Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions

Markus Kiefer a,∗ and Friedemann Pulvermüller b

a University of Ulm, Department of Psychiatry, Ulm, Germany
b Medical Research Council, Cognition and Brain Sciences Unit, Cambridge, UK

ARTICLE INFO
Article history:
Received 5 August 2010
Reviewed 29 November 2010
Revised 18 February 2011
Accepted 4 April 2011
Published online 30 April 2011

Keywords:
Semantic memory
Language
Grounded cognition
Brain-damaged patients
Neuroimaging

ABSTRACT
Conceptual representations in long-term memory crucially contribute to perception and action, language and thought. However, the precise nature of these conceptual memory traces is discussed controversially. In particular, the grounding of concepts in the sensory and motor brain systems is the focus of a current debate. Here, we review theoretical accounts of the structure and neural basis of conceptual memory and evaluate them in light of recent empirical evidence. Models of conceptual processing can be distinguished along four dimensions: (i) amodal versus modality-specific, (ii) localist versus distributed, (iii) innate versus experience-dependent, and (iv) stable versus flexible. A systematic review of behavioral and neuroimaging studies in healthy participants along with brain-damaged patients will then be used to evaluate the competing theoretical approaches to conceptual representations. These findings indicate that concepts are flexible, distributed representations comprised of modality-specific conceptual features. Conceptual features are stored in distinct sensory and motor brain areas depending on specific sensory and motor experiences during concept acquisition. Three important controversial issues are highlighted, which require further clarification in future research: the existence of an amodal conceptual representation in the anterior temporal lobe, the causal role of sensory and motor activation for conceptual processing and the grounding of abstract concepts in perception and action. We argue that an embodiment view of conceptual representations realized as distributed sensory and motor cell assemblies that are complemented by supramodal integration brain circuits may serve as a theoretical framework to guide future research on concrete and abstract concepts.

© 2011 Elsevier Srl. All rights reserved.

1. Introduction
Concepts held in semantic long-term memory are important building blocks of human cognition. They are the basis for object recognition, action planning, language, and thought, because they constitute the meaning of objects, events and abstract ideas (Humphreys et al., 1988; Levelt et al., 1999). Concepts have a central role in information processing...
because they help to interface perception (information collection from the environment) and action (information emission to the environment). As concepts are essential constituents of word meaning (see also the contribution by Meteyard et al., 2012), they provide semantic knowledge for the comprehension of verbal communication (Kiefer, 2001; McRae et al., 1997; Humphreys et al., 1999; Pulvermüller, 1999). When speaking about “semantic knowledge”, we here refer to conceptual representations systematically linked to words, their meaningful parts, or constructions composed of several words.

It is well accepted that concepts include the sum of our sensory and motor experiences with the environment in a categorial fashion. For instance, the concept “cat” includes the information that a cat has four legs, is furry, meows, can move or can be petted. Concepts refer to categories of objects, events or ideas because conceptual representations generalize across specific exemplars and situations, in which we have encountered the referent in the past (e.g., specific exemplars of cats). Thus, concepts in semantic memory are abstract in the sense that they are not specific in time or place (Tulving, 1972). In contrast, episodic memory is taken to store individual experiences from the past happening a specific time at a specific place.

Despite this general agreement regarding the content of concepts, the nature of conceptual representations is still a matter of debate. In particular, the role of sensory and motor representations in constituting concepts has been discussed controversially. This debate dates back to the ancient Greek philosophers more than 2000 years ago (for an overview, see Runes, 1962): since the early days of Platon, philosophers have speculated whether or not concepts are essentially grounded in our senses and in our actions with the environment (Markie, 2008). Platon and more modern philosophers such as Descartes, Leibnitz and Kant were very skeptical in this regard. According to their rationalist views, concepts are mental entities fundamentally distinct from sensory impressions (which, in their views, may even be a requirement for the latter). They are based upon innate categories or are formed by reasoning. These rationalist philosophers argued that concepts cannot reflect sensations because perception is too unstructured in order to provide the basis for meaningful concepts. In contrast, Aristotle and more modern empiricist philosophers, including Locke and Hume, proposed that all concepts are derived from sensory experiences (impressions) and ideas, which are faint copies of impressions. Because any significant knowledge is to be gained ultimately from one’s sense-based experience, concepts based upon reasoning can, in this view, not provide valid information about the physical and social world.

These philosophical reflections concerning the role of sensory and motor representations for establishing concepts are paralleled by a similar debate in contemporary cognitive and neuroscientific research on conceptual cognition. In the modern cognitive sciences, concepts are traditionally specified as abstract mental entities, different from the perceptual or motor brain systems (Quillian, 1969; Anderson, 1983; Tyler and Moss, 2001; Pylyshyn, 1984): sensory or motor features of objects and events are transformed into a common amodal representation format, in which original modality-specific information is lost. Although traditional theories do not deny the involvement of the sensory and motor systems in conceptual tasks, they assume that these modality-specific representations activated during language comprehension or conceptual thinking may only be auxiliary or concomitant processes, similar to imagery, while conceptual information proper is represented in an amodal fashion (Mahon and Caramazza, 2009; Machery, 2007; Weiskopf, 2010) in heteromodal association cortex (McClelland and Rogers, 2003; Rogers et al., 2004).

Challenging this classical view, recent modality-specific approaches propose close links between the sensory and motor brain systems on the one hand and the conceptual system on the other hand. They assume that concepts are essentially grounded in perception and action (Barsalou et al., 2003; Kiefer and Spitzer, 2001; Pulvermüller, 2005; Warrington and McCarthy, 1982; Martin and Chao, 2001; Lakoff and Johnson, 1999; Gallese and Lakoff, 2005). Representations of external (perception) and internal states (proprioception, emotion and introspection) as well as actions critically constitute concepts and thus play a fundamental functional role in linguistic understanding and conceptual thinking.

To put this debate in a neuroscientific context, we consider two ways of specifying the brain basis of concepts. In both a rationalist and an empiricist approach, different categories of knowledge, such as animals and tools, can be considered to be fundamentally different. In the rationalist perspective, such a difference would then have a fundamental a priori character, possibly with different brain regions being responsible for their respective processing. In the empiricist perspective, the category difference would arise from sensory-motor experience; animals are known from shape and color information whereas tools are known from acting with them and from observing them moving. These alternative views make contrary predictions on brain activation during animal and tool information processing in the brain. Clearly, the rationalist view without any neuronal specification does not make precise predictions on where in the brain conceptual information might emerge, but it implies conceptual activation outside sensory and motor areas and, crucially, independence of such activation from sensory experience and motor action. The empiricist perspective instead implies that sensory and motor systems of the brain contribute substantially to the formation of animal and tool concepts, so that animal processing is predicted to involve visual brain areas contributing to color and shape processing, whereas action-related motor areas and visual areas related to motion processing would be involved in tool concept processing. Under sensory and action deprivation, similar dissociations of cortical area involvement would, however, be difficult to explain in the empiricist framework. Recent theoretical proposals in cognitive neuroscience show indeed similarity to rationalist (Caramazza and Shelton, 1998; Mahon and Caramazza, 2009) and empiricist frameworks (Barsalou, 2008; Kiefer, 2005; Pulvermüller and Fadiga, 2010).

Behavioral and neuroimaging studies in healthy volunteers as well as in brain-damaged patients now allow addressing these fundamental questions about conceptual cognition objectively. Neuroimaging techniques can render visible the patterns of brain activation in both space and time that arise
during processing of concepts. The functional role of sensory and motor representation for conceptual processing can be assessed with experimental psychological methods and with transcranial magnetic stimulation (TMS), thus complementing the study of patients with lesions in specific brain areas. These advances in cognitive psychology and cognitive neuroscience enable us to elucidate the nature of conceptual representations in mind and brain and to experimentally test the different theoretical approaches. The present paper provides an overview of the different theoretical views on conceptual representations and discusses these theories in the light of recent behavioral and neuroscientific data. As we will demonstrate below, accumulating evidence points to close links between the conceptual and the sensory-motor brain systems in favor of the embodiment view of concepts. We close this article with a discussion of open questions and future research directions.

2. The nature of conceptual representations: a systematization of theoretical developments

Before we describe in detail recent theories of conceptual representations, we would like to introduce four important dimensions, which, as we feel, might help to systematize and evaluate these theoretical approaches. As we have already mentioned above, an important distinguishing feature for theories of conceptual representations is whether they consider conceptual representations as amodal or modality-specific entities: amodal theories assume that conceptual representations are fundamentally distinct from representations in the perceptual and motor systems of the brain. Sensory and motor information from the environment is transformed into an amodal symbolic representation format which lacks a direct representation of sensory or motor events. In contrast, modality-specific theories claim that conceptual representations are functionally and neuroanatomically grounded in, and explainable by, perceptual and motor representations (grounded cognition or embodiment theory).

The local-distributed dimension refers to the internal structure of conceptual representations. Theories with localist representations assume that a concept is coded by one representational unit such as one node within a semantic network. The idea of local representations is rooted in the idea that single nerve cells may represent concepts (Barlow, 1972), a kind of “grandmother cell theory” which most researchers are slightly hesitant to accept today (Hubel, 1995; but see Bowers, 2009). In distributed theories, each concept is coded by multiple representational units, usually using the activation vector of an entire “layer” of a network. Thus, the activation pattern across several representational units establishes the concept.

A third dimension distinguishes between theories of conceptual representations that essentially propose an experience-dependent formation of concepts from those theories that claim innate a priori categorial specializations. This dimension has already been touched upon in the discussion of “rationalist” and “empiricist” approaches to cognitive neuroscience in the previous section, although the classic rationalist position presumes both innateness and abstract symbolic status of categories, whereas the empiricist position favors experience dependence along with sensory-motor grounding.

Finally, we would like to introduce the flexible-stable dimension. In theories assuming stable conceptual representations, concepts are considered to be situational invariant mental knowledge entities. In contrast, theories assuming flexible conceptual representations view concepts as flexibly tailored to the current contextual constraints, because they are constituted of dynamically recruited features depending on the context. The stable-flexible distinction has its roots in modern analytical philosophy and linguistics, where it was once an issue whether words carry a core meaning constituted by a stable concept which is invariably accessed each time the word is used. It was, however, pointed out that the different uses and meanings of words – take the word “game” as an example – are manifold, not sharing a fixed set of conceptual features (chess, video games, football). Any definition would have difficulty capturing the actual range of meanings, which, as the philosopher Ludwig Wittgenstein pointed out, relate to each other as the members of a big family do, with some pairs exhibiting striking similarities but others varying substantially (Wittgenstein, 1953). However, as we will see, some meanwhile slightly outdated psychological theories of concepts include the idea of stable conceptual representations invariably immanent to each application of the concept.

It should be noted in this context that the four dimensions outlined above are not entirely independent of each other because some combinations of dimensions frequently go together with others: for instance, a theory assuming flexible conceptual representation necessarily entails a distributed representation because localist representations cannot subserve conceptual flexibility efficiently. In order to mimic conceptual flexibility with localist representations, several separate concepts represented as distinct nodes would have to be assumed for coding the numerous context-specific meaning variations (e.g., hammer-for-pounding-nails, hammer-for-throwing, hammer-used-as-replacement-for-a-paperweight, etc.), an assumption, which lacks both parsimony and plausibility. Table 1 provides an overview of the classes of theories described in detail below together with their classification along the four dimensions.

2.1. Localist amodal conceptual representations

The earliest and for many decades most influential theories of conceptual representation developed in cognitive psychology and artificial intelligence assume that a concept is represented as a node within a unitary semantic network (Collins and Loftus, 1975; Collins and Quillian, 1969; Quillian, 1969; Bowers, 2009; Anderson, 1983). Each conceptual node is linked to other nodes, which represent meaningfully related concepts. This structure of connected nodes within the network provides propositional knowledge about a concept (e.g., that birds can fly) in an explicit symbolic fashion. In other words, each node carries a label which specifies the particular knowledge content. As each concept is assigned to a specific representational node, which is distinct from sensory-motor representations, this classic semantic network model is a prototypical example of a theory assuming localist amodal
conceptual representations (Collins and Loftus, 1975; Collins and Quillian, 1969; Quillian, 1969). The localist representation format necessarily results in stable concepts, which are situational invariant (see also the previous paragraph). It should be noted, though, that the stability assumption is not explicitly stated in these models, which, partly due to their inflexibility, are considered slightly outdated by many psychologists today. Although classical semantic network models do not explicitly incorporate assumptions about concept acquisition, they do imply that the structure of conceptual knowledge reflects the statistical information derived from the experiences during concept acquisition, rather than an innate categorial organization.

2.2. Distributed amodal conceptual representations

The localist representation format of classical semantic network models appeared unsatisfactory to many researchers due to a lack of conceptual flexibility. For that reason, distributed semantic memory models were developed, in which concepts are coded by multiple representational units. As concepts are thought to be formed of several simple representational units, the contribution of these units to a concept can vary as a function of the context in which a concept is accessed, therefore providing the basis for conceptual flexibility. Some early semantic feature list models can be considered as an instance of a theory assuming distributed amodal conceptual representations (e.g., Smith et al., 1974). In feature list models, a concept is comprised of a set of semantic features, which code basic meaning aspects in an explicit fashion. For instance, the concept "cat" is constituted by the features "head", "four legs", "tail", "furry", "moves", "meows" and "eats mice". In other models, which build upon the parallel distributed processing (PDP) or connectionist framework of cognition (Rumelhart and McClelland, 1986), conceptual representations depend on an activation pattern of all neuron-like processing units within a network section, a so-called "layer" (Devlin et al., 1998; McClelland and Rogers, 2003; Rogers et al., 2004; Tyler and Moss, 2001; Caramazza et al., 1990).

Conceptual knowledge is not explicitly represented in form of symbolic features (such as “head” or “four legs”) or single nodes, but recovered through propagation of activation among the processing units, which are connected in a network. The connection weights (or strengths) between the processing units within the network are shaped by experience and adjusted according to learning algorithms. Common to these types of semantic feature list approaches and distributed network models are the assumption that conceptual knowledge is represented in an amodal format, separate from the perception and action brain systems, within a unitary conceptual system that stores all kinds of information irrespective of knowledge modality (e.g., visual or action-related) or category (e.g., animals or tools). It has been proposed that this unitary conceptual system has its neural substrate in anterior temporal lobe structures (see below).

It should be noted, however, that feature list and PDP models are not necessarily amodal in nature, but can also incorporate the notion of modality-specific conceptual representations as we will discuss below (Farah and...
McClelland, 1991; Vigliocco et al., 2004; Plaut, 2002; Simmons and Barsalou, 2003).

2.3. Domain-specific conceptual representations

Domain-specific theories of conceptual representations have given up the notion of a unitary conceptual memory system and propose innate distinct conceptual subsystems that dissect the semantic space in an a priori fashion, for example into animals, fruit, tools, etc (Caramazza and Mahon, 2003; Caramazza and Shelton, 1998). According to this view, evolutionary pressure has formed specialized brain circuits, which contain knowledge of one conceptual domain in order to optimally subserve rapid identification of objects, because these representations are important for survival. These innate domain-specific semantic stores represent concepts in an amodal fashion distinct from sensory and motor systems. Although domain-specific theories do not make explicit assumptions about the internal structure of concepts, they appear to endorse stable conceptual representations, which are represented locally. It has been suggested that the conceptual domain of animals is represented in left anterior temporal cortex whereas the conceptual domain of tools is represented more dorsally in parietal areas. However, these suggestions about cortical localizations of concept types do not appear to follow from any theory-immanent assumption and therefore may seem slightly a posteriori in nature.

2.4. Embodied modality-specific distributed and flexible conceptual representations

Like domain-specific theories, modality-specific theories of embodied conceptual representations propose a partitioning of conceptual memory in several subsystems. However, unlike domain-specific theories, in which innate categorial dissections are the driving force for the formation of amodal conceptual representations, modality-specific theories propose that concepts are essentially grounded in perception and action (Farah and McClelland, 1991; Plaut, 2002; Vigliocco et al., 2004; Barsalou, 2008; Humphreys and Forde, 2001; Kiefer and Spitzer, 2001; Pulvermüller and Fadiga, 2010; Warrington and Shallice, 1984; Martin, 2007). According to modality-specific approaches, representations of the perception of external and internal states (emotion, introspection) as well as actions constitute a concept and thus play a functional role in linguistic understanding and conceptual thinking. Concepts are assumed to be embodied (Lakoff and Johnson, 1999; Gallese and Lakoff, 2005) in the sense that interactions with other individuals and with objects lead to the formation of their conceptual memory traces in modality-specific brain areas, which typically process the corresponding sensory and action-related information. A similar point can be made about emotional information systematically co-occurring with, and therefore being linked to, words (Pulvermüller and Schumann, 1994; Vigliocco et al., 2009). Conceptual features (e.g., visual, acoustic, action-related, emotional) are represented by cortical cell assemblies distributed over sensory, motor, and emotional regions of the brain. These circuits are being established during concept acquisition as a consequence of correlated neuronal activity in neurons linked by way of pre-existing local connections and long range fiber bundles. Therefore, access to concepts involves a partial reinstatement of brain activity during experiences and actions the words are typically used to speak about. As distributed cell assemblies involving sensory and motor areas constitute the neural basis of concepts, these concepts themselves can be considered to be represented in a distributed fashion, also at a functional level.

Please note that distributed modality-specific representations constituting a concept can be modeled using feature lists (Vigliocco et al., 2009; Vigliocco et al., 2004) or PDP frameworks (Kiefer and Spitzer, 2001; Farah and McClelland, 1991; Plaut, 2002; Simmons and Barsalou, 2003; Pulvermüller, 2008b). Linguistic word representations (e.g., phonology, orthography) variably activate conceptual features that constitute the concept as a function of the linguistic and non-linguistic context. The mapping between word representations and conceptual representations may be achieved either directly (Kiefer, 2001; McRae et al., 1997; Humphreys et al., 1999; Pulvermüller, 1999), as we assume, or, alternatively, indirectly, via an intermediate lexico-semantic representation that is thought to support conceptual feature integration and lexicalization of concepts (Levelt et al., 1999; Vigliocco et al., 2004; Bierwisch and Schreuder, 1992). However, a detailed discussion about the existence of a lexico-semantic representation level, over and above that of conceptual representations, is beyond the scope of the present review (but see, Pulvermüller, 1999; Pulvermüller and Fadiga, 2010). Clearly, the relationship of word and conceptual representations is not necessarily 1:1. The many facets of the “game” concept, for instance, can be captured by variable and context-dependent activation of different modality-specific feature representations in sensory and motor areas. The variable and context-dependent contribution of modality-specific features to the concept thus provides the functional basis for conceptual flexibility.

We now would like to explicitly address misinterpretations previously used to reject modality-specific approaches in favor of amodal theories (Chatterjee, 2010; Machery, 2007; Weiskopf, 2010): (i) activation of modality-specific representations in conceptual task is not necessarily conscious, as in imagery, but may also occur unconsciously in a fast and automatic mode (Kiefer et al., 2008; Pulvermüller, 2007), (ii) modality-specific long-term memory conceptual representations do not resemble camera-like recordings of external and internal experiences. Instead, representations are shaped by the conjoint influence of direct experiences, imaginations, reinstated activations through language under the general influence of attention and action goals (Kiefer and Barsalou, 2011), (iii) furthermore, the notion of conceptual flexibility implies that access to a concept during language comprehension or thinking cannot be conceived as a replay of stored sensory-motor information as in a movie, but as a context-specific situation-dependent dynamic activation process (Hoenig et al., 2008).

3. Current evidence

In this section, we describe latest results from behavioral, neuropsychological, electrophysiological and neuroimaging experiments addressing the nature of conceptual
representations. This review will allow us to evaluate the different theoretical approaches. In order to facilitate appreciation of the findings described in this review, we will organize this section along the four dimensions, which we have used in the previous section to characterize the different theories. For each dimension, we will ask whether the available data already allow a definitive specification, for example whether they allow for a decision on whether representations are amodal or modality-specific. Each subsection will close with a concluding remark, which integrates the reported findings.

3.1. **Amodal – modality-specific dimension**

Some researchers posit that a center of semantic processing represents concepts across-the-board in an amodal format, functionally and neuroanatomically distinct from sensory and motor representations. Although different theories envisage the conceptual center, or “hub”, to be placed in different parts of the brain, they share the common assumption of such an amodal global hotspot for concepts. The competing perspective is that concepts are embodied as and grounded in distributed modality-specific sensory and motor brain systems. Depending on the nature of the concept (e.g., conceptual category), these sensory and motor systems differentially contribute to the conceptual representation. For instance, action-related conceptual features have been shown to be more relevant for representing objects from artifactual (e.g., tools) than from natural (e.g., animals) categories (McRae and Cree, 2002). As this controversy about the amodal versus modality-specific representation of concepts is one of the most prominent issues in current cognitive brain research, we can only discuss in depth a selection of relevant findings and refer to complementary reviews when pointing to some of the relevant findings.

Evidence for amodal conceptual representations comes from neuropsychological studies in patients with degenerative diseases of the brain. In one form of fronto-temporal dementia called Semantic Dementia, a relatively focal brain damage in the temporal pole and its vicinity results in an impairment of conceptual processing (Hodges et al., 1992; Patterson et al., 2007). Patients with Semantic Dementia exhibit a general loss of conceptual knowledge across all conceptual domains (including animal, tools etc.) or conceptual feature types (e.g., visual, auditory, action-related). Patients are more severely impaired for atypical exemplars of a category, which share just a few features with other category exemplars, than for typical exemplars, which have many features in common. The impairment in Semantic Dementia therefore demonstrates a sensitivity to abstract relations between concepts (Rogers et al., 2004; Patterson et al., 2007). Although brain-damage in Semantic Dementia is found with further progression of the disease also in posterior temporal cortex as well as in frontal areas, lesions in the temporal poles seem to be sufficient for the emergence of the typical semantic deficits, as is apparent at early stages of the disease (Patterson et al., 2007). The functional importance of anterior temporal areas for semantic processing has been recently replicated in healthy volunteers: stimulating anterior temporal areas with TMS deteriorates performance in semantic tasks for pictures and words similar to the deficit seen in Semantic Dementia (Pobric et al., 2010). As the impairment in Semantic Dementia generalizes across conceptual information derived from different sensory and motor channels, but is highly sensitive to structural relations between concepts, it has been proposed that areas within the anterior temporal cortex are the neural substrate of an amodal conceptual representation (Patterson et al., 2007; McClelland and Rogers, 2003; Rogers et al., 2004).

It is well-known that areas far from the temporal poles are also necessary for accurate conceptual processing. Lesions caused by massive stroke of the middle cerebral artery, which affect the left cortex around the sylvian fissure have sometimes been observed to lead to difficulties in processing concepts of tools and small manipulable objects (Warrington and McCarthy, 1987). Likewise, stroke primarily affecting the frontal cortex and degenerative brain disease primarily affecting the motor system have been found to lead to deficits in processing action-related verbs (Bak et al., 2001; Cotelli et al., 2006; Daniele et al., 1994). Even patients with sometimes small unilateral lesion in the motor and premotor cortex of the non-dominant right hemisphere exhibit a specific impairment in action verb processing compared with nouns (Neininger and Pulvermüller, 2003, 2001). As action verbs and object nouns differ not only with regard to the relevance of action representations, but also along other dimensions such as morpho-syntactic marking in sentences (Schnur et al., 2009; Vigliocco et al., 2011), any noun/verb dissociation can, in principle, receive multiple interpretations. Remarkably, however, these deficits in action processing can also be found when assessed non-verbally using pictures thus confirming that a genuine conceptual category deficit is the core defect (Kemmerer et al., 2011; Bak et al., 2006). Furthermore, dissociations between semantic word categories that cross grammatical boundaries are important in this context: aphasics patients showed a selective deficit for both nouns and verbs related to concepts involving a hand action (Arevalo et al., 2007). Similarly, patients with an impairment in accessing knowledge about artifactual objects (e.g., tools), for which action-related information is highly relevant, most frequently suffer from lesions in frontal and/or parietal motor areas (Gainotti, 2004; Gainotti et al., 1995). In contrast, patients who show deficits in processing natural objects, for which visual information is highly relevant (e.g., animals), most frequently have lesions in the visual association cortex within inferior occipital and temporal brain areas (Tanel et al., 1997a, 1997b; Hart and Gordon, 1992). These results further substantiate a semantic origin of category dissociations.

Not only patient studies point to a role of different cortical areas in semantic processing. Neurophysiological and neuro-imaging results have provided converging evidence on the differential involvement of brain areas in the processing of words and concepts of different kinds. Electroencephalography (EEG) and magnetoencephalography (MEG) (e.g., Kiefer, 2001, 2005; Pulvermüller et al., 1999; Pulvermüller et al., 2005b) as well as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) studies in healthy participants (e.g., Hoenig et al., 2008; Martin et al., 1996; Simmons et al., 2005) demonstrated activation in sensory and motor brain areas in a range of conceptual tasks.
There is neurophysiological evidence that concepts are functionally and neuroanatomically linked to sensory brain regions (Kiefer et al., 2008): recognition of words denoting objects, for which acoustic features are highly relevant (e.g., “telephone”), ignited cell assemblies in left posterior superior and middle temporal gyri (pSTG/MTG) that were also activated by sound perception (Fig. 1A). Importantly, activity within this part of auditory association cortex increased selectively as a function of acoustic, but not of visual and action-related feature relevance, indicating that left pSTG/MTG specifically codes acoustic conceptual knowledge. As event-related potential (ERP) recordings revealed an early onset of this activity at 150 msec after word onset (Fig. 1B), activity in pSTG/MTG most likely reflects rapid access to acoustic conceptual features and not later post-conceptual processes such as imagery. These results therefore provide direct evidence for a link between perceptual and conceptual acoustic processing. They demonstrate that access to concepts depends on a partial reinstatement of brain activity as it occurs during object perception. Recent neuro-psychological evidence suggests that left pSTG/MTG even plays a causal role in processing sound-related concepts: a patient with a focal lesion in this area was selectively impaired in accessing concepts, for which acoustic features are highly relevant (Trumpp et al., 2011). Activation of sensory brain regions in conceptual processing has been documented not only for the auditory modality, but also for visually (Kiefer, 2005; Chao et al., 1999; Sim and Kiefer, 2005; Pulvermüller and Hauk, 2006; Moscoso del Prado Martin et al., 2006) and even odor-related words (Gonzalez et al., 2006; Simmons et al., 2005). For example, reading odor-related words (“garlic”, “cinnamon”, “jasmine”) elicited activation in the primary olfactory cortex compared with neutral control words (Fig. 2) (Gonzalez et al., 2006). This shows that the

---

**Fig. 1** – The grounding of concepts in auditory brain systems: (A) Functional–anatomical overlap between conceptual and perceptual processing of acoustic features. Increased functional activation to words with acoustic conceptual features (conceptual processing) overlaps with brain activation during listening to real sounds (sound perception) in left pSTG/MTG. Shown are contiguous slices centered on the peak coordinates. (B) Time course of conceptual processing of acoustic features. Top: scalp ERPs to words with versus without acoustic features at central electrodes. Potentials are collapsed across central electrode sites. The arrow indicates the onset of the effect. Bottom: brain electrical sources of scalp ERPs: maps of cortical currents calculated according to the minimum norm algorithm from the ERP difference waves (with vs without acoustic features). Maps are shown for the respective maxima in global field power. Strongest cortical currents (visualized in blue color) were observed in and close to left pSTG/MTG. Modified after Kiefer et al. (2008).

**Fig. 2** – The grounding of concepts in olfactory brain systems: compared with matched control words, odor-related words such as “cinnamon” activate a range of brain regions implicated in odor sensation and emotion processing, including piriform cortex (BA47), insula, right amygdala, cingulate cortex, caudate nucleus and thalamus. Modified after Gonzalez et al. (2006).
modality-specific representation of conceptual features in
the corresponding sensory brain areas is a general principle
that applies to many sensory modalities.

Extended work has focused on revealing the relevance of
the motor system for conceptual processing (Boulenger et al.,
2009; Hauk et al., 2004; Hauk and Pulvermüller, 2004;
Pulvermüller et al., 2001; Shtyrov et al., 2004; Pulvermüller
et al., 2000). Conceptual processing of action words activates
motor areas as a function of the body part involved in carrying
out the action (Fig. 3): foot-related action words (e.g., “kick”) activated more dorsal parts of the motor cortex, whereas
hand-related (e.g., “pick”) and mouth-related (e.g., “lick”) action words activated lateral and ventral parts of the motor
cortex (Hauk et al., 2004; Hauk and Pulvermüller, 2004). This
activation pattern resembles the known somatotopy of the
motor cortex, suggesting that conceptual meaning of action
words is embodied in motor areas in an effector-specific
fashion. The results on brain correlates of action word
semantics are especially significant because they demonstr-
ated for the first time that the locus of category-specific
semantic activation in the brain could be predicted a priori:
the activations for arm- and leg-action concepts were found in

Fig. 3 – The grounding of concepts in motor brain systems: embodied semantic circuits are thought to bind word form
knowledge to knowledge about action programs in motor and premotor cortex. As the motor representations of different
parts of the body are located in different parts of the motor systems, the semantic circuits of face-, arm and leg-related
action words differ (Model diagram at the top). fMRI evidence shows overlapping activation during passive reading of face-, arm- and leg-related action words and related movements of body parts (Hauk et al., 2004). Stimulation of premotor arm and
leg representations using TMS specifically speeds the recognition of arm- or leg-related words (Pulvermüller et al., 2005a),
demonstrating a causal influence of motor activity on word processing. EEG and MEG studies showed early activation of
specific parts of the motor system to face-, arm- and leg-related words within 150–220 msec after stimulus information
allows for word recognition, even during tasks where subjects are distracted from language and semantic processing; this
argues for early and automatic semantic processing in the motor system. Modified after Pulvermüller et al. (2000); Hauk
and Pulvermüller (2004); Shtyrov et al. (2004); Pulvermüller et al. (2005a).
a wider area active during body movements, when the finger and foot, respectively, were moved by the participants. This local predictability is strong evidence for a sensory-motor “embodied” account, according to which the same, or closely adjacent, circuits in the motor system take a role in motor and conceptual processing. It is also noteworthy that the activations of specific motor regions by action words and sentences emerged quite rapidly, within 200 msec after the critical meaningful stimulus could be recognized. This suggests that early conceptual access is realized by the somatotopic activations, rather than a second order, late process (Pulvermüller, 2005).

Neurophysiological studies only provide correlational information about conceptual processing in the perception and action systems. To find out whether there is a functional contribution, a causal influence of specific areas, lesion studies are necessary. Above, we have already described how damage to visual, auditory or motor association cortex can impair performance in conceptual processing. Most of the available patient work, however, focuses on the results of large brain lesions and, if small lesions are present, these vary considerably between patients. However, brain models of conceptual processing make rather exact and specific predictions on contributions of circumscribed areas, which might best be tested in small-lesion studies. One solution is offered by temporarily influencing the functional state of small cortical areas using TMS. TMS to the motor cortex influences performance in conceptual tasks specifically for stimuli (e.g., action words), for which action-related information is highly relevant (Devlin and Watkins, 2007; Pulvermüller et al., 2005a; Buccino et al., 2005). Crucially, magnetic stimulation of the hand area in motor cortex tended to improve the recognition of arm related words, whereas TMS to the foot region improved leg word processing (Pulvermüller et al., 2005a). Stimulation of the motor system therefore has an effect on the processing of action-related concepts in an effector-specific fashion (Fig. 3).

Similarly, behavioral experiments using priming or interference paradigms show that previous activation of sensory (Vermeulen et al., 2008) or action-related representations (Kiefer et al., 2011; Helbig et al., 2006; Myung et al., 2006) modulates subsequent performance in conceptual tasks, in particular when the corresponding sensory or motor information is highly relevant to the target concept (Witt et al., 2010). For instance, observing an action movie as a prime (e.g., seeing a hand hammering) facilitates conceptual access to a subsequently presented manipulable target object when the target is associated with a similar action (e.g., axe) as the prime movie compared with a dissimilar one (e.g., saw) (Helbig et al., 2010). Along with a range of concordant behavioral work (for reviews see, Barsalou, 2008; Fischer and Zwaan, 2008), these findings demonstrate that sensory and motor representations play a functional role in conceptual processing and rule out the possibility that activation of modality-specific representations in conceptual tasks is epiphenomenal, reflecting for instance task-irrelevant associative processes (McClelland and Rogers, 2003; Machery, 2007).

Our review so far shows that there are arguments in favor of both amodal and modality-specific conceptual representations. It is therefore a logical consequence in recent proposals to integrate the two kinds of models with each other, envisaging a conceptual system comprised of modality-specific systems in sensory and motor areas complemented by an amodal “conceptual hub”, which integrates the distributed modality-specific representations in a common supramodal semantic space (Simmons and Barsalou, 2003; Kiefer et al., 2007a, 2007b; Patterson et al., 2007; Mahon and Caramazza, 2008; Pulvermüller et al., 2010). For a discussion of the explanatory power of such hybrid models, see the final section of this article.

3.2. Local — distributed dimension

As previously discussed, much recent evidence supports the view that concepts are established by large cell assemblies distributed over a range of cortical areas. Concepts would therefore be distributed at a neural level. In contrast, evidence supporting a local representation of concepts is difficult to find for practical reasons, especially if localist representations are understood as single neuronal units (Barlow, 1972). However, even such “grandmother cell approach” finds supporting data in recent work. Studies using single cell recordings in patients with electrodes implanted for medical reasons found neurons responding in a highly specific manner to single objects, faces, words or persons, suggesting a sparse or even localist coding by “grandmother cells” (for a review see Bowers, 2009). Neurons in the lateral temporal lobe preferentially responded to single words (Creutzfeldt et al., 1989) although it is not clear in this study whether neural responses were driven by perceptual or conceptual variables. Similarly, neurons in the medial temporal lobe were selectively activated by strikingly different pictures of a given person, landmark or object and in some cases even by person names (Quiroga et al., 2005): for instance, one neuron responded not only to a range of quite different images of Halle Barry’s face, but also to her written name, while remaining silent to other stimulus categories. At least the “Halle Barry cell” appears to be stimulated by fairly abstract memory representations rather than perceptual variables. Still, the occasional reports on grandmother cell-like neuronal activity are, of course, subject to the caveat that the specificity or the individual cells’ response patterns to stimulus types can only be compared to a relatively small number of control stimuli and comparison cells, so that the bold statement of absolute specificity can never be supported convincingly. It should also be noted that localist representations do not necessarily imply a one-neuron-one-concept correspondence, but may be realized by larger neuron populations (Bowers, 2009). For instance, similar to localist representations, a cell assembly acts as one single functional unit (Carpagnani et al., 2008; Wennekers et al., 2006): it has an activation threshold, and “ignites” as a whole when this threshold is reached. The concept of a Hebbian cell assembly therefore allows integrating critical features of localist representations within distributed neuronal networks thereby avoiding many of the problems of the one-neuron-one-concept assumption.

While there is limited evidence for local conceptual representations at the neural level, particularly in medial temporal lobe structures, there are, in contrast, a range of empirical and theoretical arguments for distributed
representations that refute purely localist theories (cf., Collins and Loftus, 1975; Collins and Quillian, 1969; Quillian, 1969; Bowers, 2009; Anderson, 1983). Firstly, the pattern of deficits produced by the widespread probabilistic loss of neurons in neurogenerative diseases such as Semantic Dementia and Alzheimer disease, which are not caused by focal lesions, can be much better accounted for by distributed than local representations at both the functional and neuroanatomical level. For instance, in Semantic Dementia Knowledge of a single concept is not entirely impaired, as a model assuming localist representations would predict, but there is a progressive deterioration from specific properties of an object concept (e.g., canary can sing) to general properties shared by many exemplars of a larger conceptual category (e.g., canary can move) during the course of the disease (Rogers et al., 2004).

Similarly, coarse superordinate categorial information for an object concept is typically relatively preserved in Semantic Dementia (animal) while more specific conceptual differentiations (bird or chicken) are more severely impaired (Hodges et al., 1995).

In Alzheimer disease, semantic deficits are sometimes category-specific (Humphreys and Forde, 2001). When many neurons are damaged during later stages of the disease, semantic impairments are typically greater for natural objects (e.g., animals, plants) than for artifact objects (e.g., tools, furniture). This category-specific impairment in Alzheimer disease can be explained by the greater intercorrelation of conceptual features for natural objects compared with artifact objects. Natural objects are more strongly impaired than artifact objects because the large amount of features that co-occur together renders differentiation and thus identification of natural objects particularly difficult if the representation of distinguishing features is damaged. This may be sufficient to illustrate the point, although, admittedly, an explanation of cases where artifacts suffer more strongly than animals is more challenging along these lines. Secondly, many different parts of the brain contribute to conceptual tasks as shown in neurophysiological studies (e.g., Martin, 2007; Pulvermüller et al., 2009; Kiefer et al., 2009). This argument is even strengthened by the observation that the activation pattern in sensory and motor areas varies as a function of the task context (Hoenig et al., 2008). Similarly, several behavioral studies also indicate that the activation of conceptual features, which code specific meaning aspects, varies depending on the situational context (Barclay et al., 1974; Barsalou, 1982). This evidence for a context-dependent contribution of meaning aspects to a concept suggests a distributed representation at the functional and neuroanatomical level: concepts are composed of several conceptual features coded by distinct cortical cell assemblies (see also the section about conceptual flexibility below). As previously mentioned, evidence of conceptual flexibility is difficult to reconcile with the notion of concepts as non-compositional localist representational entities (for a discussion of conceptual compositionality, see Goschke and Koppelberg, 1990).

Although many findings point to a distributed representation of concepts, this does not imply that conceptual features are arbitrarily distributed across the entire cortex. Instead, evidence is accumulating that conceptual features can be localized in the brain in a predictable way: they are represented within corresponding sensory or motor areas, that are also active during the perception of a referent object or the execution of a referent action (Hauk et al., 2004; Kiefer et al., 2008). In addition to the distributed representations in the sensory and motor systems, the involvement of heteromodal association cortex in conceptual processing needs also to be considered (Chatterjee, 2010). Activation of heteromodal association cortex actually follows from massive activation in sensory-motor systems, because of the well-known convergence of sensory-motor areas on higher-order areas (Pulvermüller, 2008a; Garagnani et al., 2008; Kiefer et al., 2011). As already mentioned above, the anterior temporal lobe including medial temporal lob structures have been implicated in the integration of the distributed feature representation in a supramodal semantic space (Simmons and Barsalou, 2003; Kiefer et al., 2007a, 2007b; Patterson et al., 2007; Mahon and Caramazza, 2008; Pulvermüller et al., 2010). In addition, inferior prefrontal cortex is thought to serve as conceptual working memory by guiding and controlling conceptual retrieval through attentional mechanisms (Kiefer et al., 2005; Schnur et al., 2009; Thompson-Schill et al., 1999; Wagner et al., 2001). Similarly to temporal lobe structures, which contain neurons with highly selective responses to specific objects and persons, prefrontal cortex has also been suggested to code information in a sparse or even localist fashion (Freedman et al., 2001; McClelland et al., 1995). It is therefore possible that distributed and localist conceptual representations do coexist in different brain areas (Bowers, 2009).

3.3. **Innate – experience-dependent dimension**

The debate about the existence of innate a priori categorial distinctions, which guide the acquisition of conceptual knowledge, dates back at least to the 18th century when philosophers like the rationalist Kant and the empiricist Hume controversially discussed the role of innate categories and of sensory experience for human cognition. In modern cognitive neuroscience, the notion of innate categorial distinctions has received support by the observation of category-specific conceptual deficits in brain-damaged patients who occasionally exhibit a selective impairment of semantic knowledge in some categories. Several publications described patients with a selective loss of semantic knowledge about natural (or animate/living) categories (e.g., animals, food, plants) (De Renzi and Lucchelli, 1994; Warrington and McCarthy, 1987) showing preserved knowledge about objects from artificial (or inanimate/non-living) categories (e.g., tools, furniture, kitchen utensils). The opposite pattern has also been observed, albeit less frequently (Sacchett and Humphreys, 1992; Warrington and Shallice, 1984). The category-specific impairment sometimes affects single categories such as animals (Hart and Gordon, 1992) or food items (Hillis and Caramazza, 1991) and often results in a complete loss of knowledge irrespective of the sensory or motor modality of the conceptual feature (Carbonnel et al., 1997). Therefore, it has been proposed that these impairments reflect an innate categorical organization of neural pathways for representing conceptual knowledge. These neural pathways were viewed to be shaped by evolutionary pressure to support rapid identification of objects that are...
highly relevant for survival (Caramazza and Shelton, 1998; Caramazza and Mahon, 2003).

A further argument for the notion of an innate categorial organization of conceptual knowledge comes from a case report of a selective conceptual impairment for natural objects in a 16-year-old patient whose occipito-temporal cortex was damaged at the age of 1 day (Farah and Rabinowitz, 2003). In this patient, conceptual processing of artifact objects was comparable to control participants, whereas the deficit for natural objects was apparent for both pictorial and auditory verbal stimuli and for retrieval of visual and non-visual object attributes. As brain damage, happening shortly after birth, impaired acquisition of adequate conceptual knowledge about natural objects until adolescence, this case was interpreted in terms of damage to innate category-specific neural circuits (Farah and Rabinowitz, 2003; Mahon et al., 2009). However, this interpretation is not compulsory, because this patient, in addition to his conceptual impairment, showed substantial visual field deficits along with visual object and face agnosia consistent with the widespread lesion in occipito-temporal cortex. It is known that visual recognition of natural objects is more demanding than that of artifact objects because many natural objects (e.g., different animal exemplars) have a similar overall shape rendering perceptual differentiation difficult (Lloyd-Jones and Humphreys, 1997). It is also well established that the visual channel is the most dominant sensory channel for acquiring concepts of natural objects, whereas concept acquisition of artifact objects heavily relies on action-related information (McRae and Cree, 2002; Tranel et al., 1997a, 1997b). An early visual impairment therefore more likely affects acquisition of conceptual knowledge for natural objects than for artifacts. This disadvantage for natural objects due to an early visual deficit can generalize to object properties from other sensory or motor channels because in natural objects access to concepts strongly depends on the retrieval of visual conceptual features from visual association cortex even when other conceptual features (e.g., action-related ones) are task-relevant (Hoenig et al., 2008). Thus, this case of a category-specific impairment for natural objects after perinatal brain damage presumably reflects the influence of deficient visual sensory experience on concept acquisition rather than an innate categorial specialization.

In a further attempt to provide evidence for an innate categorial organization of conceptual knowledge, category-specific activity to natural and artifact objects was assessed in congenitally blind participants with fMRI (Mahon et al., 2009). Previous studies in sighted participants indicated a categorial specialization within the ventral visual stream (for a review, see Martin, 2007): Medial regions on the ventral surface of the ventral stream (the medial fusiform gyrus, lingual gyrus, and parahippocampal cortex) showed greater activity to artifacts, whereas lateral regions on the ventral surface of the ventral stream (the lateral fusiform gyrus, inferior temporal gyrus) showed greater activity to natural objects. This medial-to-lateral organization of the ventral stream with respect to the processing of natural and artifact objects was replicated in a sample of congenitally blind participants during a size judgment task with verbally presented object names. This result demonstrates that the category-specific organization of the ventral visual pathway does not require visual experience in order to develop. However, given that the brain of congenitally blind participants has been subject to considerable plastic changes, in which visual areas are known to be activated by tactile information (Sadato et al., 1996; Roder et al., 1997) and even by shape imagery (De Volder et al., 2001), this study does not necessarily suggest the operation of innately determined domain-specific constraints on the organization of object knowledge. It is possible instead that tactile exploration of natural and artifact objects recruits "visual" areas in the course of cortical reorganization and leads to a similar extraction of category-specific object properties (e.g., regarding shape or other volumetric properties) as visual exploration. In fact, the size judgment task used in this experiment requires access to volumetric object properties, which can be comparably derived from vision and touch (De Volder et al., 2001). Finally, the authors do not take account of the fact that a major difference between animals and tools lies in their action relevance and action effects. Tools are typically used in motor actions, which in turn produce characteristic sounds (e.g., the hammering noise when using a hammer). Visually impaired people are not excluded from such activity, allowing them to acquire differential knowledge for animals and tools. Any category difference in the organization of temporal cortex could be driven by action or sound representations and therefore be related to differential action or sound knowledge about tools and animals acquired by blind individuals. Note that different action word categories not only activated motor cortex in a somatotopic manner, but also evoked differential activation in temporal cortex, which can be explained by fronto-temporal connections (Pulvermüller et al., 2009). Furthermore, sound stimulation has been shown to sharpen category-selective responses in the ventral visual stream (Adam and Noppeney, 2010). Thus, category-specific differences in congenitally blind participants can be driven by experience, specifically by touch, action and sound, and do not necessarily reflect innate categorial specialization. Organization of these higher-level visual areas can be influenced by long-distance connections converging on ventral temporal cortex (Adam and Noppeney, 2010; Noppeney et al., 2006).

While evidence for an innate categorial specialization of neural circuits involved in concept representation is limited, there are convincing demonstrations that concepts are represented in sensory and motor areas through the learning-based formation of cortical cell assembles. Three training studies investigated the experience-dependent acquisition of conceptual representations for novel objects. In the first study (James and Gauthier, 2003), participants learned associations between novel objects ("greebles") and verbal labels of object features referring to a given modality (auditory and object motion). In a sequential matching task at test, the authors found stronger activity to objects associated with auditory words ("buzzes") in the superior temporal gyrus, which responded to sounds in general. Activity in the superior posterior temporal sulcus, which was sensitive to motion processing, was greater for objects associated with motion words ("hops"). In the second study (Weisberg et al., 2007) the learning of tool-like functions for novel objects was assessed.
During training, participants had the opportunity to interact with the objects. In a sequential matching task with photographs of the objects, activity in brain regions (left middle temporal gyrus, left intraparietal sulcus and premotor area) previously implicated in the processing of manipulable objects such as tools (Kellenbach et al., 2003; Noppeney et al., 2006) was increased after training in comparison to a pre-training baseline. In the third study (Kiefer et al., 2007b), plasticity of conceptual representations was investigated by training human participants with novel objects (“nobjects”) under different training conditions (Fig. 4A). During training, participants were asked to either make an action pantomime towards the detail feature of the novel object or to point to it. In a categorization task at test, the neural correlates of the acquired conceptual representations were assessed by measuring electrical brain activity. Only in the pantomime group, in which a meaningful action was performed towards the object during training, early activation in frontal motor regions and later activation in occipito-parietal visual-motor regions was found (Fig. 4B). In the pointing training group, in which the action during training was not meaningfully related to the object, these effects were absent. These results show that action information contributes to conceptual processing depending on the specific learning experience.

Experience-dependent plasticity of conceptual representations can not only be demonstrated with novel, but also with real objects. Professional musicians constitute a model par excellence for understanding plasticity in the human brain, particularly in the auditory domain. It was therefore determined whether orchestra musicians’ intensive sensory-motor experiences with musical instruments shape their conceptual representations of this object class (Hoening et al., 2011). Using a picture-word matching task with musical instruments and matched control objects as stimuli, it was shown that visual recognition of musical instruments automatically activates right pSTG/MTG only in professional musicians, but not in musical laypersons. These areas in auditory association cortex were not only recruited by the conceptual processing of musical instruments during visual object recognition, but also by the auditory perception of real sounds. This finding unequivocally demonstrates experience-driven neuroplasticity of conceptual representations of musical instruments embodied in auditory cortex.

In conclusion, several lines of evidence clearly indicate experience-dependent formation of cell assemblies in sensory and motor areas, which code conceptual features in a modality-specific fashion. Findings supporting the notion of an innate categorial specialization of the conceptual system are inconclusive because sensory or motor experience as the driving force for conceptual organization most likely influenced the results in all the described empirical reports.

Fig. 4 – Experience-dependent plasticity of conceptual representations in the visuo-motor system: (A) Examples of the novel 3D objects (“nobjects”), for which conceptual knowledge was acquired during training and demonstrations of the different training conditions. Top: actions associated with the detail feature in the pantomime-training group. Bottom: pictures highlighting the detail feature to be pointed to in the pointing training group. (B) Minimum norm source estimates obtained from grand-mean scalp ERPs at test after the training during a categorization task. Shown are difference maps between the pantomime and the pointing condition at 117 msec after picture onset. Note the fronto-central activation (yellow circle) close to the premotor cortex in the pantomime group. Modified after Kiefer et al. (2007b).
3.4. Stable–flexible dimension

Concepts are traditionally assumed to be stable mental knowledge entities that are retrieved from memory in a situational invariant fashion. By conceptual invariance or stability, we mean here that, if a sign, picture or word is linked to a concept, the evoked concept will always be the same, thus implying that the meaning of “cow” or “game” does not change with contexts. Although entirely lacking direct empirical support, conceptual stability is commonly indirectly implicated in the basic premises that underlie the investigation of the neurocognitive architecture of conceptual knowledge. For instance, category-related imaging findings (i.e., activation differences between conceptual categories such as animals and tools) are generally requested to be task-independent in order to be considered as valid index for conceptual processing. Artifact objects like tools should consistently activate motor regions, and natural objects like animals should consistently activate visual regions, irrespectively and independently of the task (Devlin et al., 2002). Consequently, observations of task-dependent modulations of category effects are either explained by task-specific processes (instead of conceptual effects) or they are considered as spurious epiphenomena (Gerlach, 2007). However, the theoretical basis of the request to find main effects of category and task in the absence of any interaction between these factors critically depends on the tacit assumption of conceptual stability. This theoretical stance includes the notion that processing of a particular concept is performed by an invariant pattern of activated brain areas irrespective of task demands.

This tacit assumption is problematic already on theoretical grounds. As already mentioned above, the conceptual representations associated with a given word presented in different contexts are best characterized in terms of family resemblance as the search for core meanings for many words either leads to very unspecific conceptual feature sets or to no specification at all (Wittgenstein, 1953). Take again the example of the concept “game”, where the context determines whether a running-after-ball type competitive human interaction or rather a solitary thoughtful activity is meant. The relation between words and concepts is even more complicated because the word “game” is semantically ambiguous as it also refers to animals being hunted and is thus connected to two different concepts.

The phenomena of family resemblance and lexical ambiguity can be explained best by theories assuming conceptual flexibility. According to this view concepts and thus word meanings are constituted of dynamically recruited features depending on the context (Barsalou, 1982; Kiefer, 2005): the activity level of features contributing to a concept differs as a function of the dominance or weighting of conceptual features and of the contextual constraints. Theoretical investigations and behavioral findings consistently indicate that the contribution of features to a concept are context-dependent: Barsalou (1982) showed that non-dominant conceptual features were in need for contextual support to be verified rapidly while the accessibility of dominant conceptual features was independent of such support (see also Barclay et al., 1974). This suggests that some object features (e.g., red) belong to the conceptual core (e.g., tomato) and are thus more constitutive for that concept (dominant features) than others (e.g., to throw). Although non-dominant conceptual features can likewise be associated with a given object they are less relevant in constituting the concept. Their activation depends more strongly on the semantic context.

At a neural level, the use of a concept in different situations is modeled as the context-specific ignition of cell assemblies, which is constrained by both established connections between neurons that constitute conceptual long-term memory traces and the variable modulatory influence of the situation, which primes different sets of neural populations (Hoening et al., 2008; Kiefer, 2005; Pulvermüller, 1999). Concepts and correspondingly word meaning as well as their neurobiological underpinnings should therefore be viewed as context-dependent.

The notion of flexible concepts was further tested in a combined fMRI and ERP study (Hoening et al., 2008). Participants performed verifications of two object attribute types (visual, action-related) for words denoting artificial and natural objects. Functional imaging predominantly revealed cross-over interactions between category and attribute type in visual, motor and motion-related brain areas indicating that access to conceptual knowledge is strongly modulated by attribute type: activation in these modality-specific brain areas was increased when the non-dominant conceptual features, which do not belong to the conceptual core and are therefore less primed (i.e., visual features for artifacts and action-related features for natural kinds) had to be verified (Fig. 5A). ERPs indicated that these category–attribute interactions emerged as early as 116 msec after stimulus onset suggesting that they reflect rapid access to conceptual features rather than post-conceptual processing (Fig. 5B). These results indicate that concepts are situation-dependent mental entities. They are comprised of semantic features which are flexibly recruited from distributed, yet localized semantic maps in modality-specific brain regions depending on contextual constraints.

4. Synthesis and future directions

In the previous section, we have reviewed evidence in light of four important dimensions to characterize conceptual representations. Based on this evidence, we conclude that concepts are modality-specific representations grounded in perception and action and realized in the brain as action-perception circuits, that is, cortical cell assemblies that also contribute to motor and sensory processes. Concepts are distributed representations, which include conceptual features coded in distinct sensory and motor areas of the human brain. The representation of concepts depends on previous experience with the referent because sensory and motor interactions with objects shape conceptual memory traces in the corresponding modality-specific brain areas. Finally, concepts are flexible mental entities whose features are activated depending on situational constraints. To summarize, converging evidence indicates that concepts are flexible, experience-dependent modality-specific representations distributed across sensory and motor systems. Although
Fig. 5 — Flexible retrieval of conceptual representations from the visual, motor and motion-related brain systems: (A) Illustration of functional interaction effects between semantic category (artifactual vs natural) and attribute type (action-related vs visual) in modality-specific (visual, motor and motion-related) brain regions of the left hemisphere. Bar charts next to each interaction cluster depict the size of effect for each of the four conditions at the peak voxel within each cluster. Functional activation in these modality-specific brain areas followed a systematic pattern as a function of the dominance of a given conceptual feature for a category: The precise pattern of these cross-over interactions differed significantly both within posterior parietal cortex and between this action-related brain region and the visual brain region in inferior posterior temporal cortex. Small vertical bars indicate the standard error of means (s.e.m.). LH = left hemisphere; IFG = inferior frontal gyrus; ITG = inferior temporal gyrus; IMTG = inferior/middle temporal gyrus; SPG = superior parietal gyrus. (B) ERP time courses. Interaction effects between semantic category and attribute type are found as early as 116 msec (150 and 200 msec) over fronto-central (occipito-parietal) scalp regions, strongly suggesting a conceptual origin of the functional interactions. The graphs depict voltages (average-referenced data; n = 20) at representative occipito-parietal (P9) and fronto-central (FC3) electrodes as a function of category and attribute type. Negativity is plotted downwards and amplitudes are in μV. Modified after Hoenig et al. (2008).
much progress has been made in elucidating the nature of concepts several issues require further clarification in future research.

4.1. Is there a conceptual hub in the temporal pole?

One open question concerns the precise role of anterior temporal areas including the temporal pole for conceptual processing (for reviews, see Simmons and Martin, 2009; Visser et al., 2010). As described above, this area is most frequently damaged in patients with Semantic Dementia resulting in a general impairment regardless of conceptual domain or feature type. According to a purely amodal account, the anterior temporal cortex is the neural substrate of an amodal conceptual representation (McClelland and Rogers, 2003; Rogers et al., 2004), which serves as unitary conceptual store representing conceptual knowledge proper. In this perspective, representations in the sensory and motor systems would only serve as peripheral input and output systems for this amodal conceptual store (Rogers et al., 2004): sensory systems provide the input to conceptual memory, and the motor system permits the expression of conceptual knowledge in behavior.

Although this proposal is convincing to some extent, as it is rooted in patient work, neuroimaging results and neural network simulations, we wish to point out alternative possibilities, which, in our views, would also be consistent with the pre-existing evidence. Conceptual knowledge proper may be grounded and critically represented in sensory and motor areas, whereas the anterior temporal cortex merely facilitates and influences conceptual processing in sensory and motor areas, but does not store conceptual knowledge per se. This view is taken by hybrid models, which combine a modality-specific framework with the notion of a central conceptual hub. Two further interpretations are conceivable, which differ slightly with regard to the modulatory role of anterior temporal areas: (i) anterior temporal cortex acts as a conceptual convergence zone (Damasio, 1989a, 1989b; Damasio and Damasio, 1994) which integrates distributed modality-specific conceptual features into a common semantic space (Pulvermüller et al., 2010; Patterson et al., 2007; Kiefer et al., 2007a, 2007b; Simmons and Barsalou, 2003). This integration might be achieved by a supramodal higher-level representation, which indicates where in sensory and motor areas features for constituting a specific concept are represented. Such supramodal higher-level representations have been proposed for explaining the role of prefrontal cortex in working memory (Kessler and Kiefer, 2005) and of the hippocampus in episodic memory functions (McClelland et al., 1995): these higher-level representations do not store the content per se, but are thought to guide the retrieval of stored information by re-instating and stabilizing activity in sensory and motor areas (Kessler and Kiefer, 2005). If the anterior temporal cortex provides such higher-level representation for conceptual memory, damage to this region would result in a general loss of conceptual knowledge as it is observed in Semantic Dementia, (ii) a further possible interpretation assumes that anterior temporal regions (along with the basal forebrain) act as a power station which is necessary for conceptual processing, but does not perform the conceptual computations (Pulvermüller and Schumann, 1994). Accordingly, together with the fronto-orbital regions, the temporal pole plays a role in connecting cortex to the subcortical parts of the limbic system, including the amygdala. In animals, removal of both temporal lobes causes the Klüver–Bucy syndrome, which includes a lack of knowledge about the meaning of objects and actions of others, which once was called “psychic blindness” (Klüver and Bucy, 1939). It has been pointed out that Semantic Dementia indeed includes symptoms of the Klüver–Bucy syndrome (Hodges et al., 1992). In particular, “psychic blindness” is reminiscent of the semantic deficit in Semantic Dementia. Results can be explained by viewing the temporal lobes either as conceptual hub or as power supply station for conceptual processing, with conceptual circuits being localized elsewhere (for the principal argument, see Hughlings Jackson; Jackson, 1958).

Irrespective of whether anterior temporal cortex serves as convergence zone or as power station, both interpretations converge on the assumption that functional activation of modality-specific areas should also be necessary for conceptual processing. This prediction most clearly distinguishes our embodiment account from the amodal account, which claims that activity in sensory and motor areas during conceptual tasks is concomitant and reflects input or output processes, but does not contribute to conceptual performance directly.

4.2. Do sensory and motor representations play a causal role in conceptual processing?

Although previous behavioral and TMS studies in healthy participants as well as neuropsychological patient studies indicated a functional role of motor representations for conceptual processing (e.g., Helbig et al., 2010; Pulvermüller et al., 2005a, 2005b), more future work is needed for clarifying the causal relation between modality-specific representations and conceptual processing. In particular, more direct evidence for a causal role of sensory areas for performance in conceptual tasks is missing to date (but see, Gainotti, 2004; Gainotti et al., 1995). It would therefore be interesting to investigate whether damage to higher-level sensory areas in patients or temporal deactivation of these areas using TMS selectively impairs processing of concepts, for which a given modality is important (for a recent theory-driven approach in this direction, see Trumpp et al., 2011). Likewise the precise role of the anterior temporal lobe as a central conceptual hub could be further elucidated using patient studies or TMS (Pobric et al., 2010). A recent study (Pulvermüller et al., 2010) found fine grained differences in processing semantic word categories — such as form and color words or face and arm related action words — in the general semantic deficit characteristic of Semantic Dementia, a pattern which the authors interpret as support for a joint contribution of temporal pole and sensory-motor circuits to conceptual processing (Patterson et al., 2007).

4.3. Are abstract concepts grounded in the sensory and motor brain systems?

Most previous work has focused on the representation of object concepts or action-related concepts (action words). In
contrast to these well-investigated concrete concepts, the representation of abstract concepts, including social terms ("justice", "freedom"), scientific expressions ("gravitation", "quantum ergodicity") but also inner state terms ("desire", "pity"), is poorly understood and represents a crucial issue in the research on the nature of concepts. By definition, abstract concepts do not refer to physical objects that can be directly experienced by the senses and their action relationship is, if it exists at all, very complex. At the first glance, it is therefore hard to imagine how such concepts could be grounded in the sensory and motor brain systems. Hence, the mere existence of abstract concepts appears to falsify modality-specific theories and points to an amodal symbolic representation.

In fact, past and current research is almost exclusively dominated by the view that abstract concepts depend on amodal, verbal-symbolic representations. This theoretical stance dates back to the influential work by Paivio on his Dual Code Theory (Paivio, 1986). Paivio (1986) assumed that abstract concepts are stored in a verbal-symbolic code within left hemisphere regions. Concrete concepts are assumed to rely in addition to a visual imaginary code that involves right hemisphere regions. Similarly, one could suggest embodied sensory-motor representations for concrete object and action concepts, but add a symbolic semantic processor that handles all the concepts, including the tough ones with abstract meaning (Mahon and Caramazza, 2008). Although this proposal may appear as a nice compromise, this “solution” does not provide empirically informed assumptions about the cognitive mechanisms and the neuronal circuitry involved in representing abstract concepts. Admittedly, some abstract concepts are still a challenge for modality-specific theories and a priori knowledge seems to be needed to represent aspects thereof, but reverting to a non-mechanistic position allowing abstraction to appear as a *deus ex machina* in an abstract symbolic cognitive system does not appeal as a convincing solution.

Previous research on abstract concepts was restricted to proving the validity of Paivio’s Dual Code Theory. Earlier studies aimed at showing differences in the processing between abstract and concrete concepts at a behavioral and neural level. (i) There is a processing advantage for words referring to concrete concepts in comparison to words referring to abstract concepts: concrete words are remembered better (Marshark and Paivio, 1977) and recognized faster than abstract words (James, 1975). This so-called “concreteness effect” is consistent with the idea of a dual coding for concrete concepts (visual and verbal-symbolic codes) and the exclusively verbal-symbolic coding of abstract concepts. This concreteness effect, however, is diminished when target words were presented within a sentence context (Schwanenflugel and Akin, 1994). It has therefore been argued that concrete and abstract words only differ with regard to the availability of contextual information, e.g., situations in which the concept is encountered, but do not differ with regard to representational codes.

(ii) Words with abstract meaning were seen to elicit relatively stronger left-lateralized brain responses whereas concrete words were associated with more symmetric hemispheric responses (Neville et al., 1992) or even greater right hemisphere activation (Binder et al., 2005) as predicted by the Dual Coding Theory (for a review, see Pulvermüller, 2007). However, a range of different neurophysiological and neuro-imaging studies revealed quite diverging results on the “cortical seat” of abstract concepts. Some studies found left hemisphere activity for both word classes (Sabsevitz et al., 2005) or even greater right hemisphere involvement for abstract words (Kiehl et al., 1999). Such variability of results can falsify any theory about a singular symbolic center treating abstract terms.

As previous attempts elucidating the nature of abstract concepts were not successful, we believe that a new approach in addressing these issues is necessary and propose a modality-specific approach as an integrative conceptual theory. We assume that abstract and concrete concepts are similarly grounded in perception and action (Kiefer and Barsalou, 2011; Barsalou and Wiemer-Hastings, 2005; Pulvermüller, 2008b). Although abstract concepts may more strongly depend on lexical associations between words as well as on emotional and introspective representations in addition to sensory and motor representations than concrete concepts (Barsalou et al., 2008; Vigliocco et al., 2009), the dichotomy between these types of concepts does not appear justified: even abstract concepts are typically embedded into concrete situations which can be experienced and constitute the content of the concept. Our approach to the processing of abstract concepts is therefore in fact very similar to the one offered for concrete concepts. In contrast to a concrete term, which has one typical referent (“sun”) or a small set of typical ones, an abstract term such as “beautiful” can be viewed as linked to a set of different possible instantiations of the concept. In the visual domain, such instantiations may include a beautiful sunset as it may cover a beautiful person, an object or event.

At the neural level, the embodied perceptual representations of all of these possible instantiations may in fact play a role, and a general neural mechanism that connects these instantiations in a modality-specific schema-like representation may apply. Abstract concepts have presumably to be embedded in a conceptual context that selects one of the competing instantiations and complements missing information in order to reach an unequivocal interpretation (Schwanenflugel and Akin, 1994). For action-related abstract concepts such as “to free” – which can refer to a wide range of basic actions associated with freeing such as the removal of constraints of some sort – action representations related to these different kinds of freeing actions are obviously most crucial and should recruit motor circuits. Complementing sensory-motor representations, abstract concepts such as “to free”, but also “truth” and “relationship” are typically strongly associated with emotions and may also include introspective information about internal states experienced in corresponding situations (e.g., in a situation, in which an individual felt freed in the past). Hence, abstract concepts may also depend on brain circuits typically involved in emotion (Rolls, 1999) and introspection (Van Overwalle and Baetens, 2009). The relative contribution of sensory-motor, emotional, introspective and lexical representations to an abstract concept may highly variable across different abstract concepts and situations, in which these concepts are instantiated. The beauty of the modality-specific account is that it explains
representation of abstract concepts mechanistically with reference to known cortical structure and function.

At present, evidence regarding a modality-specific grounding of abstract concepts in sensory-motor, emotional and introspective brain circuits is scarce. Moderately abstract form and color related terms activated frontal (close to pre-motor cortex) and anterior temporal areas (close to visual cortex) differentially suggesting a grounding in action and perception (Pulvermüller and Hauk, 2006). For every-day abstract concepts, preliminary data suggest that participants refer to concrete situations when thinking of the content of abstract concepts (Barsalou and Wiemer-Hastings, 2005). For instance, the word “freedom” may apply to a situation, in which a teenager spends her or his first vacation without parents and enjoys doing what she or he wants to do. Within this context, the concept “freedom” is defined by sensory and action properties of the vacation situation as well as by emotional states of the actor derived from introspection. Recent neuroimaging studies confirmed this observation and found activation in sensory-motor as well as in emotional brain regions during the processing of abstract concepts (Pexman et al., 2007; Wilson-Mendenhall et al., 2011).

To further test the embodiment view of abstract concepts, the development of new experimental paradigms is needed, which are suited to demonstrate a possible involvement of the perception, action and emotional systems in the representation of abstract concepts. Research on the grounding of abstract concepts in perception, action, emotion and introspection would open a novel promising field, which helps to establish a coherent concept. Another, but related issue concerns the causal role of sensory and motor representations in conceptual processing. Convergent results indicate that concepts are flexible mental entities that are constituted by distributed representational features. Concepts are embodied in the sense that their conceptual features are represented in sensory and motor brain areas in an experience-dependent fashion. Although many studies support a modality-specific view on conceptual representations, we have identified three issues that deserve further clarification in future studies. It has to be determined whether anterior temporal lobe structures, which are damaged in Semantic Dementia, code conceptual knowledge proper in an amodal representation format. Alternatively, these structures may serve as central hub, or merely facilitate the integration of the distributed sensory and motor representations into a coherent concept. Another, but related issue concerns the causal role of sensory and motor representation for conceptual processing, which calls for more future work. Only embodiment accounts of conceptual representation, but not purely amodal accounts assume that modality-specific representations are essential for performing a conceptual task. Finally, the probably most crucial issue to be resolved in future research is the representation of abstract concepts. While the embodiment theory grounds abstract concepts in perception, action and emotion through their reference to concrete situations that can be experienced, the representation of abstract concepts in the sensory and motor systems of the brain has to be further elucidated for a broad range of concepts. This future research will decide whether the embodiment theory of conceptual representations can serve as an integrated framework for both concrete and abstract concepts.

4.4. Conclusions

Evidence from behavioral, neuroimaging and neuropsychological studies highlights the important role of sensory and motor representations in conceptual processing. Convergent results indicate that concepts are flexible mental entities that are constituted by distributed representational features. Concepts are embodied in the sense that their conceptual features are represented in sensory and motor brain areas in an experience-dependent fashion. Although many studies support a modality-specific view on conceptual representations, we have identified three issues that deserve further clarification in future studies. It has to be determined whether anterior temporal lobe structures, which are damaged in Semantic Dementia, code conceptual knowledge proper in an amodal representation format. Alternatively, these structures may serve as central hub, or merely facilitate the integration of the distributed sensory and motor representations into a coherent concept. Another, but related issue concerns the causal role of sensory and motor representation for conceptual processing, which calls for more future work. Only embodiment accounts of conceptual representation, but not purely amodal accounts assume that modality-specific representations are essential for performing a conceptual task. Finally, the probably most crucial issue to be resolved in future research is the representation of abstract concepts. While the embodiment theory grounds abstract concepts in perception, action and emotion through their reference to concrete situations that can be experienced, the representation of abstract concepts in the sensory and motor systems of the brain has to be further elucidated for a broad range of concepts. This future research will decide whether the embodiment theory of conceptual representations can serve as an integrated framework for both concrete and abstract concepts.

Acknowledgments

This research was supported by a grant of the German Research Community (DFG 804/5-1) to MK. The authors thank two anonymous reviewers for their helpful comments on an earlier draft of this manuscript.

References


Pulvermüller F, Cooper-Pye E, Dine C, Hauk O, Nestor P, and Patterson K. The word processing deficit in semantic dementia: All categories are equal but some categories are more equal than others. Journal of Cognitive Neuroscience, 22(9): 2027–2041, 2010.


