press, Cambridge.
- Christiansen, M.H., Chater, N., 1999. Toward a connectionist model of recursion in human linguistic performance. *Cognit. Sci.* 23, 157–205.
- Citron, F.M., Goldberg, A.E., 2014. Metaphorical sentences are more emotionally engaging than their literal counterparts. *J. Cogn. Neurosci.* 26, 2585–2595.
- Clark, E., 1987. The principle of contrast: a constraint on language acquisition. In: MacWhinney, B. (Ed.), *Mechanisms of Language Acquisition*. Erlbaum, Hillsdale, NJ.
- Cook, R., Bird, G., Catmur, C., Press, C., Heyes, C., 2014. Mirror neurons: from origin to function. *Behav. Brain Sci.* 37, 177–192.
- Cotelli, M., Borroni, B., Manenti, R., Alberici, A., Calabria, M., Agosti, C., Arevalo, A., Ginex, V., Ortelli, P., Binetti, G., Zanetti, O., Padovani, A., Cappa, S.F., 2006. Action and object naming in frontotemporal dementia progressive supranuclear palsy, and corticobasal degeneration. *Neuropsychology* 20, 558–565.
- Cotelli, M., Borroni, B., Manenti, R., Zanetti, M., Arevalo, A., Cappa, S.F., Padovani, A., 2007. Action and object naming in Parkinson's disease without dementia. *Eur. J. Neurol.* 14, 632–637.
- D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., Fadiga, L., 2009. The motor somatotopy of speech perception. *Curr. Biol.* 19, 381–385.
- Damasio, H., Grabowski, T.J., Tranel, D., Hichwa, R.D., Damasio, A.R., 1996. A neural basis for lexical retrieval. *Nature* 380, 499–505.
- Damasio, A.R., 1989. Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. *Cognition* 33, 25–62.
- de Boysson-Bardies, B., Vihman, M.M., 1991. Adaptation to language: evidence from babbling and first words in four languages. *Language* 67, 297–319.
- de Saussure, F., 1916. *Cours de linguistique generale*. Payot, Paris.
- de Zubicaray, G., Arciuli, J., McMahon, K., 2013. Putting an end to the motor cortex representations of action words. *J. Cogn. Neurosci.* 25, 1957–1974.
- Deco, G., Jirsa, V.K., McIntosh, A.R., 2013. Resting brains never rest: computational insights into potential cognitive architectures. *Trends Neurosci.* 36, 268–274.
- Dehaene-Lambertz, G., 1997. Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport* 8, 919–924.
- Desai, R.H., Conant, L.L., Binder, J.R., Park, H., Seidenberg, M.S., 2013. A piece of the action: modulation of sensory-motor regions by action idioms and metaphors. *Neuroimage* 83, 862–869.
- Devlin, J.T., Matthews, P.M., Rushworth, M.F., 2003. Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *J. Cogn. Neurosci.* 15, 71–84.
- Dick, A.S., Bernal, B., Tremblay, P., 2014. The language connectome: new pathways new concepts. *Neuroscientist* 20, 453–467.
- Dreyer, F.R., Frey, D., Arana, S., von Saldern, S., Picht, T., Vajkoczy, P., Pulvermüller, F., 2015. Is the motor system necessary for processing action and abstract emotion words? Evidence from focal brain lesions. *Front. Psychol.* 6, 1661.



- Dubois, J., Hertz-Pannier, L., Cachia, A., Mangin, J.F., Le Bihan, D., Dehaene-Lambertz, G., 2009. Structural asymmetries in the infant language and sensori-motor networks. *Cereb. Cortex* 19, 414–423.
- Dubois, J., Dehaene-Lambertz, G., Kulikova, S., Poupon, C., Hüppi, P., Hertz-Pannier, L., 2014. The early development of brain white matter: a review of imaging studies in fetuses newborns and infants. *Neuroscience* 276, 48–4871.
- Dum, R.P., Strick, P.L., 2002. Motor areas in the frontal lobe of the primate. *Physiol. Behav.* 77, 677–682.
- Dunbar, R.I., 1993. Coevolution of neocortical size: group size and language in humans. *Behav. Brain Sci.* 16, 681–694.
- Dunbar, R.I., 2016. Group size, vocal grooming and the origins of language. *Psychon. Bull. Rev.* 24, 209–212.
- Duncan, J., 2006. EPS mid-Career award 2004: brain mechanisms of attention. *Q. J. Exp. Psychol. (Colchester)* 59, 2–27.
- Duncan, J., 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn. Sci.* 14, 172–179.
- Dunn, M., Greenhill, S.J., Levinson, S.C., Gray, R.D., 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473, 79–82.
- Edwards, E., Nagarajan, S.S., Dalal, S.S., Canolty, R.T., Kirsch, H.E., Barbaro, N.M., Knight, R.T., 2010. Spatiotemporal imaging of cortical activation during verb generation and picture naming. *Neuroimage* 50, 291–301.
- Egorova, N., Shtyrov, Y., Pulvermüller, F., 2013. Early and parallel processing of pragmatic and semantic information in speech acts: neurophysiological evidence. *Front. Hum. Neurosci.* 7, 1–13.
- Egorova, N., Pulvermüller, F., Shtyrov, Y., 2014. Neural dynamics of speech act comprehension: an MEG study of Naming and Requesting. *Brain Topogr.* 27, 375–392.
- Egorova, N., Shtyrov, Y., Pulvermüller, F., 2016. Brain basis of communicative actions in language. *Neuroimage* 125, 857–867.
- Eimas, P.D., Siqueland, E.R., Jusczyk, P.W., Vigorito, J., 1971. Speech perception in infants. *Science* 171, 303–306.
- Elbert, T., Rockstroh, B., 1987. Threshold regulation – a key to the understanding of the combined dynamics of EEG and event related potentials. *J. Psychophysiol.* 4, 317–333.
- Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., Taub, E., 1995. Increased cortical representation of the fingers of the left hand in string players. *Science* 270, 305–307.
- Eldredge, D.H., Miller, J.D., 1971. Physiology of hearing. *Annu. Rev. Physiol.* 33, 281–310.
- Elman, J.L., Bates, L., Johnson, M., Karmiloff-Smith, A., Parisi, D., Plunkett, K., 1996. Rethinking Innateness. A Connectionist Perspective on Development. MIT Press, Cambridge, MA.
- Elman, J.L., 1990. Finding structure in time. *Cogn. Sci.* 14, 179–211.
- Elowson, M.A., Snowdon, C.T., Lazaro-Perea, C., 1998. Babbling and social context in infant monkeys: parallels to human infants. *Trends Cogn. Sci.* 2, 31–37.
- Engel, A.K., Maye, A., Kurthen, M., König, P., 2013. Where's the action? The pragmatic turn in cognitive science. *Trends Cogn. Sci.* 17, 202–209.
- Evans, S., Davis, M.H., 2015. Hierarchical organization of auditory and motor representations in speech perception: evidence from searchlight similarity analysis. *Cereb. Cortex* 25, 4772–4788.
- Fadiga, L., Craighero, L., Buccino, G., Rizzolatti, G., 2002. Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur. J. Neurosci.* 15, 399–402.
- Fadiga, L., Roy, A.C., Fazio, P., Craighero, L., 2007. From hand actions to speech: evidence and speculations. In: *Attention and Performance XXII: Sensori-Motor Foundations of Higher Cognition*. pp. 409–434. Eds. P. Haggard, Y. Rossetti, M. Kawato. OUP: Oxford.
- Fernandino, L., Binder, J.R., Desai, R.H., Pendl, S.L., Humphries, C.J., Gross, W.L., Conant, L.L., Seidenberg, M.S., 2015. Concept representation reflects multimodal abstraction: a framework for embodied semantics. *Cereb. Cortex* .
- Ferrari, P.F., Tramacere, A., Simpson, E.A., Iriki, A., 2013. Mirror neurons through the lens of epigenetics. *Trends Cogn. Sci.* 17, 450–457.
- Fillmore, C.J., 1967. The case for case. In: Bach, E., Harms, R. (Eds.), *Proceedings of the Texas Symposium, on Language Universals*, April 13–15. Holt, Rinehart and Winston, New York, pp. 1967.
- Fischer, M.H., Zwaan, R.A., 2008. Embodied language: a review of the role of the motor system in language comprehension. *Q. J. Exp. Psychol. (Colchester)* 61, 825–850.
- Fletcher, P.T., Whitaker, R.T., Tao, R., DuBray, M.B., Froehlich, A., Ravichandran, C., Alexander, A.L., Bigler, E.D., Lange, N., Lainhart, J.E., 2010. Microstructural connectivity of the arcuate fasciculus in adolescents with high-functioning autism. *Neuroimage* 51, 1117–1125.
- Fodor, J.A., Pylyshyn, Z.W., 1988. Connectionism and cognitive architecture: a critical analysis. *Cognition* 28, 3–71.
- Fodor, J.A., 1983. *The Modularity of Mind*. MIT Press, Cambridge, MA.
- Fornito, A., Zalesky, A., Breakspear, M., 2015. The connectomics of brain disorders. *Nat. Rev. Neurosci.* 16, 150–172.
- Fowler, C.A., 1986. An event approach to the study of speech perception from a direct-realist perspective. *J. Phon.* 14, 3–28.
- Freud, S., 1891. *Zur Auffassung der Aphasien*. Franz Deuticke, Leipzig, Wien.
- Friston, K., 2010. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138.
- Fritz, G., 2013. *Dynamische Texttheorie*. Gießener Elektronische Bibliothek, Gießen.
- Fuster, J.M., Bressler, S.L., 2012. Cognit activation: a mechanism enabling temporal integration in working memory. *Trends Cogn. Sci.* 16, 207–218.
- Fuster, J.M., Bressler, S.L., 2015. Past makes future: role of pFC in prediction. *J. Cogn. Neurosci.* 27, 639–654.
- Fuster, J.M., 1995. Memory in the cerebral cortex. *An Empirical Approach to Neural Networks in the Human and Nonhuman Primate*. MIT Press, Cambridge, MA.
- Fuster, J.M., 2003. *Cortex and Mind: Unifying Cognition*. Oxford University Press, Oxford.
- Fuster, J.M., 2009. Cortex and memory: emergence of a new paradigm. *J. Cogn. Neurosci.* 21, 2047–2072.
- Galantucci, B., Fowler, C.A., Turvey, M.T., 2006. The motor theory of speech perception reviewed. *Psychon. Bull. Rev.* 13, 361–377.
- Gallese, V., Lakoff, G., 2005. The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 22, 455–479.
- Garagnani, M., Pulvermüller, F., 2013. Neuronal correlates of decisions to speak and act: spontaneous emergence and dynamic topographies in a computational model of frontal and temporal areas. *Brain Lang.* 127, 75–85.
- Garagnani, M., Pulvermüller, F., 2016. Conceptual grounding of language in action and perception: a neurocomputational model of the emergence of category specificity and semantic hubs. *Eur. J. Neurosci.* 43, 721–737.
- Garagnani, M., Wennekers, T., Pulvermüller, F., 2007. A neuronal model of the language cortex. *Neurocomputing* 70, 1914–1919.
- Garagnani, M., Wennekers, T., Pulvermüller, F., 2008. A neuroanatomically-grounded Hebbian learning model of attention-language interactions in the human brain. *Eur. J. Neurosci.* 27, 492–513.
- Garagnani, M., Shtyrov, Y., Pulvermüller, F., 2009a. Effects of attention on what is known and what is not: MEG evidence for functionally discrete memory circuits. *Front. Hum. Neurosci.* 3, 1–8.
- Garagnani, M., Wennekers, T., Pulvermüller, F., 2009b. Recruitment and consolidation of cell assemblies for words by way of Hebbian learning and competition in a multi-layer neural network. *Cogn. Comput.* 1, 160–176.
- Garagnani, M., Lucchese, G., Tomasello, R., Wennekers, T., Pulvermüller, F., 2017. A spiking neurocomputational model of high-frequency oscillatory brain responses to words and pseudowords. *Front. Comput. Neurosci.* 10, 145.
- Gentner, T.Q., Fenn, K.M., Margoliash, D., Nusbaum, H.C., 2006. Recursive syntactic pattern learning by songbirds. *Nature* 440, 1204–1207.
- Gerfo, E.L., Oliveri, M., Torriero, S., Salerno, S., Koch, G., Caltagirone, C., 2008. The influence of rTMS over prefrontal and motor areas in a morphological task: grammatical vs. semantic effects. *Neuropsychologia* 46, 764–770.
- Gervain, J., Macagno, F., Cogoi, S., Pena, M., Mehler, J., 2008. The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U. S. A.* 105, 14222–14227.
- Geschwind, N., 1970. The organization of language and the brain. *Science* 170, 940–944.
- Giese, M.A., Rizzolatti, G., 2015. Neural and computational mechanisms of action processing: interaction between visual and motor representations. *Neuron* 88, 167–180.
- Gisladottir, R.S., Chwilla, D.J., Levinson, S.C., 2015. Conversation electrified: ERP correlates of speech act recognition in underspecified utterances. *PLoS One* 10, e0120068.
- Glenberg, A.M., Gallese, V., 2012. Action-based language: a theory of language acquisition comprehension, and production. *Cortex* 48, 905–922.
- Glenberg, A.M., Sato, M., Cattaneo, L., 2008. Use-induced motor plasticity affects the processing of abstract and concrete language. *Curr. Biol.* 18, R290–291.
- Goldberg, A.E., 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press, Chicago.
- Goldberg, A.E., 2006. *Constructions at Work: The Nature of Generalisation in Language*. Oxford University Press, Oxford.
- Goldberg A.E., in press. Explain me this: Creativity, competition and the partial productivity of constructions. Princeton University Press: Princeton.
- Goldman-Rakic, P.S., 1995. Cellular basis of working memory. *Neuron* 14, 477–485.
- Gonzalez, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuan, A., Belloch, V., Avila, C., 2006. Reading cinnamon activates olfactory brain regions. *Neuroimage* 32, 906–912.
- Goodale, M.A., Milner, A.D., 1992. Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25.
- Graziano, M.S., Aflalo, T.N., 2007. Mapping behavioral repertoire onto the cortex. *Neuron* 56, 239–251.
- Griffin, D.M., Hoffman, D.S., Strick, P.L., 2015. Corticomotoneuronal cells are functionally tuned. *Science* 350, 667–670.
- Griffiths, T.D., Warren, J.D., 2002. The planum temporale as a computational hub. *Trends Neurosci.* 25, 348–353.
- Grisoni, L., Dreyer, F.R., Pulvermüller, F., 2016. Somatotopic semantic priming and prediction in the motor system. *Cereb. Cortex* 26, 2353–2366.
- Guenther, F.H., Gajda, M.N., 1996. The perceptual magnet effect as an emergent property of neural map formation. *J. Acoust. Soc. Am.* 100, 1111–1121.
- Guenther, F.H., Vladusich, T., 2012. A neural theory of speech acquisition and production. *J. Neurolinguist.* 25, 408–422.
- Grisoni, L., Miller, T.M., Pulvermüller, F., 2017. Neural Correlates of Semantic Prediction and Resolution in Sentence Processing. *J. Neurosci.* 37, 4848–4858.
- Gustafsson, B., Wigström, H., Abraham, W.C., Huang, Y.Y., 1987. Long term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials. *J. Neurosci.* 7, 774–780.
- Gleichgerricht, E., Fridriksson, J., Rorden, C., Nesland, T., Desai, R., Bonilha, L., 2016. eparate neural systems support representations for actions and objects during narrative speech in post-stroke aphasia. *Neuroimage Clin* 10, 140–145.

- Hagoort, P., Levinson, S.C., 2014. *Neuropragmatics. The Cognitive Neurosciences*. MIT Press, Boston, MA, pp. 667–674.
- Hare, M., Elman, J.L., Daugherty, K.G., 1995. Default generalisation in connectionist networks. *Lang. Cogn. Processes* 10, 601–630.
- Harley, T.A., 2008. *The Psychology of Language*. Psychology Press, Taylor & Francis Group, Hove, UK.
- Harnad, S., 1990. The symbol grounding problem. *Physica D* 42, 335–346.
- Harris, K.D., 2005. Neural signatures of cell assembly organization. *Nat. Rev. Neurosci.* 6, 399–407.
- Hartwigsen, G., Weigel, A., Schuschan, P., Siebner, H.R., Weise, D., Classen, J., Saur, D., 2015. Dissociating parieto-frontal networks for phonological and semantic word decisions: a condition-and-perturb TMS study. *Cereb. Cortex* 26, 2590–2601.
- Haug, H., 1987. Brain sizes, surfaces, and neuronal sizes of the cortex cerebri: a stereological investigation of man and his variability and a comparison with some mammals (primates, whales, marsupials, insectivores, and one elephant). *Am. J. Anat.* 180, 126–142.
- Hauk, O., Pulvermüller, F., 2004. Neurophysiological distinction of action words in the fronto-central cortex. *Hum. Brain Mapp.* 21, 191–201.
- Hauk, O., Tschentscher, N., 2013. The body of evidence: what can neuroscience tell us about embodied semantics? *Front. Psychol.* 4, 50.
- Hauk, O., Johnsrude, I., Pulvermüller, F., 2004. Somatotopic representation of action words in the motor and premotor cortex. *Neuron* 41, 301–307.
- Hauk, O., Davis, M.H., Ford, M., Pulvermüller, F., Marslen-Wilson, W.D., 2006a. The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage* 30, 1383–1400.
- Hauk, O., Shtyrov, Y., Pulvermüller, F., 2006b. The sound of actions as reflected by mismatch negativity: rapid activation of cortical sensory-motor networks by sounds associated with finger and tongue movements. *Eur. J. Neurosci.* 23, 811–821.
- Hauser, M.D., Newport, E.L., Aslin, R.N., 2001. Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* 78, B53–64.
- Hayon, G., Abeles, M., Lehmann, D., 2005. A model for representing the dynamics of a system of synfire chains. *J. Comput. Neurosci.* 18, 41–53.
- Hebb, D.O., 1949. *The Organization of Behavior. A Neuropsychological Theory*. John Wiley, New York.
- Herculano-Houzel, S., 2009. The human brain in numbers: a linearly scaled-up primate brain. *Front. Hum. Neurosci.* 3, 31.
- Heyes, C., 2010. Where do mirror neurons come from? *Neurosci. Biobehav. Rev.* 34, 575–583.
- Hickok, G., 2009. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* 21, 1229–1243.
- Hickok, G., 2010. The role of mirror neurons in speech perception and action word semantics. *Lang. Cogn. Processes* 25, 749–776.
- Hickok, G., 2012. Computational neuroanatomy of speech production. *Nat. Rev. Neurosci.* 13, 135–145.
- Hickok, G., 2014. *The Myth of Mirror Neurons: the Real Neuroscience of Communication and Cognition*. WW Norton & Company, New York.
- Hickok, G., 2015. The motor system's contribution to perception and understanding actions: clarifying mirror neuron myths and misunderstandings. *Lang. Cogn.* 7, 476–484.
- Hillis, A.E., Wityk, R.J., Tuffiash, E., Beauchamp, N.J., Jacobs, M.A., Barker, P.B., Selnes, O.A., 2001. Hypoperfusion of Wernicke's area predicts severity of semantic deficit in acute stroke. *Ann. Neurol.* 50, 561–566.
- Hoenig, K., Sim, E.J., Bochev, V., Herrnberger, B., Kiefer, M., 2008. Conceptual flexibility in the human brain: dynamic recruitment of semantic maps from visual motor, and motion-related areas. *J. Cogn. Neurosci.* 20, 1799–1814.
- Hoffmann, T., Trousdale, G. (Eds.), 2013. *Oxford Handbook of Construction Grammar*. Oxford University Press, Oxford.
- Hommel, B., Musseler, J., Aschersleben, G., Prinz, W., 2001. The Theory of Event Coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–878 discussion 878–937.
- Hubel, D., 1995. *Eye, Brain, and Vision*. Scientific American Library, New York.
- Huth, A.G., de Heer, W.A., Griffiths, T.L., Theunissen, F.E., Gallant, J.L., 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532, 453–458.
- Huyck, C.R., Passmore, P.J., 2013. A review of cell assemblies. *Biol. Cybern.* 107, 263–288.
- Ito, T., Tiede, M., Ostry, D.J., 2009. Somatosensory function in speech perception. *Proc. Natl. Acad. Sci. U. S. A.* 106, 1245–1248.
- Indefrey, P., Levelt, W.J., 2004. The spatial and temporal signatures of word production components. *Cognition* 92, 101–144.
- Jackendoff, R., 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford University Press, Oxford, UK.
- Jacob, P., Jeannerod, M., 2005. The motor theory of social cognition: a critique. *Trends Cogn. Sci.* 9, 21–25.
- Jacobs, B., Batal, H.A., Lynch, B., Ojemann, G., Ojemann, L.M., Scheibel, A.B., 1993. Quantitative dendritic and spine analyses of speech cortices: a case study. *Brain Lang.* 44, 239–253.
- Jeannerod, M., Arbib, M.A., Rizzolatti, G., Sakata, H., 1995. Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci.* 18, 314–320.
- Jeannerod, M., 2006. *Motor Cognition: What Actions Tell to the Self*. Oxford University Press, Oxford.
- Jerison, H., 1975. *Evolution of the Brain and Intelligence*. University of Chicago Press, Chicago.
- Joseph, R.M., Fricker, Z., Fenoglio, A., Lindgren, K.A., Knaus, T.A., Tager-Flusberg, H., 2013. Structural asymmetries of language-related gray and white matter and their relationship to language function in young children with ASD. *Brain Imaging Behav.*
- Kakei, S., Hoffman, D.S., Strick, P.L., 1999. Muscle and movement representations in the primary motor cortex. *Science* 285, 2136–2139.
- Kaplan, S., Sonntag, M., Chown, E., 1991. Tracing recurrent activity in cognitive elements (TRACE): a model of temporal dynamics in a cell assembly. *Connect. Sci.* 3, 179–206.
- Kemmerer, D., Gonzalez-Castillo, J., 2010. The Two-Level Theory of verb meaning: an approach to integrating the semantics of action with the mirror neuron system. *Brain Lang.* 112, 54–76.
- Kemmerer, D., Castillo, J.G., Talavage, T., Patterson, S., Wiley, C., 2008. Neuroanatomical distribution of five semantic components of verbs: evidence from fMRI. *Brain Lang.* 107, 16–43.
- Kemmerer, D., Rudrauf, D., Manzel, K., Tranel, D., 2012. Behavioural patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of action. *Cortex* 48, 826–848.
- Kemmerer, D., 2015a. Are the motor features of verb meanings represented in the precentral motor cortices? Yes, but within the context of a flexible, multilevel architecture for conceptual knowledge. *Psychon. Bull. Rev.* 22, 1068–1075.
- Kemmerer, D., 2015b. Does the motor system contribute to the perception and understanding of actions? Reflections on Gregory Hickok's The myth of mirror neurons: the real neuroscience of communication and cognition. *Lang. Cogn.* 7, 1068–1075.
- Keysers, C., Gazzola, V., 2014. Hebbian learning and predictive mirror neurons for actions, sensations and emotions. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 369, 20130175.
- Keysers, C., Perrett, D.I., 2002. Visual masking and RSVP reveal neural competition. *Trends Cogn. Sci.* 6, 120–125.
- Keysers, C., Perrett, D.I., 2004. Demystifying social cognition: a Hebbian perspective. *Trends Cogn. Sci.* 8, 501–507.
- Kiefer, M., Pulvermüller, F., 2012. Conceptual representations in mind and brain: theoretical developments: current evidence and future directions. *Cortex* 48, 805–825.
- Kiefer, M., Sim, E.J., Herrnberger, B., Grothe, J., Hoenig, K., 2008. The sound of concepts: four markers for a link between auditory and conceptual brain systems. *J. Neurosci.* 28, 12224–12230.
- Kiefer, M., Adams, S.C., Zovko, M., 2012. Attentional sensitization of unconscious visual processing: top-down influences on masked priming. *Adv. Cogn. Psychol.* 8, 50–61.
- Kiefer, M., Liegel, N., Zovko, M., Wentura, D., 2016. Mechanisms of masked evaluative priming: task sets modulate behavioral and electrophysiological priming for picture and words differentially. *Soc. Cogn. Affect. Neurosci.*
- Kilner, J.M., Friston, K.J., Frith, C.D., 2007. Predictive coding: an account of the mirror neuron system. *Cogn. Process.* 8, 159–166.
- Kilner, J., Hommel, B., Bar, M., Barsalou, L.W., Friston, K.J., Jost, J., Maye, A., Metzinger, T., Pulvermüller, F., Sánchez-Fibla, M., Tsotsos, J., Vigliocco, G., 2016. A little less cogitation, a little more action please: action-Oriented models of cognitive processing. In: Engel, A.K., Friston, K.J., Kragic, D. (Eds.), *The Pragmatic Turn: Toward Action-oriented Views in Cognitive Science*. MIT Press, Boston, MA, pp. 159–174.
- Knoblauch, A., Palm, G., 2001. Pattern separation and synchronization in spiking associative memories and visual areas. *Neural Netw.* 14, 763–780.
- Knoblauch, A., Palm, G., 2002. Scene segmentation by spike synchronization in reciprocally connected visual areas. I. Local effects of cortical feedback. *Biol. Cybern.* 87, 151–167.
- Knoblauch, A., Pulvermüller, F., 2005. Sequence detector networks and associative learning of grammatical categories. In: Wermter, S.G., Palm, M. Elshaw (Eds.), *Biomimetic Neural Learning for Intelligent Robots*. Springer, Berlin, pp. 31–53.
- Kuhl, P.K., Miller, J.D., 1978. Speech perception by the chinchilla: identification function for synthetic VOT stimuli. *J. Acoust. Soc. Am.* 63, 905–917.
- Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., Lindblom, B., 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608.
- Kuhl, P.K., 2004. Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843.
- Kuipers, J.R., van Koningsbruggen, M., Thierry, G., 2013. Semantic priming in the motor cortex: evidence from combined repetitive transcranial magnetic stimulation and event-related potential. *Neuroreport* 24, 646–651.
- Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647.
- Lahav, A., Saltzman, E., Schlaug, G., 2007. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J. Neurosci.* 27, 308–314.
- Lakoff, G., 1987. *Women, Fire, and Dangerous Things. What Categories Reveal About the Mind*. University of Chicago Press, Chicago.
- Landauer, T.K., Dumais, S.T., 1997. A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition induction, and representation of knowledge. *Psychol. Rev.* 104, 211–240.
- Lansner, A., 2009. Associative memory models: from the cell-assembly theory to biophysically detailed cortex simulations. *Trends Neurosci.* 32, 178–186.
- Lashley, K.S., 1951. The problem of serial order in behavior. In: Jeffress, L.A. (Ed.), *Cerebral Mechanisms in Behavior. The Hixon Symposium*. John Wiley, New York, pp. 112–136.

- Lebois, L.A., Wilson-Mendenhall, C.D., Barsalou, L.W., 2015. Are automatic conceptual cores the gold standard of semantic processing? The context-dependence of spatial meaning in grounded congruency effects. *Cogn. Sci.* 39, 1764–1801.
- Lenneberg, E.H., 1967. *Biological Foundations of Language*. Wiley, New York.
- Leshinskaya, A., Caramazza, A., 2016. For a cognitive neuroscience of concepts: moving beyond the grounding issue. *Psychon. Bull. Rev.* 23, 991–1001.
- Liberman, A.M., Cooper, F.S., Shankweiler, D.P., Studdert-Kennedy, M., 1967. Perception of the speech code. *Psychol. Rev.* 74, 431–461.
- Lichtheim, L., 1885. On aphasia. *Brain* 7, 433–484.
- Liljestrom, M., Tarkiainen, A., Parviainen, T., Kujala, J., Numminen, J., Hiltunen, J., Laine, M., Salmelin, R., 2008. Perceiving and naming actions and objects. *Neuroimage* 41, 1132–1141.
- Liuzzi, G., Freundlieb, N., Ridder, V., Hoppe, J., Heise, K., Zimerman, M., Dobel, C., Enriquez-Geppert, S., Gerloff, C., Zwitserlood, P., Hummel, F.C., 2010. The involvement of the left motor cortex in learning of a novel action word lexicon. *Curr. Biol.* 20, 1745–1751.
- Locatelli, M., Gatti, R., Tettamanti, M., 2012. Training of manual actions improves language understanding of semantically related action sentences. *Front. Psychol.* 3, 547.
- Locke, J.L., Pearson, D.M., 1990. Linguistic significance of babbling: evidence from a tracheostomized infant. *J. Child Lang.* 17, 1–16.
- Locke, J.L., 1993. *The Child's Path to Spoken Language*. Harvard University Press, Cambridge, MA.
- Lopez-Barroso, D., Catani, M., Ripolles, P., Dell'Acqua, F., Rodriguez-Fornells, A., de Diego-Balaguer, R., 2013. Word learning is mediated by the left arcuate fasciculus. *Proc. Natl. Acad. Sci. U. S. A.* 110, 13168–13173.
- Luria, A.R., 1980. *Higher Cortical Functions in Man*. Basic Books, New York.
- Lutzenberger, W., Pulvermüller, F., Birbaumer, N., 1994. Words and pseudowords elicit distinct patterns of 30-Hz activity in humans. *Neurosci. Lett.* 176, 115–118.
- Lyons, I.M., Mattarella-Micke, A., Cieslak, M., Nusbaum, H.C., Small, S.L., Beilock, S.L., 2010. The role of personal experience in the neural processing of action-related language. *Brain Lang.* 112, 214–222.
- Möttönen, R., Watkins, K.E., 2009. Motor representations of articulators contribute to categorical perception of speech sounds. *J. Neurosci.* 29, 9819–9825.
- Möttönen, R., Dutton, R., Watkins, K.E., 2013. Auditory-motor processing of speech sounds. *Cereb. Cortex* 23, 1190–1197.
- MacGregor, L.J., Pulvermüller, F., van Casteren, M., Shtyrov, Y., 2012. Ultra-rapid access to words in the brain. *Nat. Commun.* 3, 711.
- MacWhinney, B., 2014. What we have learned. *J. Child Lang.* 41 (Suppl. 1), 124–131.
- Mahon, B.Z., Caramazza, A., 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70.
- Mahon, B.Z., Hickok, G., 2016. Arguments about the nature of concepts: symbols embodiment, and beyond. *Psychon. Bull. Rev.* 23, 941–958.
- Mahon, B.Z., 2015. What is embodied about cognition? *Lang. Cogn. Neurosci.* 30, 420–429.
- Marie, P., 1906. Révision de la question de l'aphasie: la troisième circonvolution frontale gauche ne joue aucun rôle spécial dans la fonction du langage. *Semaine Médicale (Paris)* 26, 241–247.
- Markram, H., Lubke, J., Frotscher, M., Sakmann, B., 1997. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275, 213–215.
- Marshall, P.J., Meltzoff, A.N., 2014. Neural mirroring mechanisms and imitation in human infants. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 20130620.
- Marslen-Wilson, W., Warren, P., 1994. Levels of perceptual representation and process in lexical access: words, phonemes, and features. *Psychol. Rev.* 101, 653–675.
- Martin, A., Wiggs, C.L., Ungerleider, L.G., Haxby, J.V., 1996. Neural correlates of category-specific knowledge. *Nature* 379, 649–652.
- Martin, A., 2007. The representation of object concepts in the brain. *Annu. Rev. Psychol.* 58, 25–45.
- Matsumoto, R., Nair, D.R., LaPresto, E., Najm, I., Bingaman, W., Shibusaki, H., Luders, H.O., 2004. Functional connectivity in the human language system: a cortico-cortical evoked potential study. *Brain* 127, 2316–2330.
- Matsumoto, R., Nair, D.R., LaPresto, E., Bingaman, W., Shibusaki, H., Luders, H.O., 2007. Functional connectivity in human cortical motor system: a cortico-cortical evoked potential study. *Brain* 130, 181–197.
- Matsuzaka, Y., Picard, N., Strick, P.L., 2007. Skill representation in the primary motor cortex after long-term practice. *J. Neurophysiol.* 97, 1819–1832.
- Meltzoff, A.N., Moore, M.K., 1983. Newborn infants imitate adult facial gestures. *Child Dev.* 54, 702–709.
- Menenti, L., Pickering, M.J., Garrod, S.C., 2012. Toward a neural basis of interactive alignment in conversation. *Front. Hum. Neurosci.* 6, 185.
- Mesgarani, N., Cheung, C., Johnson, K., Chang, E.F., 2014. Phonetic feature encoding in human superior temporal gyrus. *Science* 343, 1006–1010.
- Meteyard, L., Cuadrado, S.R., Bahrami, B., Vigliocco, G., 2012. Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex* 48, 788–804.
- Miller, R., 1996. *Axonal Conduction Times and Human Cerebral Laterality*. A Psychobiological Theory. Harwood Academic Publishers, Amsterdam.
- Milner, P.M., 1996. Neural representation: some old problems revisited. *J. Cogn. Neurosci.* 8, 69–77.
- Miozzo, M., Pulvermüller, F., Hauk, O., 2015. Early parallel activation of semantics and phonology in picture naming: evidence from a multiple linear regression MEG study. *Cereb. Cortex* 25, 3343–3355.
- Mitchell, T.M., Shinkareva, S.V., Carlson, A., Chang, K.M., Malave, V.L., Mason, R.A., Just, M.A., 2008. Predicting human brain activity associated with the meanings of nouns. *Science* 320, 1191–1195.
- Moineau, S., Dronkers, N.F., Bates, E., 2005. Exploring the processing continuum of single-word comprehension in aphasia. *J. Speech Lang. Hear. Res.* 48, 884–896.
- Moscato Del Prado Martin, F., Hauk, O., Pulvermüller, F., 2006. Category specificity in the processing of color-related and form-related words: an ERP study. *Neuroimage* 29, 29–37.
- Moseley, R., Carota, F., Hauk, O., Mohr, B., Pulvermüller, F., 2012. A role for the motor system in binding abstract emotional meaning. *Cereb. Cortex* 22, 1634–1647.
- Moseley, R.L., Pulvermüller, F., Shtyrov, Y., 2013. Sensorimotor semantics on the spot: brain activity dissociates between conceptual categories within 150 ms. *Sci. Rep.* 3, 1928.
- Moseley, R.L., Correia, M.M., Baron-Cohen, S., Shtyrov, Y.Y., Pulvermüller, F., Mohr, B., 2016. Reduced volume of the arcuate fasciculus in adults with high-functioning autism spectrum conditions. *Front. Hum. Neurosci.* 10, 214.
- Mollo, G., Pulvermüller, F., Hauk, O., 2016. Movement priming of EEG/MEG brain responses for action-words characterizes the link between language and action. *Cortex* 74, 262–276.
- Mukamel, R., Ekstrom, A.D., Kaplan, J., Iacoboni, M., Fried, I., 2010. Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* 20, 750–756.
- Murakami, T., Kell, C.A., Restle, J., Ugawa, Y., Ziemann, U., 2015. Left dorsal speech stream components and their contribution to phonological processing. *J. Neurosci.* 35, 1411–1422.
- Näätänen, R., Lehtikoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, A., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik, J., Sinkkonen, J., Alho, K., 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434.
- Näätänen, R., 1992. *Attention and Brain Function*. Erlbaum, Hillsdale, NJ.
- Näätänen, R., 2001. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* 38, 1–21.
- Nadeau, S.E., Crosson, B., 1997. Subcortical aphasia. *Brain Lang.* 58, 355–402 discussion 418–323.
- Nagy, E., Compagne, H., Orvos, H., Pal, A., Molnar, P., Janszky, I., Loveland, K.A., Bardos, G., 2005. Index finger movement imitation by human neonates: motivation learning, and left-hand preference. *Pediatr. Res.* 58, 749–753.
- Nagy, E., Pilling, K., Orvos, H., Molnar, P., 2013. Imitation of tongue protrusion in human neonates: specificity of the response in a large sample. *Dev. Psychol.* 49, 1628–1638.
- Neininger, B., Pulvermüller, F., 2001. The right hemisphere's role in action word processing: a double case study. *Neurocase* 7, 303–317.
- Noordzij, M.L., Newman-Norlund, S.E., de Ruiter, J.P., Hagoort, P., Levinson, S.C., Toni, I., 2009. Brain mechanisms underlying human communication. *Front. Hum. Neurosci.* 3, 14.
- Noveck, I.A., Reboul, A., 2008. Experimental pragmatics: a Gricean turn in the study of language. *Trends Cogn. Sci.* 12, 425–431.
- Nasir, S.M., Ostry, D.J., 2009. Auditory plasticity and speech motor learning. *Proc Natl Acad Sci U S A* 106, 20470–20475.
- O'Reilly, R.C., 1998. Six principles for biologically based computational models of cortical cognition. *Trends Cogn. Sci.* 2, 455–562.
- Obleser, J., Eisner, F., 2009. Pre-lexical abstraction of speech in the auditory cortex. *Trends Cogn. Sci.* 13, 14–19.
- Obleser, J., Boecker, H., Drzezga, A., Haslinger, B., Hennenlotter, A., Roetinger, M., Eulitz, C., Rauschecker, J.P., 2006. Vowel sound extraction in anterior superior temporal cortex. *Hum. Brain Mapp.* 27, 562–571.
- Ostry, D.J., Gribble, P.L., 2016. Sensory Plasticity in Human Motor Learning. *Trends Neurosci.* 39, 114–123.
- Palm, G., Sommer, F.T., 1995. Associative data storage and retrieval in neural networks. In: Domany, E., van Hemmen, J.L., Schulten, K. (Eds.), *Models of Neural Networks III*. Springer Verlag, New York, pp. 79–118.
- Palm, G., Knoblauch, A., Hauser, F., Schuz, A., 2014. Cell assemblies in the cerebral cortex. *Biol. Cybern.* 108, 559–572.
- Pinker, S., 1989. *Learnability and cognition: the acquisition of argument structure*. MIT Press, Cambridge, MA.
- Palm, G., 1982. *Neural Assemblies*. Springer, Berlin.
- Papeo, L., Corradi-Dell'Acqua, C., Rumiat, R.I., 2011. She is not like I: the tie between language and action is in our imagination. *J. Cogn. Neurosci.* 23, 3939–3948.
- Papeo, L., Pascual-Leone, A., Caramazza, A., 2013. Disrupting the brain to validate hypotheses on the neurobiology of language. *Front. Hum. Neurosci.* 7, 148.
- Patten, E., Belardi, K., Baranek, G.T., Watson, L.R., Labban, J.D., Oller, D.K., 2014. Vocal patterns in infants with autism spectrum disorder: canonical babbling status and vocalization frequency. *J. Autism Dev. Disord.* 44, 2413–2428.
- Patterson, K., Nestor, P.J., Rogers, T.T., 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* 8, 976–987.
- Pelvig, D.P., Pakkenberg, H., Stark, A.K., Pakkenberg, B., 2008. Neocortical glial cell numbers in human brains. *Neurobiol. Aging* 29, 1754–1762.
- Penfield, W., Rasmussen, T., 1950. *The Cerebral Cortex of Man*. Macmillan, New York.
- Perin, R., Berger, T.K., Markram, H., 2011. A synaptic organizing principle for cortical neuronal groups. *Proc. Natl. Acad. Sci.* 108, 5419–5424.
- Perlovsky, L., 2011. Language and cognition interaction neural mechanisms. *Comput Intell Neurosci* 2011, 454587.
- Petitot, I.A., Marentette, P.F., 1991. Babbling in manual mode: evidence for the ontogeny of language. *Science* 251, 1493–1496.



- Petrides, M., Pandya, D.N., 2009. Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biol.* 7, e1000170.
- Pickering, M.J., Clark, A., 2014. Getting ahead: forward models and their place in cognitive architecture. *Trends Cogn. Sci.* 18, 451–456.
- Pickering, M.J., Garrod, S., 2004. Toward a mechanistic psychology of dialogue. *Behav. Brain Sci.* 27, 169–190 discussion 190–226.
- Plaut, D.C., Shallice, T., 1993. Deep dyslexia: a case study of connectionist neuropsychology. *Cogn. Neuropsychol.* 10, 377–500.
- Plenz, D., Thiagarajan, T.C., 2007. The organizing principles of neuronal avalanches: cell assemblies in the cortex? *Trends Neurosci.* 30, 101–110.
- Postle, N., McMahon, K.L., Ashton, R., Meredith, M., de Zubicaray, G.I., 2008. Action word meaning representations in cytoarchitecturally defined primary and premotor cortices. *Neuroimage* 43, 634–644.
- Price, C., Thierry, G., Griffiths, T., 2005. Speech-specific auditory processing: where is it? *Trends Cogn. Sci.* 9, 271–276.
- Price, C.J., 2000. The anatomy of language: contributions from functional neuroimaging. *J. Anat.* 197 (Pt 3), 335–359.
- Price, C.J., 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech spoken language and reading. *Neuroimage* 62, 816–847.
- Prinz, W., 1990. *A Common Coding Approach to Perception and Action*. Springer, Heidelberg, New York.
- Pulvermüller, F., Fadiga, L., 2010. Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci.* 11, 351–360.
- Pulvermüller, F., Garagnani, M., 2014. From sensorimotor learning to memory cells in prefrontal and temporal association cortex: a neurocomputational study of disembodiment. *Cortex* 57, 1–21.
- Pulvermüller, F., Hauk, O., 2006. Category-specific processing of color and form words in left fronto-temporal cortex. *Cereb. Cortex* 16, 1193–1201.
- Pulvermüller, F., Knoblauch, A., 2009. Discrete combinatorial circuits emerging in neural networks: a mechanism for rules of grammar in the human brain? *Neural Netw.* 22, 161–172.
- Pulvermüller, F., Preissl, H., 1991. A cell assembly model of language. *Network: Computation in Neural Systems* 2, 455–468.
- Pulvermüller, F., Shtyrov, Y., 2006. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. *Prog. Neurobiol.* 79, 49–71.
- Pulvermüller, F., Shtyrov, Y., 2009. Spatio-temporal signatures of large-scale synfire chains for speech: MEG evidence. *Cereb. Cortex* 19, 79–88.
- Pulvermüller, F., Hummel, F., Härle, M., 2001a. Walking or Talking? Behavioral and neurophysiological correlates of action verb processing. *Brain Lang.* 78, 143–168.
- Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Martinkauppi, S., Ilmoniemi, R.J., Näätänen, R., 2001b. Memory traces for words as revealed by the mismatch negativity. *Neuroimage* 14, 607–616.
- Pulvermüller, F., Shtyrov, Y., Ilmoniemi, R.J., 2003. Spatio-temporal patterns of neural language processing: an MEG study using Minimum-Norm Current Estimates. *Neuroimage* 20, 1020–1025.
- Pulvermüller, F., Hauk, O., Nikulin, V.V., Ilmoniemi, R.J., 2005a. Functional links between motor and language systems. *Eur. J. Neurosci.* 21, 793–797.
- Pulvermüller, F., Shtyrov, Y., Ilmoniemi, R.J., 2005b. Brain signatures of meaning access in action word recognition. *J. Cogn. Neurosci.* 17, 884–892.
- Pulvermüller, F., Huss, M., Kherif Moscoso del Prado Martin, F.F., Hauk, O., Shtyrov, Y., 2006. Motor cortex maps articulatory features of speech sounds. *Proc. Natl. Acad. Sci. U. S. A.* 103, 7865–7870.
- Pulvermüller, F., Kherif, F., Hauk, O., Mohr, B., Nimmo-Smith, I., 2009a. Cortical cell assemblies for general lexical and category-specific semantic processing as revealed by fMRI cluster analysis. *Hum. Brain Mapp.* 30, 3837–3850.
- Pulvermüller, F., Shtyrov, Y., Hauk, O., 2009b. Understanding in an instant: neurophysiological evidence for mechanistic language circuits in the brain. *Brain Lang.* 110, 81–94.
- Pulvermüller, F., Cooper-Pye, E., Dine, C., Hauk, O., Nestor, P.J., Patterson, K., 2010. The word processing deficit in Semantic Dementia: all categories are equal but some categories are more equal than others. *J. Cogn. Neurosci.* 22, 2027–2041.
- Pulvermüller, F., Cook, C., Hauk, O., 2012a. Inflection in action: semantic motor system activation to noun- and verb-containing phrases is modulated by the presence of overt grammatical markers. *Neuroimage* 60, 1367–1379.
- Pulvermüller, F., Kiff, J., Shtyrov, Y., 2012b. Can language-action links explain language laterality? An ERP study of perceptual and articulatory learning of novel pseudowords. *Cortex* 48, 471–481.
- Pulvermüller, F., Cappelle, B., Shtyrov, Y., 2013. Brain basis of meaning, words, constructions, and grammar. In: Hoffmann, T., Trousdale, G. (Eds.), *Oxford Handbook of Construction Grammar*. Oxford University Press, Oxford, pp. 397–416.
- Pulvermüller, F., Garagnani, M., Wennekers, T., 2014a. Thinking in circuits: towards neurobiological explanation in cognitive neuroscience. *Biol. Cybern.* 108, 573–593.
- Pulvermüller, F., Moseley, R., Egorova, N., Shebani, Z., Boulenger, V., 2014b. Motor cognition – motor semantics: action-perception theory of cognitive and communicative cortical function. *Neuropsychologia* 55, 71–84.
- Pulvermüller, F., 1996. Hebb's concept of cell assemblies and the psychophysiology of word processing. *Psychophysiology* 33, 317–333.
- Pulvermüller, F., 1999. Words in the brain's language. *Behav. Brain Sci.* 22, 253–336.
- Pulvermüller, F., 2002a. A brain perspective on language mechanisms: from discrete neuronal ensembles to serial order. *Prog. Neurobiol.* 67, 85–111.
- Pulvermüller, F., 2002b. *The Neuroscience of Language*. Cambridge University Press, Cambridge.
- Pulvermüller, F., 2005. Brain mechanisms linking language and action. *Nat. Rev. Neurosci.* 6, 576–582.
- Pulvermüller, F., 2010. Brain embodiment of syntax and grammar: discrete combinatorial mechanisms spelt out in neuronal circuits. *Brain Lang.* 112, 167–179.
- Pulvermüller, F., 2012. Meaning and the brain: the neurosemantics of referential interactive, and combinatorial knowledge. *J. Neurolinguist.* 25, 423–459.
- Pulvermüller, F., 2013a. How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends Cogn. Sci.* 17, 458–470.
- Pulvermüller, F., 2013b. Semantic embodiment, disembodiment or misembodiment? In search of meaning in modules and neuron circuits. *Brain Lang.* 127, 86–103.
- Rüschmeyer, S.A., Brass, M., Friederici, A.D., 2007. Comprehending prehending: neural correlates of processing verbs with motor stems. *J. Cogn. Neurosci.* 19, 855–865.
- Ralph, M.A., Jefferies, E., Patterson, K., Rogers, T.T., 2017. The neural and computational bases of semantic cognition. *Nat. Rev. Neurosci.* 18, 42–55.
- Raposo, A., Moss, H.E., Stamatakis, E.A., Tyler, L.K., 2009. Modulation of motor and premotor cortices by actions: action words and action sentences. *Neuropsychologia* 47, 388–396.
- Reichardt, W., Varju, D., 1959. Übertragungseigenschaften im Auswertesystem für das Bewegungsehen. *Zeitschrift für Naturforschung* 14b, 674–689.
- Repetto, C., Colombo, B., Cipresso, P., Riva, G., 2013. The effects of rTMS over the primary motor cortex: the link between action and language. *Neuropsychologia* 51, 8–13.
- Rilling, J.K., Glasser, M.F., Preuss, T.M., Ma, X., Zhao, T., Hu, X., Behrens, T.E., 2008. The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat. Neurosci.* 11, 426–428.
- Rilling, J.K., Glasser, M.F., Jbabdi, S., Andersson, J., Preuss, T.M., 2012. Continuity, divergence, and the evolution of brain language pathways. *Front. Evol. Neurosci.* 3, 11.
- Rilling, J.K., 2014. Comparative primate neuroimaging: insights into human brain evolution. *Trends Cogn. Sci.* 18, 46–55.
- Rizzolatti, G., Arbib, M.A., 1998. Language within our grasp. *Trends Neurosci.* 21, 188–194.
- Rizzolatti, G., Craighero, L., 2004. The mirror-neuron system. *Ann. Rev. Neurosci.* 27, 169–192.
- Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L., 1996. Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* 3, 131–141.
- Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., Rozzi, S., 2014. Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol. Rev.* 94, 655–706.
- Rodriguez-Ferreiro, J., Gennari, S.P., Davies, R., Cuetos, F., 2011. Neural correlates of abstract verb processing. *J. Cogn. Neurosci.* 23, 106–118.
- Rogalsky, C., Love, T., Driscoll, D., Anderson, S.W., Hickok, G., 2011. Are mirror neurons the basis of speech perception? Evidence from five cases with damage to the purported human mirror system. *Neurocase* 17, 178–187.
- Rogers, J.C., Möttönen, R., Boyles, R., Watkins, K.E., 2014. Discrimination of speech and non-speech sounds following theta-burst stimulation of the motor cortex. *Front. Psychol.* 5, 754.
- Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S., Rauschecker, J.P., 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* 2, 1131–1136.
- Rosenbek, J.C., LaPointe, L.L., Wertz, R., 1995. *Aphasia: A Clinical Approach*. College-Hill Press, Boston.
- Rossini, P.M., Burke, D., Chen, R., Cohen, L.G., Daskalakis, Z., Di Iorio, R., Di Lazzaro, V., Ferreri, F., Fitzgerald, P.B., George, M.S., Hallett, M., Lefaucheur, J.P., Languth, B., Matsumoto, H., Miniussi, C., Nitsche, M.A., Pascual-Leone, A., Paulus, W., Rossi, S., Rothwell, J.C., Siebner, H.R., Ugawa, Y., Walsh, V., Ziemann, U., 2015. Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: basic principles and procedures for routine clinical and research application. An updated report from an I. F. C. N. Committee. *Clin. Neurophysiol.* 126, 1071–1107.
- Rueschemeyer, S.A., Gardner, T., Stoner, C., 2015. The Social N400 effect: how the presence of other listeners affects language comprehension. *Psychon. Bull. Rev.* 22, 128–134.
- Rumelhart, D.E., McClelland, J.L., 1987. Learning the past tense of English verbs: implicit rules or parallel distributed processing. In: MacWhinney, B. (Ed.), *Mechanisms of Language Acquisition*. Erlbaum, Hillsdale, NJ.
- Rupp, A., Uppenkamp, S., Gutschalk, A., Beucker, R., Patterson, R.D., Dau, T., Scherg, M., 2002. The representation of peripheral neural activity in the middle-latency evoked field of primary auditory cortex in humans(1). *Hear. Res.* 174, 19–31.
- Saffran, J.R., Aslin, R.N., Newport, E.L., 1996. Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.
- Schüz, A., Braitenberg, V., 2002. The human cortical white matter: quantitative aspects of cortico-cortical long-range connectivity. In: Schüz, A., Miller, R. (Eds.), *Cortical Aras: Unity and Diversity*. Taylor & Francis, London, New York, pp. 377–385.
- Schmid, H.-J., 2007. Entrenchment, salience, and basic levels. In: Geeraerts, D., Cuyckens, H. (Eds.), *The Oxford Handbook of Cognitive Linguistics*. Oxford University Press, Oxford, pp. 117–138.
- Schüz, A., Sultan, F., 2009. Brain connectivity and brain size. In: Squire, L.R. (Ed.), *Encyclopedia of Neuroscience*. Academic Press, Oxford, pp. 317–326.

- Schomers, M., Pulvermüller, F., 2016. Is the sensorimotor cortex relevant for speech perception and understanding? An integrative review. *Front. Hum. Neurosci.* 10, 435.
- Schomers, M., Kirilina, E., Weigand, A., Bajbouj, M., Pulvermüller, F., 2015. Causal influence of articulatory motor cortex on comprehending single spoken words: TMS evidence. *Cereb. Cortex* 25, 3894–3902.
- Schomers, M., Garagnani, M., Pulvermüller, F., 2017. Neurocomputational consequences of evolutionary connectivity changes in perisylvian language cortex. *J. Neurosci.* 37, 3045–3055.
- Searle, J.R., 1979. *Expression and Meaning*. Cambridge University Press, Cambridge.
- Searle, J.R., 1980. *Minds brains, and programs*. *Behav. Brain Sci.* 3, 417–457.
- Searle, J.R., 1984. *Minds, Brains, and Science*. Harvard University Press, Cambridge, MA.
- Shallice, T., 1988. *From Neuropsychology to Mental Structure*. Cambridge University Press, New York.
- Shebani, Z., Pulvermüller, F., 2013. Moving the hands and feet specifically impairs working memory for arm- and leg-related action words. *Cortex* 49, 222–231.
- Shebani, Z., Patterson, K., Nestor, P.J., Diaz-de-Greñu, L.Z., Dawson, K., Pulvermüller, F., 2017. Semantic word category processing in semantic dementia and posterior cortical atrophy. *Cortex* 93, 92–106.
- Shew, W.L., Plenz, D., 2013. The functional benefits of criticality in the cortex. *Neuroscientist* 19, 88–100.
- Shtyrov, Y., Lenzen, M., 2017. First-pass neocortical processing of spoken language takes only 30 msec: electrophysiological evidence. *Cogn. Neurosci.* 8, 24–38.
- Shtyrov, Y., Hauk, O., Pulvermüller, F., 2004. Distributed neuronal networks for encoding category-specific semantic information: the mismatch negativity to action words. *Eur. J. Neurosci.* 19, 1083–1092.
- Shtyrov, Y., Kujala, T., Pulvermüller, F., 2010. Interactions between language and attention systems: early automatic lexical processing? *J. Cogn. Neurosci.* 22, 1465–1478.
- Shtyrov, Y., Butorina, A., Nikolaeva, A., Stroganova, T., 2014. Automatic ultrarapid activation and inhibition of cortical motor systems in spoken word comprehension. *Proc. Natl. Acad. Sci. U. S. A.* 111, E1918–1923.
- Simmons, W.K., Ramjee, V., Beauchamp, M.S., McRae, K., Martin, A., Barsalou, L.W., 2007. A common neural substrate for perceiving and knowing about color. *Neuropsychologia* 45, 2802–2810.
- Skinner, B.F., 1957. *Verbal Behavior*. Appleton-Century-Crofts, New York.
- Skipper, J.L., Devlin, J.T., Lametti, D.R., 2017. The hearing ear is always found close to the speaking tongue: review of the role of the motor system in speech perception. *Brain Lang.* 164, 77–105.
- Smalle, E.H., Rogers, J., Möttönen, R., 2015. Dissociating contributions of the motor cortex to speech perception and response bias by using transcranial magnetic stimulation. *Cereb. Cortex* 25, 3690–3698.
- Sougné, J., 1998. Connectionism and the problem of multiple instantiation. *Trends Cogn. Sci.* 2, 183–189.
- Sporns, O., Honey, C.J., Kotter, R., 2007. Identification and classification of hubs in brain networks. *PLoS One* 2, e1049.
- Stalnaker, R.C., 2002. Common ground. *Linguist. Philos.* 25, 701–721.
- Steinmetz, H., Volkman, J., Jancke, L., Freund, H.J., 1991. Anatomical left-right asymmetry of language-related temporal cortex is different in left- and right-handers. *Ann. Neurol.* 29, 315–319.
- Steinschneider, M., Schroeder, C.E., Arezzo, J.C., Vaughan Jr., H.G., 1994. Speech-evoked activity in primary auditory cortex: effects of voice onset time. *Electroencephalogr. Clin. Neurophysiol.* 92, 30–43.
- Steinschneider, M., Nourski, K.V., Fishman, Y.I., 2013. Representation of speech in human auditory cortex: is it special? *Hear. Res.* 305, 57–73.
- Strauss, A., Kotz, S.A., Scharinger, M., Obleser, J., 2014. Alpha and theta brain oscillations index dissociable processes in spoken word recognition. *Neuroimage* 97, 387–395.
- Strijkers, K., Costa, A., 2016. The cortical dynamics of speaking: present shortcomings and future avenues. *Lang. Cogn. Neurosci.* 31, 484–503.
- Strijkers, K., Costa, A., Pulvermüller, F., in press. The cortical dynamics of speaking: Lexical and phonological knowledge simultaneously recruit the frontal and temporal cortex within 200 ms. *Neuroimage*.
- Szenkovits, G., Peelle, J.E., Norris, D., Davis, M.H., 2012. Individual differences in premotor and motor recruitment during speech perception. *Neuropsychologia* 50, 1380–1392.
- Tesnière, L., 2015. *Elements of Structural Syntax*. John Benjamins Publishing Company, Amsterdam.
- Tettamanti, M., Buccino, G., Saccuman, M.C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S.F., Perani, D., 2005. Listening to action-related sentences activates fronto-parietal motor circuits. *J. Cogn. Neurosci.* 17, 273–281.
- Thiebaut de Schotten, M., Dell'Acqua, F., Valabregue, R., Catani, M., 2012. Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex* 48, 82–96.
- Tomasello, M., Kruger, A.C., 1992. Joint attention on actions: acquiring verbs in ostensive and non-ostensive contexts. *J. Child Lang.* 19, 311–333.
- Tomasello, R., Garagnani, M., Wennekers, T., Pulvermüller, F., 2017. Brain connections of words: perceptions and actions: a neurobiological model of spatio-temporal semantic activation in the human cortex. *Neuropsychologia* 98, 111–129.
- Tomasello, M., 1995. Pragmatic contexts for early verb learning. In: Tomasello, M., Merriman, W.E. (Eds.), *Beyond Names for Things: Young Children's Acquisition of Verbs*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 115–146.
- Tomasello, M., 2005. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press, Cambridge, MA.
- Tomasello, M., 2008. *Origins of Human Communication*. MIT Press, Cambridge, MA.
- Tomasello, M., 2014. *A Natural History of Human Thinking*. Harvard University Press.
- Tomasino, B., Werner, C.J., Weiss, P.H., Fink, G.R., 2007. Stimulus properties matter more than perspective: an fMRI study of mental imagery and silent reading of action phrases. *Neuroimage* 36 (Suppl. 2), T128–141.
- Tranel, D., Kemmerer, D., Adolphs, R., Damasio, H., Damasio, A.R., 2003. Neural correlates of conceptual knowledge for actions. *Cogn. Neuropsychol.* 20, 409–432.
- Trumpp, N.M., Kliese, D., Hoenig, K., Haarmeier, T., Kiefer, M., 2013. Losing the sound of concepts: damage to auditory association cortex impairs the processing of sound-related concepts. *Cortex* 49, 474–486.
- Tyler, L.K., Stamatakis, E.A., Dick, E., Bright, P., Fletcher, P., Moss, H., 2003. Objects and their actions: evidence for a neurally distributed semantic system. *Neuroimage* 18, 542–557.
- Ueno, T., Saito, S., Rogers, T.T., Lambon Ralph, M.A., 2011. Lichtheim 2: synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron* 72, 385–396.
- Ulrich, M., Hoenig, K., Gron, G., Kiefer, M., 2013. Brain activation during masked and unmasked semantic priming: commonalities and differences. *J. Cogn. Neurosci.* 25, 2216–2229.
- Uppenkamp, S., Johnsrude, I.S., Norris, D., Marslen-Wilson, W., Patterson, R.D., 2006. Locating the initial stages of speech-sound processing in human temporal cortex. *Neuroimage* 31, 1284–1296.
- Vallabha, G.K., McClelland, J.L., Pons, F., Werker, J.F., Amano, S., 2007. Unsupervised learning of vowel categories from infant-directed speech. *Proc. Natl. Acad. Sci. U. S. A.* 104, 13273–13278.
- van Ackeren, M.J., Casasanto, D., Bekkering, H., Hagoort, P., Rueschemeyer, S.A., 2012. Pragmatics in action: indirect requests engage theory of mind areas and the cortical motor network. *J. Cogn. Neurosci.* 24, 2237–2247.
- van Ackeren, M.J., Smaragdi, A., Rueschemeyer, S.A., 2016. Neuronal interactions between mentalising and action systems during indirect request processing. *Soc. Cogn. Neurosci.* 11, 1402–1410.
- van Dam, W.O., Rueschemeyer, S.A., Bekkering, H., 2010. How specifically are action verbs represented in the neural motor system: an fMRI study. *Neuroimage* 53, 1318–1325.
- van Dam, W.O., van Dijk, M., Bekkering, H., Rueschemeyer, S.A., 2012. Flexibility in embodied lexical-semantic representations. *Hum. Brain Mapp.* 33, 2322–2333.
- van Dam, W.O., Brazil, I.A., Bekkering, H., Rueschemeyer, S.A., 2014. Flexibility in embodied language processing: context effects in lexical access. *Top Cogn. Sci.* 6, 407–424.
- van Elk, M., van Schie, H.T., Zwaan, R.A., Bekkering, H., 2010. The functional role of motor activation in language processing: motor cortical oscillations support lexical-semantic retrieval. *Neuroimage* 50, 665–677.
- van den Heuvel, M.P., Sporns, O., 2013. Network hubs in the human brain. *Trends Cogn. Sci.* 17, 683–696.
- von der Malsburg, C., 1999. The what and why of binding: the modeler's perspective. *Neuron* 24, 95–104 111–125.
- Van Berkum, J.J.A., 2009. The neuropragmatics of 'simple' utterance comprehension: an ERP review. In: Sauerland, U., Yatushiro, K. (Eds.), *Semantics and Pragmatics: From Experiment to Theory*. Palgrave Macmillan, London, pp. 276–316.
- Vannuscorps, G., Caramazza, A., 2016. Typical action perception and interpretation without motor simulation. *Proc. Natl. Acad. Sci. U. S. A.* 113, 86–91.
- Venezia, J.H., Saberi, K., Chubb, C., Hickok, G., 2012. Response bias modulates the speech motor system during syllable discrimination. *Front. Psychol.* 3, 157.
- Verduzco-Flores, S., Bodner, M., Ermentrout, B., Fuster, J.M., Zhou, Y., 2009. Working memory cells' behavior may be explained by cross-regional networks with synaptic facilitation. *PLoS One* 4, e6399.
- Vigliocco, G., Kousta, S.T., Della Rosa, P.A., Vinson, D.P., Tettamanti, M., Devlin, J.T., Cappa, S.F., 2014. The neural representation of abstract words: the role of emotion. *Cereb. Cortex* 24, 1767–1777.
- Vigneau, M., Beaucousin, V., Herve, P.Y., Duffau, H., Crivello, F., Houde, O., Mazoyer, B., Tzourio-Mazoyer, N., 2006. Meta-analyzing left hemisphere language areas: phonology semantics, and sentence processing. *Neuroimage* 30, 1414–1432.
- Vincent-Lamarre, P., Masse, A.B., Lopes, M., Lord, M., Marcotte, O., Harnad, S., 2016. The Latent Structure of Dictionaries. *Top Cogn Sci* 8, 625–659.
- Warrington, E.K., Shallice, T., 1984. Category specific semantic impairments. *Brain* 107, 829–854.
- Wassermann, E., Epstein, C., Ziemann, U., 2008. *Oxford Handbook of Transcranial Stimulation*. Oxford University Press.
- Watkins, K.E., Strafella, A.P., Paus, T., 2003. Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41, 989–994.
- Watson, C.E., Cardillo, E.R., Ianni, G.R., Chatterjee, A., 2013. Action concepts in the brain: an activation likelihood estimation meta-analysis. *J. Cogn. Neurosci.* 25, 1191–1205.
- Watson, J.B., 1925. *Behaviorism*. Transaction Publishers, New York.
- Weber, C., Wermter, S., Elshaw, M., 2006. A hybrid generative and predictive model of the motor cortex. *Neural Netw.* 19, 339–353.
- Wennekers, T., Palm, G., 2009. Syntactic sequencing in Hebbian cell assemblies. *Cogn Neurodyn* 3, 429–441.
- Wennekers, T., Garagnani, M., Pulvermüller, F., 2006. Language models based on Hebbian cell assemblies. *J. Physiol. Paris* 100, 16–30.
- Werker, J.F., Tees, R.C., 1999. Influences on infant speech processing: toward a new synthesis. *Annu. Rev. Psychol.* 50, 509–535.

- Werker, J.F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., Amano, S., 2007. Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition* 103, 147–162.
- Westermann, G., Reck Miranda, E., 2004. A new model of sensorimotor coupling in the development of speech. *Brain Lang.* 89, 393–400.
- Willems, R.M., Casasanto, D., 2011. Flexibility in embodied language understanding. *Front. Psychol.* 2, 116.
- Willems, R.M., Labruna, L., D'Esposito, M., Ivry, R., Casasanto, D., 2011. A functional role for the motor system in language understanding: evidence from theta-burst transcranial magnetic stimulation. *Psychol. Sci.* 22, 849–854.
- Willshaw, D.J., Buneman, O.P., Longuet-Higgins, H.C., 1969. Non-holographic associative memory. *Nature* 222, 960–962.
- Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csepe, V., Ilmoniemi, R.J., Näätänen, R., 1999. Brain responses reveal the learning of foreign language phonemes. *Psychophysiology* 36, 638–642.
- Wittgenstein, L., 1953. *Philosophical Investigations*. Blackwell Publishers, Oxford.
- Wierzbicka, A., 1985. Different cultures, different languages, different speech acts: Polish vs. English. *J Pragmatics* 9, 145–178.
- Wierzbicka, A., 2003. *Cross-cultural pragmatics: The semantics of human interaction..* Walter de Gruyter, Berlin.
- Young, M.P., Scannell, J.W., Burns, G.A., Blakemore, C., 1994. Analysis of connectivity: neural systems in the cerebral cortex. *Rev. Neurosci.* 5, 227–250.
- Yeung, H.H., Werker, J.F., 2013. Lip movements affect infants' audiovisual speech perception. *Psychol Sci* 24, 603–612.
- Zipser, D., Kehoe, B., Littlewort, G., Fuster, J.M., 1993. A spiking network model of short-term active memory. *J. Neurosci.* 13, 3406–3420.
- Zwaan, R.A., 2014. Embodiment and language comprehension: reframing the discussion. *Trends Cogn. Sci.* 18, 229–234.