Are abstract action words embodied? An fMRI investigation at the interface between language and motor cognition

Katrin Sakreida[1]*, Claudia Scorolli[2], Mareike M. Menz[3], Stefan Heim[1,4,5], Anna M. Borghi[2,6] and Ferdinand Binkofski[1,7,8]

[1] Division of Clinical and Cognitive Neurosciences, Department of Neurology, Medical School, RWTH Aachen University, Aachen, Germany
[2] Department of Psychology, University of Bologna, Bologna, Italy
[3] Department of Systems Neuroscience and Neuroimage Nord, University Medical Center Hamburg Eppendorf, Hamburg, Germany
[4] Section Structural Functional Brain Mapping, Department of Psychiatry, Psychotherapy and Psychosomatics, Medical School, RWTH Aachen University, Aachen, Germany
[5] Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich, Germany
[6] Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy
[7] Institute of Neuroscience and Medicine (INM-4), Research Centre Jülich, Jülich, Germany
[8] Translational Brain Medicine, Jülich Aachen Research Alliance, Jülich, Germany

The cognitive and neural representation of abstract words is still an open question for theories of embodied cognition. Generally, it is proposed that abstract words are grounded in the activation of sensorimotor or at least experiential properties, exactly as concrete words. Further behavioral theories propose multiple representations evoked by abstract and concrete words. We conducted a functional magnetic resonance imaging (fMRI) study to investigate the neural correlates of concrete and abstract multi-word expressions in an action context. Participants were required to read simple sentences which combined each concrete noun with an adequate concrete verb and an adequate abstract verb, as well as an adequate abstract noun with either kind of verbs previously used. Thus, our experimental design included a continuum from pure concreteness to mere abstractness. As expected, comprehension of both concrete and abstract language content activated the core areas of the sensorimotor neural network namely the left lateral (precentral gyrus) and medial (supplementary motor area) premotor cortex. While the purely concrete multi-word expressions elicited activations within the left inferior frontal gyrus (pars triangularis) and two foci within the left inferior parietal cortex, the purely abstract multi-word expressions were represented in the anterior part of left middle temporal gyrus that is part of the language processing system. Although the sensorimotor neural network is engaged in both concrete and abstract language contents, the present findings show that concrete multi-word processing relies more on the sensorimotor system, and abstract multi-word processing relies more on the linguistic system.

Keywords: language comprehension, abstract, concrete, fMRI, sensorimotor cortex

INTRODUCTION

Embodied and grounded cognition theories such as Theories of Situated Action, Cognitive Linguistics Theories, Cognitive and Social Simulation Theories (for a review, see Barsalou, 2008), are becoming increasingly popular in cognitive neuroscience. This approach extends to different domains (e.g., perception, action, language, decision-making etc.) and crosses different disciplines, from philosophy (e.g., Clark, 1999), developmental psychology (e.g., Smith, 2005), and social psychology (e.g., Semin and Smith, 2008), to computer science and robotics (e.g., Nolfi and Floreano, 2000; Ziemke, 2002). In contrast to the classical cognitivism that is based on representational systems of symbolic information processing, which distinguishes between so-called high and low cognitive processes, embodied views propose that high-level cognitive processes, such as language comprehension, are grounded in the lower-level processes of perception and action. A wide range of publications within the last decade demonstrates this interest in embodied cognition (for an analysis, see Chatterjee, 2010; Gentner, 2010; Jirak et al., 2010).

When embodied cognition approaches are applied to language comprehension, the notion of "simulation" becomes a prominent feature (e.g., Gallese, 2008). Here, simulation refers to the process of internally representing (or simulating) the content that a word or sentence describes. Thus, the simulation process involves the same sensorimotor neural correlates as during the action execution or when interacting with the actual object or entity itself (Zwaan, 2004). Behavioral and neural evidence has reliably shown that the process of language comprehension elicits activations within primary and secondary motor areas, thus prompting an explanation in terms of embodied simulation (for reviews, see Pulvermüller, 2005; Barsalou, 2008; Fischer and Zwaan, 2008; Toni et al., 2008).

Recently, the limitations of embodied motor simulation have been examined by studies using transcranial magnetic stimulation (TMS). Papeo et al. (2011) showed that enhanced TMS-induced motor-evoked potentials do reflect motor simulation, but that these are restricted to the experimental condition when hand-action verbs were presented in first person, i.e., when the self was recruited as agent, compared to third person verbs and non-action verbs. Moreover, a comparable limb-specific effect for processing of hand-action verbs was found when TMS was applied at 500 ms post-stimulus. This finding indicates that the activity of primary motor cortex was involved in post-conceptual processing, which follows the retrieval of motor representations, rather than in initial lexical-semantic processing (Papeo et al., 2009). As such, the extent to which language comprehension is actually embodied is still the focus of intense debate.

In opposition to a strong embodied approach, some authors propose that sensorimotor system activation during language processing is not necessary for comprehension, since this occurs after the context and the meaning of the information has been computed. Hence, those authors suggest a dynamical interaction among the multimodal modules of language, perception, and action (Mahon and Caramazza, 2008; Bedny and Caramazza, 2011) or gradations from embodied to disembodied cognition (Chatterjee, 2010).

Beyond the discussion on embodiment of language comprehension in general, the debate focuses on disentangling concrete vs. abstract word representations. Hence, abstract word semantics constitute a specific challenge for embodied cognition theories (for a recent review, see Pecher et al., 2011). Embodied representations of abstract words are proposed to underlie activation of sensorimotor, or at least experiential properties, exactly as concrete words. In support of this, Glenberg et al. (2008) used combined behavioral and TMS data to demonstrate that abstract transfer sentences (e.g., "to give some news") activate motor areas in the same way as concrete transfer sentences (e.g., "to give a pizza"; see also Glenberg and Kaschak, 2002). Further evidence was provided by Barsalou and Wiemer-Hastings (2005) who showed that abstract concepts focus rather on settings and events as well as introspective states than on purely perceptual properties. Further, Kousta et al. (2009, 2011) have demonstrated that abstract as compared to concrete words involve more emotional aspects.

An additional proposal is that multiple representations are evoked by words (for a non-embodied version of this view see Dove, 2009). According to the *Language And Situated Simulation* (LASS) theory, the left-hemispheric language areas are mainly involved during superficial linguistic processing. This consists of word recognition and the immediately subsequent generation of associated word forms (Barsalou et al., 2008). These associated words in turn provide a linguistic context that can be sufficient to perform a wide variety of tasks, such as lexical decision-making tasks. Nevertheless, these superficial strategies may prevent deeper conceptual processing. The conceptual content of properties and relations reside in associated simulations (Barsalou et al., 2008) involving bilateral perceptual and motor neural networks. However, these two systems are not modular,

rather they interact in a continuous way. Differently from LASS, the *Words As Tools* (WAT) theory suggests that, in simulation, the linguistic form representation is not superficial and does not prevent deeper conceptual processing. According to WAT words can be conceived as tools that are useful in interacting with the world. During language comprehension a combination of both linguistic and non-linguistic sensorimotor experiences is early on activated and weighted depending not just on the task but also on the kind of considered words (Borghi and Cimatti, 2009, 2010). In fact, the WAT proposal differs from the LASS theory as the former ascribes more relevance to different lexical categories within language, e.g., concrete vs. abstract words, whereas the latter focuses more on the different levels of language processing required for the task, e.g., lexical decision vs. conceptual task.

Both LASS and WAT are in line with the *Dual Coding* theory. This approach ascribes the effect according to which concrete words are memorized better than abstract words to the existence of both a linguistic and a sensorimotor imagery code. Both codes would be activated by concrete words, whereas processing of mere verbal information would be necessary for encoding of abstract words (Paivio, 1971, 1986). Recent functional magnetic resonance imaging (fMRI) studies (e.g., Binder et al., 2005; for a review see Sabsevitz et al., 2005) endorse the *Dual Coding* assumption by showing an activation pattern that confines representation of abstract words to the left hemisphere, whereas it is bilateral for processing of concrete words (for contrasting evidence, see Rodríguez-Ferreiro et al., 2011). Moreover, Desai et al. (2010) found pronounced left-hemispheric superior temporal (BA 22) and inferior frontal (BA 44/45/47) areas activated while processing of abstract sentences (e.g., "use the opportunity"), thus, suggesting that abstract words primarily activate and are understood through verbal associations with other words. However, the embodied multiple representations proposals LASS and WAT extend the Dual Coding theory insofar as both linguistic and sensorimotor information are crucial for not just concrete words but also abstract words.

Since concrete and abstract words rely on different acquisition mechanisms (Borghi and Cimatti, 2010; Borghi et al., 2011) linguistic experience with its social aspects is more important for the acquisition of abstract rather than concrete words, given that abstract words refer to more sparse and diverse experiences than concrete ones. Thus, in line with the WAT proposal it can be assumed that the neural language network predominantly supports processing of abstract words, while concrete words are embedded mainly within the sensorimotor neural network. Evidence supporting an assumption of distributed semantic networks was recently provided by several lesion studies (Mårtensson et al., 2011; Arévalo et al., 2012; Kemmerer et al., 2012).

A number of behavioral (e.g., Day, 1979; Chiarello et al., 1987; Deloche et al., 1987), electrophysiological (e.g., Holcomb et al., 1999; Kellenbach et al., 2002; Nittono et al., 2002), and functional imaging (e.g., Kiehl et al., 1999; Perani et al., 1999; Friederici et al., 2000; Grossman et al., 2002; Noppeney and Price, 2004; Binder et al., 2005; Desai et al., 2010) studies have investigated linguistic abstractness, but the majority of these focused only on single word processing. However, it is clear that human communication

consists of much more than an individual word, since words are combined in sentences and these in turn lead to the emergence of word meaning.

The aim of our study was to dissociate neural correlates of concrete and abstract multi-word expressions, focusing on natural linguistic stimuli. To this end we experimentally manipulated very simple sentences composed by a concrete vs. abstract noun and verb. To generate a novel experimental design that encompasses a continuum from pure concreteness to mere abstractness, nouns referring to graspable/non-graspable (concrete/abstract) objects or entities were combined with motor/non-motor (concrete/abstract) verbs. Thus, at one end of the spectrum, a combination of a noun referring to a graspable object with a motor verb (CC) generates a concrete meaning. At the other end, a combination of a noun referring to a non-graspable entity with a non-motor verb (AA) leads to an unambiguous abstract content. The mixed combinations (CA, AC) served to further differentiate between the role of verb and noun in abstract contents processing. Our stimuli and the experimental design were the same as those used in previous behavioral (Scorolli et al., 2011) and TMS (Scorolli et al., 2012) studies, the results of which are expanded on in the discussion.

Based on the embodied approach, our first anatomical prediction concerned activations within the sensorimotor neural network during language processing, regardless of mere concrete or abstract content. Against the background of the LASS and the WAT proposals, our second anatomical hypothesis focused on the dissociation of core areas for pure concrete and mere abstract expression: While concrete noun-verb combinations (CC) should activate pronounced sensorimotor areas, both mixed combinations (CA, AC) and abstract noun-verb combinations (AA) should elicit stronger activations within the neural language network, especially semantic processing areas with their crucial role in the representation of concept meaning.

MATERIALS AND METHODS

PARTICIPANTS

We obtained written informed consent from 25 participants (16 female, 9 male, age range 20–36 years, mean age 25.2 ± 3.6) prior to the scanning session. All participants were right-handed with a Laterality Index >0.7 (Annett, 1970) and had normal or corrected-to-normal visual acuity. Only native speakers of German participated in the study. The experimental standards were approved by the local ethics committee. Data were handled anonymously.

STIMULUS GENERATION AND STANDARDIZATION

Noun-verb combinations for fMRI stimulus generation included 96 German nouns—48 (concrete) graspable objects and 48 (abstract) non-graspable entities—and 96 German verbs—48 (concrete) motor verbs and 48 (abstract) non-motor verbs—Note that according to the German word order, the noun is presented first followed by the verb. Each noun referring to a graspable object (C), preceded by a determinative or non-determinative article, was combined with an adequate motor verb (C) as well as an adequate non-motor verb (A), and an adequate noun referring to a non-graspable entity (A) was combined with the same

verbs previously used, e.g., "einen Schmetterling malen" (to draw a butterfly), CC—"einen Schmetterling bestaunen" (to marvel at a butterfly), CA—"den Sonnenuntergang malen" (to draw the sunset), AC—"den Sonnenuntergang bestaunen" (to marvel at the sunset), AA (see **Figure 1A**). Thus, 48 quadruples of pairs were created that were formed by two nouns and two verbs each, resulting in 192 noun-verb combination stimuli. This particular paradigm encompasses a concreteness-to-abstractness continuum. Any metaphorical or idiomatic combinations, as for instance "to kick in the dugout" or "to kick the bucket" were avoided.

Twenty German students were asked to judge the familiarity of each noun-verb combination and for the degree of probability they would use it. Ratings were given by making a cross on a continuous line scale (not familiar—very familiar; not probably—very probably). Subsequently, 18 quadruples with lowest scores and highest standard deviations for both familiarity and probability of use were removed, thus, finally 30 quadruples including 120 noun-verb pairs were selected for the fMRI experiment.

Due to the peculiarity of the linguistic material, in a first step, the lexical frequency of all stimuli in both written and spoken German was assessed using the CELEX database (Baayen et al., 1996). The averages for all four stimuli types were above 400/million, i.e., in the range of high-frequency words (concrete nouns: 404/million; concrete verbs: 530/million; abstract nouns:

FIGURE 1 | Experimental design with an exemplary stimuli combination (A) and trial procedure (B). Each noun referring to a graspable object, preceded by a determinative or non-determinative article, was combined with an adequate motor verb as well as an adequate non-motor verb, and an adequate noun referring to a non-graspable entity was combined with the same verbs previously used, e.g., "einen Schmetterling malen" (to draw a butterfly), CC—"einen Schmetterling bestaunen" (to marvel at a butterfly), CA—"den Sonnenuntergang malen" (to draw the sunset), AC—"den Sonnenuntergang bestaunen" (to marvel at the sunset), AA. Note that according to the German word order the noun is presented first followed by the verb. Due to the variable onset delay (jitter) the trial duration was 2500 ms at minimum to 4500 ms at maximum.

587/million; abstract verbs: 3132/million). The relatively high frequency of abstract verbs resulted from the item "haben" (to have). Scores were submitted to a 2 (concrete vs. abstract) × 2 (noun vs. verb) ANOVA. Analysis yielded no significant effects [no main effect concrete vs. abstract: $F_{(1, 29)} = 1.87$; $p = 0.182$, no main effect noun vs. verb: $F_{(1, 29)} = 1.83$; $p = 0.186$, no interaction: $F_{(1, 29)} = 1.52$; $p = 0.228$]. In a second step, the frequency of each noun-verb combination in written German was assessed by using the search engine "Google" with each multi-word expression within quotations marks as search terms (updated at March 2012). The 2 (kind of noun: concrete vs. abstract) × 2 (kind of verb: concrete vs. abstract) ANOVA did not show any significant difference across the four conditions [no main effect concrete vs. abstract noun: $F_{(1, 29)} = 0.09$; $p = 0.763$, no main effect concrete vs. abstract verb: $F_{(1, 29)} = 0.96$; $p = 0.335$, no interaction: $F_{(1, 29)} = 2.01$; $p = 0.167$].

Furthermore, in an additional study the linguistic material was standardized for imageability, literality, quantity of motion as well as for age of acquisition. Methods and results are reported in the Supplementary Material to this article and discussed in detail by Scorolli et al. (2011).

EXPERIMENTAL PARADIGM AND TASK

Task instructions were implemented in a go-nogo paradigm. Participants were asked to carefully read the 30 visually presented critical quadruples that demanded no motor response. To sustain attention, a button press was required toward oddball multi-word expressions that were 30 combinations of foot-related nouns with foot-related motor verbs, e.g., "einen Ball schießen," (to kick a ball). Hence, participants were instructed to press a button as fast as possible if the read sentence referred to an action typically performed with the foot and/or leg. Responses were collected with a custom-made four-buttons response-box.

The experimental design included 150 stimuli which were visually presented as white writing on a black background using VisuaStim VGA goggles (Resonance Technology Inc., Northridge, CA, USA) in a pseudo-randomized order. There were five different conditions: noun referring to a graspable object/motor verb (CC), noun referring to a graspable object/non-motor verb (CA), noun referring to a non-graspable entity/motor verb (AC), noun referring to a non-graspable entity/non-motor verb (AA), and oddball condition. Finally, experimental stimuli were supplemented by 15 empty trials used as a low-level baseline (rest condition).

Each trial started with the presentation of a noun and verb simultaneously for a duration of 800 ms, followed by a fixation cross for a duration of 1700 ms. A variable onset delay (jitter) of every stimulus in relation to the acquisition time (0, 500, 1000, 1500, or 2000 ms) produced an oversampling of the actual image acquisition time of 2500 ms by a factor of five, consequently leading to an acquisition sampling rate of 500 ms. Thus, the trial duration was 2500 ms at minimum to 4500 ms at maximum (see **Figure 1B**).

FUNCTIONAL LOCALIZER TASK

Subsequent to the main experimental task we ran a finger tapping task in 2 × 4 blocks, which is known to produce robust activation of motor areas (Moritz et al., 2000). Stimuli were green squares presented for 150 ms with either rhythmic or regular intervals (250–1010 ms). Each block started with 23.4 s stimulus presentation, followed by 23.4 s without any visual stimulation ending with a temporary presented yellow square. Participants were instructed to tap a button with their right index finger as accurate as possible synchronous to the duration of the visual pacing (green squares) and to continue tapping throughout the following second unpaced period until the yellow square appeared. Blocks were separated by a 23.4 s rest period.

DATA ACQUISITION

Imaging was performed at 3 T on a Philips magnetic resonance imaging scanner equipped with an 8-channel head coil (Philips). A fast single-shot echo-planar imaging (EPI)-sequence (echo time 30 ms, 90° flip angle, repetition time = 2500 ms) sensitive to blood oxygenation level-dependent (BOLD) contrast was used for acquiring 35 axial slices (240 mm field of view, 80 × 80 pixel matrix, 3 mm thickness, 10% spacing) covering the whole brain. Two functional runs with 330 (localizer task) and 230 (experimental task) T2* scans were performed, with each scan sampling over the 35 slices. The first five volumes of each subject's scan were removed to allow for full T2 saturation. Subsequently, a set of anatomical T1-images (240 mm field of view, 240 × 240 pixel matrix, 164 slices, 1 mm thickness, no gap, echo time = 3.7 ms, repetition time = 8100 ms) was acquired.

DATA ANALYSIS

fMRI-data were analyzed using the Statistical Parametric Mapping software SPM8 (Wellcome Department of Cognitive Neurosciences, London, UK) running under Matlab 7.10 (MathWorks Inc., Natick, MA, USA). Spatial preprocessing included realignment to the first scan, coregistration to the T1 anatomical volume images. T1-weighted images were segmented into gray and white matter. This segmentation was the basis for spatial normalization to the Montreal Neurological Institute (MNI) template, which was then resliced and smoothed with a 9 × 9 × 9 mm full width at half maximum Gaussian Kernel filter to improve the signal-to-noise ratio. To correct for low-frequency components, a temporal high-pass filter with a cut-off frequency of 1/128 Hz (=128 s) was applied.

Statistical analyses were performed using the general linear model as implemented in SPM8. In the first-level experimental task analysis for each subject onsets of picture presentation with a duration of 800 ms were used as regressors to the model including the four conditions (CC, CA, AC, and AA) as well as the oddball condition. In the functional localizer task, event related regressors to the model were the response onsets to paced and unpaced—rhythmic and regular—finger tapping.

The second-level analysis was carried out using the flexible factorial design with the first factor SUBJECT and the second factor CONDITION (CC, CA, AC, AA, Tapping). The significance level was set to $p < 0.05$, FWE corrected. Additionally, a cluster size of ≥ 5 contiguous voxels (40 mm³) extended the threshold. The SPM Anatomy toolbox v1.8 (Eickhoff et al., 2005) was employed for anatomical assignments.

RESULTS

BEHAVIORAL RESULTS

Behavioral performance was assessed by correct responses (mean percentage = 95.2%, mean percentage omissions = 4.8%) and reaction times (mean = 966 ms, standard deviation = 407 ms) to oddball multi-word expressions. Thus, as task performance was appropriate, participants' attention was directed toward comprehension of the linguistic material.

Although mean false positive rate was only 2.4% of all nogo trials, most of them occurred when the verb was a concrete one (condition CC: 58.0% of all false positive responses, condition CA: 7.4% of all false positive responses, condition AC: 25.9% of all false positive responses, condition AA: 7.4% of all false positive responses).

FUNCTIONAL IMAGING RESULTS

Whole brain analysis

The functional localizer task revealed broad activations in left primary and secondary motor cortex including lateral motor/premotor cortex and supplementary motor area, as well as subcortically in the thalamus, the putamen, and the right cerebellum as shown by contrasting the finger tapping periods minus the rest periods. Other activation clusters were located in right postcentral gyrus, right inferior frontal gyrus (pars opercularis), left middle frontal gyrus as well as bilateral in temporal areas and visual cortex (see **Figure 2**, red-colored, and **Table 1**).

Overlapping areas of the functional localizer task and the experimental conditions as revealed by a conjunction analysis (finger tapping > rest ∩ [CC + AA] > rest) are also depicted in **Figure 2**, blue color indicating experimental task activations and green color indicating overlapping areas, and listed in **Table 2**. Two left-lateralized activation clusters encompassed the lateral (precentral gyrus) and medial (supplementary motor area) premotor cortex.

As the study focused on differential neural correlates of abstract and concrete contents of language the main effect of interest was achieved by contrasting condition CC (noun referring to a graspable object/motor verb) and condition AA (noun
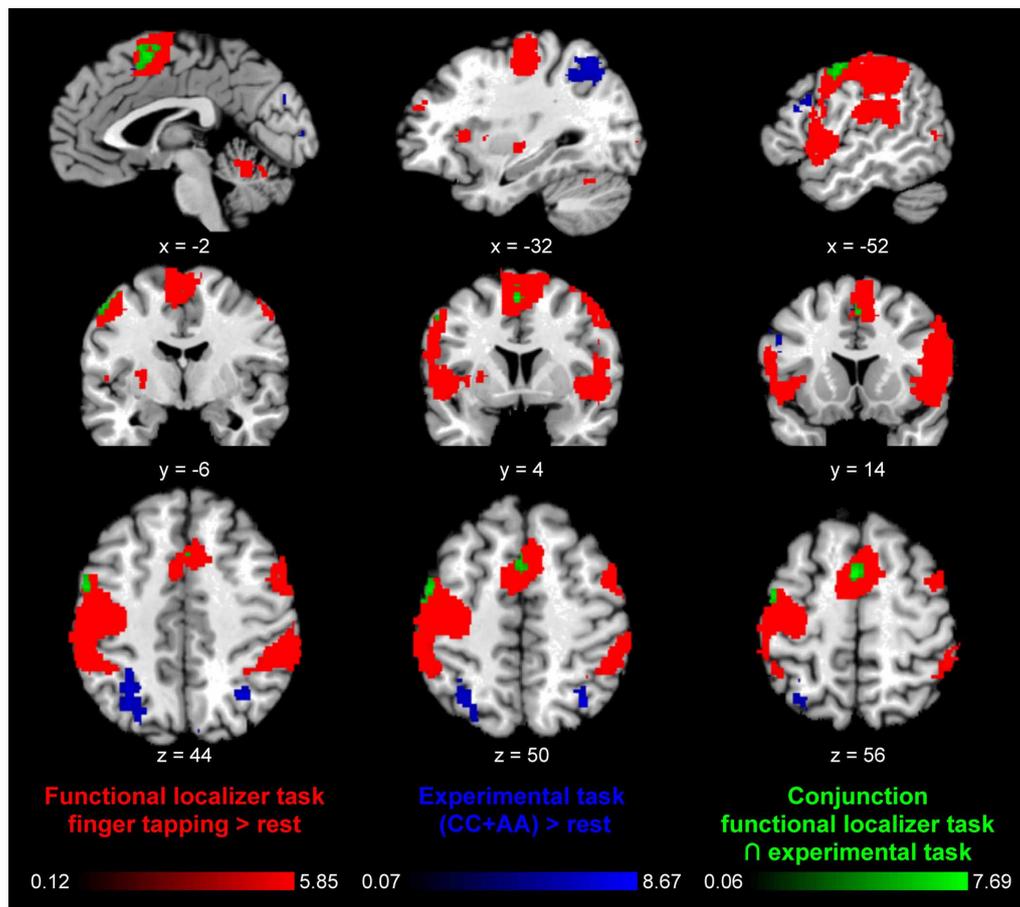
FIGURE 2 | Absolute activations resulting from functional localizer task and experimental task. Activations from finger tapping task compared to rest (red), activations from the presentation of abstract and concrete multi-word expressions (CC+AA) in the experimental task compared to rest (blue), and overlapping areas of the functional localizer task and the experimental tasks (green) as revealed by a conjunction analysis. Images were thresholded at $p < 0.05$, FWE corrected for the whole brain volume, superimposed on representative sagittal, coronal and axial slices of the MNI template using the software MRIcron Version 12/2012 (http://www.mccauslandcenter.sc.edu/ mricro/mricron/index.html).

Table 1 | Macroanatomical structure, cytoarchitectonical area (Area$_{cyto}$), percent overlap of cluster with cytoarchitectonical area, cluster size in voxel, MNI coordinates (x, y, z), and maximum T value (T$_{max}$) of the local maxima for the functional localizer task compared to rest.

Local maximum in macroanatomical structure	Area$_{cyto}$	Percent overlap of cluster with cytoarchitectonical area	Cluster size (voxel)	MNI coordinates			T$_{max}$
				x	y	z	
L. Precentral gyrus	Area 4a	3.8	5490	−42	−18	54	13.65
R. Inferior frontal gyrus (Pars opercularis)			3719	52	14	0	14.97
L. Supplementary motor area (SMA)	Area 6	44.4	1830	−4	−2	62	14.03
R. Superior temporal gyrus	IPC (PF)	28.4	1472	66	−36	16	11.40
R. Cerebellum	Lobule VI (Hem)	43.6	1157	14	−58	−24	14.12
R. Inferior occipital gyrus			363	42	−84	−6	8.00
L. Putamen			196	−26	−8	6	7.01
L. Middle frontal gyrus			125	−36	50	22	6.76
L. Middle occipital gyrus			101	−24	−90	2	7.04
R. Postcentral gyrus	OP 4	60.5	76	62	−14	20	7.21
L. Thalamus	Th-prefrontal	67.4	38	−12	−18	2	6.46
L. Middle temporal gyrus			12	−52	−68	4	5.70
R. Cerebellum			12	44	−62	−26	6.20
R. Inferior temporal gyrus			8	56	−62	−12	5.93

The significance level was set to p < 0.05, FWE corrected for the whole brain volume. A cluster size of ≥ 5 contiguous voxels (40 mm³) extended the threshold. Abbreviations: L, left; R, right.

Table 2 | Macroanatomical structure, cytoarchitectonical area (Area$_{cyto}$), percent overlap of cluster with cytoarchitectonical area, cluster size in voxel, MNI coordinates (x, y, z), and maximum T value (T$_{max}$) of the local maxima of the conjunction: functional localizer task (finger tapping > rest) ∩ experimental tasks ([CC + AA] > rest).

Local maximum in macroanatomical structure	Area$_{cyto}$	Percent overlap of cluster with cytoarchitectonical area	Cluster size (voxel)	MNI coordinates			T$_{max}$
				x	y	z	
L. Supplementary motor area (SMA)	Area 6	69.1	125	−2	4	56	7.72
L. Precentral gyrus	Area 6	93.3	85	−52	−6	50	7.27

The significance level was set to p < 0.05, FWE corrected for the whole brain volume. A cluster size of ≥ 5 contiguous voxels (40 mm³) extended the threshold. Abbreviations: L, left; R, right.

referring to a non-graspable entity/non-motor verb) and vice versa. The direct contrasts CC > AA and AA > CC ($p < 0.05$, FWE corrected for small volumes using the image masks of the SPM Anatomy toolbox v1.8 and a mask of the temporal lobe generated by the WFU PickAtlas Toolbox v3.0.4, respectively) yielded significant activation clusters within a fronto-parietal-temporal network (**Figure 3** and **Table 3**). In the contrast CC > AA the left inferior frontal gyrus (pars triangularis) and two foci within the left inferior parietal cortex were activated, whereas the reverse contrast AA > CC yielded only one suprathreshold activation cluster in the anterior part of left middle temporal gyrus.

Regions of interest analysis
The statistical comparisons of the mixed conditions with the pure conditions (CA > CC, AC > CC, CA > AA, AC > AA) showed no significant effects at the chosen threshold. However, to evaluate the effect sizes for the mixed conditions, the group-averaged contrast values of the maximally activated voxel of the frontal (CC > AA) and temporal (AA > CC) activation cluster were

statistically compared between all four conditions using repeated-measures ANOVAs and are also displayed in **Figure 3**. A significant main effect were analyzed within both regions of interest, i.e., left inferior frontal gyrus [$F_{(3, 72)} = 14.27$; $p < 0.001$] and left middle temporal gyrus [$F_{(3, 72)} = 6.77$; $p < 0.001$]. As condition CA did not significantly differ from condition AC in both of the activation peaks, both mixed conditions were combined by averaging. Thus, *post-hoc* paired T-tests were calculated for comparisons of three conditions (CC, MIX, AA) with an adjusted significance level to $p = 0.05$ (corresponding to uncorrected $p = 0.016$) by applying a Bonferroni-correction for multiple comparisons (here three). The analysis yielded a significant difference between the MIX condition vs. the AA condition in both regions of interest [left inferior frontal gyrus: $T_{(24)} = 4.64$; $p < 0.001$, left middle temporal gyrus: $T_{(24)} = 3.16$; $p = 0.004$]. The comparison of the MIX condition vs. the CC condition were not significant in the region of interest left inferior frontal gyrus [$T_{(24)} = 1.71$; $p = 0.101$], whereas the differences in contrast values between the MIX and the CC condition in the region of
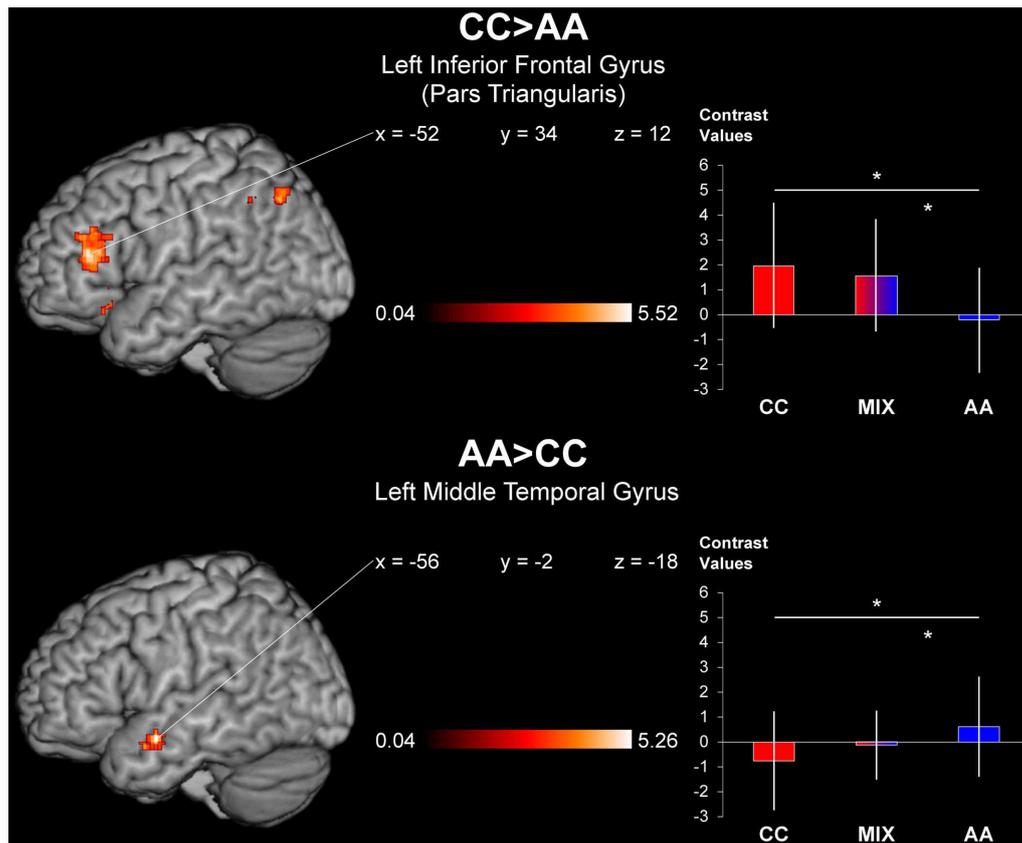
FIGURE 3 | Direct contrasts of concreteness vs. abstractness.
Differences between processing concrete noun-verb combinations (top panel) compared to abstract noun-verb combinations (bottom panel) and extracted contrast values for the pure abstract, the summarized mixed conditions and the pure abstract condition from defined local maxima. Note that for visualization the statistical images were thresholded at $p < 0.001$, uncorrected, with an extended cluster size of ≥ 45 contiguous voxels ($360\,mm^3$), superimposed on the MNI template using the software MRIcron Version 12/2012 (http://www.mccauslandcenter.sc.edu/mricro/mricron/index.html). The contrast values were extracted from the individual beta images and are depicted as group mean with standard deviation of the mean. Asterisks indicate statistical differences of *post-hoc* paired t tests ($p < 0.05$, Bonferroni-corrected for multiple comparisons).

Table 3 | Macroanatomical structure, cytoarchitectonical area (Area_cyto), percent overlap of cluster with cytoarchitectonical area, cluster size in voxel, MNI coordinates (x, y, z), and maximum T value (T_{max}) of the local maxima from the direct contrasts of concrete noun-verb combinations compared to abstract noun-verb combinations (CC > AA) and vice versa (AA > CC).

Local maximum in macroanatomical structure	Area_cyto	Percent overlap of cluster with cytoarchitectonical area	Cluster size (voxel)	MNI coordinates			T_{max}
				x	y	z	
CC > AA							
L. Inferior frontal gyrus (Pars triangularis)	Area 45	63.3	207	−52	34	12	5.55
L. Inferior parietal lobule	hIP1	70.2	80	−36	−50	40	4.89
L. Inferior parietal lobule	IPC (PGa)	54.7	59	−34	−68	42	4.57
AA > CC							
L. Middle temporal gyrus*			45	−56	−2	−18	5.28

The significance level was set to $p < 0.05$, FWE corrected for small volumes using the image masks of the SPM Anatomy toolbox v1.8 (Eickhoff et al., 2005). A cluster size of ≥ 45 contiguous voxels ($360\,mm^3$) extended the threshold. Abbreviations: L, left; R, right.

*Note that due to the non-availability of a cytoarchitectonical map for that area a mask of the temporal lobe was generated using the WFU PickAtlas Toolbox v3.0.4 (Maldjian et al., 2003) which was applied within the small volume correction and improved the significance.

interest left middle temporal gyrus did not reach the corrected significance threshold [$T_{(24)} = 2.17$; $p = 0.040$]. Note that based on the above mentioned main effects, condition CC differed significantly from condition AA in all clusters, i.e., left inferior frontal gyrus [$T_{(24)} = 5.34$; $p < 0.001$] and left middle temporal gyrus [$T_{(24)} = 3.21$; $p = 0.004$].

DISCUSSION

Embodied cognition theories propose that during language comprehension an internal simulation of the content of the word or sentence occurs. Thus, involvement of the same sensorimotor neural network is assumed during the simulation process as while interacting with an object or entity or while executing the action, the word refers to Zwaan (2004).

One of the core open questions in this area concerns the difference in neural representations of concrete and abstract words, as for instance "cake" vs. "theme." This functional imaging study addressed this question by presenting participants with combinations of nouns referring to graspable/non-graspable objects/entities and motor/non-motor verbs within a concreteness-to-abstractness continuum in order to generate a novel experimental design which also allows to differentiate between the role of verbs and nouns in abstract contents processing.

Firstly, our imaging results replicate previous findings that demonstrate the involvement of motor areas in language comprehension. In contrast to previous studies we chose regular and rhythmic paced and unpaced finger-tapping as a functional localizer. It is important to note that this task excited both primary motor areas and adjacent regions, which are referred to as secondary motor areas that are consequently involved in action observation and language processing. Results show a significant overlap of activations that were evoked by the functional localizer task and also as a result of perceiving abstract and concrete multi-word expressions. The regions involved included the left lateral (precentral gyrus) and medial (supplementary motor area) premotor cortex.

Importantly, our first anatomical prediction was confirmed for the motor areas recruited by language stimuli. Specifically, the activations elicited by the concrete and abstract multi-word expressions were not significantly different from each other within the motor areas identified by the localizer task. Hence, processing of both concrete and abstract language content is crucially supported by the sensorimotor neural network.

Regarding our second anatomical hypothesis, which focused on the dissociation of core areas for both pure concrete and mere abstract multi-word expressions, considerable differences in neural correlates were identified in the present study. Processing of concrete compared to abstract multi-word content predominantly employed a fronto-parietal network, which is a well-known circuit for object perception and manipulation (see early fMRI studies by Binkofski et al., 1999; Buccino et al., 2001). This shows that this network could also be activated by reading nouns that refer to graspable objects, which might reflect the possible nature of the interaction with the object. Conversely, processing of abstract noun-verb combinations compared to concrete language content showed a pronounced activation in the left anterior middle temporal gyrus. Crucially this area is close to the language processing system (see Price, 2010).

The finding of left-lateralized contribution of middle temporal gyrus to abstract rather than to concrete words comprehension is in line with several functional imaging studies on the ability to mentally imagine concrete vs. abstract nouns (Mellet et al., 1998; Sabsevitz et al., 2005). In previous studies, anterior middle temporal gyrus has consistently been activated during categorization of unique entities, such as famous faces (Sergent et al., 1992; Gorno-Tempini et al., 1998; Leveroni et al., 2000; Martin and Chao, 2001). In recent TMS studies, left posterior middle temporal gyrus was shown to be a crucial part of a distributed network for semantic control (Whitney et al., 2011, 2012). Interestingly, (posterior) middle temporal gyrus was found to be recruited during processing of semantic jokes (Goel and Dolan, 2001), while the same area is also involved in control for action rationality, as for instance when goal-directed actions are violated due to contextual constraints, as in the presence of an obstacle (Jastorff et al., 2011). The responsiveness to the violation of contextual constraints seem to be critical in communication and in social interaction. In turn, this appears to be in line with the WAT proposal that the social context of language acquisition impacts upon representations of abstract rather than concrete language content.

Crucially, the pronounced involvement of left middle temporal gyrus in abstract language processing supports multiple representation theories like LASS and WAT. This is insofar as although concrete and abstract language content engages the sensorimotor neural network, abstract word processing relies more on the linguistic neural system. The idea of parallel systems, the language and the motor one, for preferentially processing concreteness and abstractness has been already discussed by Scorolli et al. (2011). Their reaction time study used the same linguistic material but implemented a sentence evaluation task. Reaction times were significantly faster while processing pure concrete and pure abstract language content (CC, AA) compared to the mixed conditions CA and AC (see also a recent TMS study with the same paradigm, Scorolli et al., 2012). However, disambiguation of processing within linguistic neural system preceding or resulting from sensorimotor processing is constrained by the poor temporal resolution of fMRI data in principle.

As well as mere abstract language content, the mixed combinations (CA, AC) also characterized representations of concept meaning. Thus, according to our second anatomical hypothesis, the neural language network was also assumed to be involved while semantically processing nouns referring to a graspable object combined with non-motor verbs or nouns referring to a non-graspable entity combined with motor verbs. Indeed, the mixed combinations reveal activations of the same areas as recruited by the pure abstract and concrete condition, even if to a greater or lesser extent. To point out in detail, analysis of effect sizes within the frontal and temporal regions of interest showed that the pure concrete condition (CC) did not differ from the summarized mixed conditions (CA + AC), but rather conversely, the pure abstract condition (AA) differed significantly from the summarized mixed conditions (CA + AC). Consequently, if either the noun or the verb becomes an abstract meaning, semantic processing predominantly changes to ambiguity. These results

in part support the ability of our paradigm to implement a continuum from concreteness to abstractness, and thus, future research could adopt a similar approach by finding appropriate linguistic material.

One point is worth of notice. Even if the present study did not aim to investigate the different representations elicited by multi-word expressions with various degrees of metaphoricity, we cannot exclude a variation of our materials along the continuum of literal-figurative language. Related theories like the coarse-semantic-coding theory (Jung-Beeman, 2005) and the graded salience hypothesis (Giora, 1997) focused directly on the different neural underpinnings of literal and figurative language. The former suggests a right hemispheric advantage for tasks requiring both the integration of distant semantic concepts and for the understanding of figurative language whereas the left hemisphere, instead, would be specialized in analytic tasks that require the processing of literal semantic associations. The latter assumes the dimension of novelty-conventionality to be more salient than the distinction between literal and metaphorical language. Both theories predict that literal language is processed primarily in the dominant left hemisphere, while novel figurative language has faster access to the right hemisphere. But the two theories differ with respect to conventional figurative expressions: According to the coarse-semantic-coding theory the right hemisphere is rather involved in conventional metaphors than in literal expressions, while the graded salience hypothesis proposes conventional metaphors to be rather processed in the left hemisphere. Due to inconsistent evidence provided by functional imaging studies, Bohrn et al. (2012a) conducted a meta-analysis on neuroimaging studies and have found asides bilateral frontal activations the left middle temporal gyrus to be involved in figurative as compared to literal language processing and in conventional as compared to novel metaphors processing, in line with the graded salience hypothesis. The authors concluded that literal and figurative language processing elicit shared neural correlates, but figurative language requires more cognitive resources to integrate words at the phrase or sentence level, thus, recruit more widespread activations. In sum, the meta-analytic results are convergent with our finding of distinct left lateralized activation within the middle temporal gyrus for abstract sentences, which to some extent resemble conventional metaphors. Interestingly, pronounced extensive left middle temporal gyrus activation was also found by another study by Bohrn et al. (2012b) to be correlated with unfamiliar as compared to familiar proverbs.

Thereby, an interesting idea would be to investigate whether any of the activated areas are triggered just by the noun or the verb depending on the emphasis of concreteness or abstractness on the noun or verb. However, as the noun and the verb were presented simultaneously in our paradigm, this issue remains rather speculative. Interestingly, the study by Rüschemeyer et al. (2007) demonstrated the impact of processing the meaning of the entire word (e.g., "be-greifen," to comprehend) compared to the meaning of its morphological concrete or abstract components (e.g., "greifen," to grasp). The components identified by Rüschemeyer et al. (2007) concerned single words rather than

word combinations, even though it is possible that the same principle could also be applied to the combination of words. With respect to our results, another interesting open question arises regarding sentence comprehension, i.e., whether a single abstract word, independently of its grammatical class, could shift the whole sentence meaning to a mere abstract one.

Moreover, our imaging data showed a clear pattern of left-lateralized neural correlates associated with both concrete and abstract language content. This is, in part, in line with the findings by Binder et al. (2005), as they reported left-lateralized processing of abstract words and bilateral activations of associative areas during the processing of concrete words. The discrepant evidence might be due to differences in both the tasks and the linguistic materials. Compared to the study of Binder et al. (2005) in which a lexical decision task on word-nonword categorization implies superficial processing of concrete highly imageable words (e.g., "cloud") vs. abstract low imageable words (e.g., "dogma") and non-words, our paradigm required semantic processing by combining a noun with a verb within a concreteness-to-abstractness continuum. However, in line with our study Desai et al. (2010) employed a sentence sensibility task implying semantic processing of the linguistic stimuli by manipulation just of the verb meaning (e.g., "use the hammer"—"use the opportunity"), thus, resulting in bilateral activation associated with abstract language, even if stronger for the left hemisphere. Consequently, it remains a matter of debate whether processing of abstract language content rests more on a bilateral than on a left-lateralized neural network.

Similarities in design were highly visible in a study by Christoff et al. (2009) in which anagrams with concrete words (as for instance "desk"), moderately abstract words (as for instance "dance"), and highly abstract words (as for instance "myth") were used and yielded a functional topography in the prefrontal cortex with relative stronger recruitment of left hemispheric ventrolateral, dorsolateral and rostrolateral prefrontal cortex, respectively, with an curvilinear direction of increase in representational abstraction. Interestingly, even if below the used significance and cluster threshold, we have also found an activation within the rostrolateral prefrontal cortex elicited by the contrast AA > CC (cluster size = 6 voxels; MNI coordinates: $x = 30$, $y = 46$, $z = -4$; maximum T value = 3.95). Moreover, our activation was right hemispheric. In contrast to our study, Christoff et al. (2009) employed a task and induced mindsets by cuing the participants to the degree of abstraction of the anagram solution, thus, this might have contributed to the missing significance in our rostrolateral prefrontal activation. Also the application of both a mask of the frontal lobe and a mask of the Brodman area 10, which were generated with the WFU PickAtlas Toolbox v3.0.4 (Maldjian et al., 2003), did not yield significance within a small volume correction. However, the differential representations invoked by reading the word combinations including abstract language content might fit well to the hypothesis of hierarchical processing within the prefrontal cortex (for review, see Badre, 2008; Botvinick, 2008).

CONCLUSION

The discussion on cognitive as well as neural representations of concrete vs. abstract linguistic stimuli is still a matter of

keen debate. The present fMRI study addressed this question by using a novel paradigm to demonstrate considerable functional dissociations in the neural correlates associated with the concrete and abstract contents of language. In contrast to previous studies that have generally focused on single words in rather superficial lexical-semantic decision-making tasks (for a review, see Sabsevitz et al., 2005), our paradigm implemented semantic processing of multi-word expressions within a concreteness-to-abstractness continuum. To this end, each concrete noun (denoting graspable objects) was combined with an adequate concrete motor verb and an adequate abstract non-motor verb. Likewise, adequate abstract nouns (denoting a non-graspable entity) were combined with either kind of verbs previously used.

First of all, both concrete and abstract multi-word expressions activated the core areas of the sensorimotor neural network. Hence, this is in line with embodied cognition theories. The finding suggests that internal simulation results in the activation of sensorimotor representations, wherein the grounding is in the sensorimotor system for not only concrete but also for abstract language content. In order to show dissociative neural correlates, direct contrasts of pure concrete vs. mere abstract noun-verb combinations and *vice versa* were used. Concrete stimuli revealed adjacent activations to the sensorimotor system whereas abstract stimuli elicited pronounced activation of areas known to underlie lexical and phonological processing. Multiple representations like this in turn are predicted by embodied cognition theories including LASS and WAT proposals. Since both LASS and WAT rely on the idea that multiple representations are activated, both theories are compatible with the results we obtained.

However, only WAT makes specific predictions concerning the difference between concrete and abstract words. As explained in more detail in the introduction, the two theories differ in the role they ascribe to linguistic processes. LASS is focused on lexical vs. conceptual levels of language processing, as it assumes that linguistic processes might be rather superficial, while conceptual processes are not. WAT, instead, does not treat linguistic processing as superficial since it can convey meaning. This theoretical difference has lead the former to put emphasis on the differences between more deep and more superficial tasks and processes (e.g., on the difference between lexical decision and picture naming), without focusing on the differences between semantic categories. In contrast, specific predictions concerning the representation of concrete and abstract words derive from the WAT proposal. More specifically, according to WAT the sensorimotor neural network is engaged by both concrete and abstract words, but in particular by concrete words, while the linguistic neural network is pronounced activated by abstract words. According to WAT, the differences in the representation between these two kinds of words are due to their different acquisition modality, since the absence of a concrete word referent with abstract words needs to be compensated by the use of linguistic labels and explanations. This has been demonstrated through recent behavioral evidence in which new concrete vs. abstract words are learnt (Borghi et al., 2011), but further neural evidence could complement this behavioral data. In sum, even if our findings are compatible with both LASS and WAT theories, the WAT proposal can better predict and account for the dissociation of concrete and abstract language content that we presented in our study.

REFERENCES

Annett, M. (1970). A classification of hand preference by association analysis. *Br. J. Psychol.* 61, 303–321.

Arévalo, A. L., Baldo, J. V., and 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.

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1996). *CELEX2.* Philadelphia, PA: Linguistic Data Consortium.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn. Sci.* 12, 193–200.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645.

Barsalou, L. W., Santos, A., Kyle Simmons, W., and Wilson, C. D. (2008). "Language and simulations in conceptual processing," in *Symbols, Embodiment and Meaning,* eds M. de Vega, A. M. Glenberg, and A. C. Graesser (Oxford: Oxford University Press), 245–283.

Barsalou, L. W., and Wiemer-Hastings, K. (2005). "Situating abstract concepts," in *Grounding Cognition: The Role of Perception and Action in Memory, Language and Thinking,* eds D. Pecher and R. A. Zwaan (New York, NY: Cambridge University Press), 129–163.

Bedny, M., and 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.

Binder, J. R., Westbury, C. F., McKiernam, K. A., Possing, E. T., and Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *J. Cogn. Neurosci.* 17, 905–917.

Binkofski, F., Buccino, G., Posse, S., Seitz, R. J., Rizzolatti, G., and Freund, H. J. (1999). A fronto-parietal circuit for object manipulation in man: evidence from an fMRI-study. *Eur. J. Neurosci.* 11, 3276–3286.

Bohrn, I. C., Altmann, U., and Jacobs, A. M. (2012a). Looking at the brains behind figurative language—a quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia* 50, 2669–2683.

Bohrn, I. C., Altmann, U., Lubrich, O., Menninghaus, W., and Jacobs, A. M. (2012b). Old proverbs in new skins—an fMRI study on defamiliarization. *Front. Psychol.* 3:204. doi: 10.3389/fpsyg.2012.00204

Borghi, A. M., and Cimatti, F. (2009). "Words as tools and the problem of abstract words meanings," in *Proceedings of the 31st Annual Conference of the Cognitive Science Society,* eds N. Taatgen and H. van Rijn (Amsterdam: Cognitive Science Society), 2304–2309.

Borghi, A. M., and Cimatti, F. (2010). Embodied cognition and beyond: acting and sensing the body. *Neuropsychologia* 48, 763–773.

Borghi, A. M., Flumini, A., Cimatti, F., Marocco, D., and Scorolli, C. (2011). Manipulating objects and telling words: a study on concrete and abstract words

acquisition. *Front. Psychol.* 2:15. doi: 10.3389/fpsyg.2011.00015

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends Cogn. Sci.* 12, 201–208.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404.

Chatterjee, A. (2010). Disembodying cognition. *Lang. Cogn.* 2, 79–116.

Chiarello, C., Senehi, J., and Nuding, S. (1987). Semantic priming with abstract and concrete words: differential asymmetry may be postlexical. *Brain Lang.* 31, 43–60.

Christoff, K., Keramatian, K., Gordon, A. M., Smith, R., and Mädler, B. (2009). Prefrontal organization of cognitive control according to levels of abstraction. *Brain Res.* 1286, 94–105.

Clark, A. (1999). "Embodied, situated, and distributed cognition," in *A Companion to Cognitive Science,* eds W. Betchel and G. Graham (Malden, MA: Blackwell Publishing), 506–517.

Day, J. (1979). Visual half-field word recognition as a function of syntactic class and imageability. *Neuropsychologia* 17, 515–519.

Deloche, G., Seron, X., Scius, G., and Segui, J. (1987). Right hemisphere language processing: lateral difference with imageable and nonimageable ambiguous words. *Brain Lang.* 30, 197–205.

Desai, R. H., Binder, J. R., Conant, L. L., and Seidenberg, M. S. (2010). Activation of sensory-motor areas in sentence comprehension. *Cereb. Cortex* 20, 468–478.

Dove, G. (2009). Beyond conceptual symbols: a call for representational pluralism. *Cognition* 110, 412–431.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335.

Fischer, M., and Zwaan, R. (2008). Embodied language: a review of the role of the motor system in language comprehension. *Q. J. Exp. Psychol.* 61, 825–850.

Friederici, A. D., Opitz, B., and von Cramon, D. Y. (2000). Segregating semantic and syntactic aspects of processing in the human brain: an fMRI investigation of different word types. *Cereb. Cortex* 10, 698–705.

Gallese, V. (2008). Mirror neurons and the social nature of language: the neural exploitation hypothesis. *Soc. Neurosci.* 3, 317–333.

Gentner, D. (2010). Psychology in cognitive science: 1978–2038. *Top. Cogn. Sci.* 2, 328–344.

Giora, R. (1997). Understanding figurative and literal language: the graded salience hypothesis. *Cogn. Linguist.* 8–3, 183–206.

Glenberg, A. M., and Kaschak, M. P. (2002). Grounding language in action. *Psychon. Bull. Rev.* 9, 558–565.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., and Buccino, G. (2008). Processing abstract language modulates motor system activity. *Q. J. Exp. Psychol.* 61, 905–919.

Goel, V., and Dolan, R. J. (2001). The functional anatomy of humor: segregating cognitive and affective components. *Nat. Neurosci.* 4, 237–238.

Gorno-Tempini, M. L., Price, C. J., Josephs, O., Vandeberghe, R., Cappa, S. F., Kapur, N., et al. (1998). The neural systems sustaining face and proper name processing. *Brain* 121(Pt 11), 2103–2118.

Grossman, M., Koenig, P., DeVita, C., Glosser, G., Alsop, D., Detre, J., et al. (2002). The neural basis for category-specific knowledge: an fMRI study. *Neuroimage* 15, 936–948.

Holcomb, P. J., Kounios, J., Anderson, J. E., and West, W. C. (1999). Dual-coding, context availability, and concreteness effects in sentence comprehension: an electrophysiologic investigation. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 721–742.

Jastorff, J., Clavagnier, S., Gergely, G., and Orban, G. A. (2011). Neural mechanisms of understanding rational actions: middle temporal gyrus activation by contextual violation. *Cereb. Cortex* 21, 318–329.

Jirak, D., Menz, M. M., Buccino, G., Borghi, A. M., and Binkofski, F. (2010). Grasping language – a short story on embodiment. *Conscious. Cogn.* 19, 711–720.

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends Cogn. Sci.* 9, 512–518.

Kellenbach, M. L., Wijers, A. A., Hovius, M., Mulder, J., and Mulder, G. (2002). Neural differentiation of lexico-syntactic categories or semantic features? Event-related potential evidence for both. *J. Cogn. Neurosci.* 14, 561–577.

Kemmerer, D., Rudrauf, D., Manzel, K., and Tranel, D. (2012). Behavioral patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions. *Cortex* 48, 826–848.

Kiehl, K. A., Liddle, P. F., Smith, A. M., Mendrek, A., Forster, B. B., and Hare, R. D. (1999). Neural pathways involved in the processing of concrete and abstract words. *Hum. Brain Mapp.* 7, 225–233.

Kousta, S., Vigliocco, G., Vinson, D., and Andrews, M. (2009). "Happiness is … an abstract word: the role of affect in abstract knowledge representation," in *Proceedings of the 31st Annual Conference of the Cognitive Science Society,* eds N. Taatgen and H. van Rijn (Amsterdam: Cognitive Science Society).

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., and Del Campo, E. (2011). The representation of abstract words: why emotion matters. *J. Exp. Psychol. Gen.* 140, 14–34.

Leveroni, C. L., Seidenberg, M., Mayer, A. R., Mead, L. A., Binder, J. R., and Rao, S. M. (2000). Neural systems underlying the recognition of familiar and newly learned faces. *J. Neurosci.* 20, 878–886.

Mahon, B. Z., and 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.

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., and Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19, 1233–1239.

Mårtensson, F., Roll, M., Apt, P., and Horne, M. (2011). Modeling the meaning of words: neural correlates of abstract and concrete noun processing. *Acta Neurobiol. Exp. (Wars.)* 71, 455–478.

Martin, A., and Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Curr. Opin. Neurobiol.* 11, 194–201.

Mellet, E., Tzourio, N., Denis, M., and Mazoyer, B. (1998). Cortical anatomy of mental imagery of concrete nouns based on their dictionary definition. *Neuroreport* 9, 803–808.

Moritz, C. H., Haughton, V. M., Cordes, D., Quigley, M., and Meyerand, M. E. (2000). Whole-brain functional MR imaging activation from a finger-tapping task examined with independent component analysis. *AJNR Am. J. Neuroradiol.* 21, 1629–1635.

Nittono, H., Suehiro, M., and Hori, T. (2002). Word imageability and N400 in an incidental memory paradigm. *Int. J. Psychophysiol.* 44, 219–229.

Nolfi, S., and Floreano, D. (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines.* Cambridge, MA: MIT Press.

Noppeney, U., and Price, C. J. (2004). Retrieval of abstract semantics. *Neuroimage* 22, 164–170.

Paivio, A. (1971). *Imagery and Verbal Processes.* New York, NY: Holt, Rinehart and Winston.

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach.* New York, NY: Oxford University Press.

Papeo, L., Corradi-Dell'Acqua, C., and Rumiati, 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., Vallesi, A., Isaja, A., and Rumiati, R. I. (2009). Effects of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PLoS ONE* 4:e4508. doi: 10.1371/journal.pone.0004508

Pecher, D., Boot, I., and Van Dantzig, S. (2011). "Abstract concepts: sensorimotor grounding, metaphors, and beyond," in *The Psychology of Learning and Motivation,* Vol. 54, ed B. Ross (Burlington, MA: Academic Press), 217–248.

Perani, D., Cappa, S. F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M. M., et al. (1999). The neural correlates of verb and noun processing: a PET study. *Brain* 122(Pt 12), 2337–2344.

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann. N.Y. Acad. Sci.* 1191, 62–88.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nat. Rev. Neurosci.* 6, 576–582.

Rodríguez-Ferreiro, J., Gennari, S. P., Davies, R., and Cuetos, F. (2011). Neural correlates of abstract verb processing. *J. Cogn. Neurosci.* 23, 106–118.

Rüschemeyer, S. A., Brass, M., and Friederici, A. D. (2007). Comprehending prehending: neural correlates of processing verbs with motor stems. *J. Cogn. Neurosci.* 19, 855–865.

Sabsevitz, D. S., Medler, D. A., Seidenberg, M., and Binder, J. R. (2005). Modulation of the semantic system by word

imageability. *Neuroimage* 27, 188–200.

Scorolli, C., Binkofski, F., Buccino, G., Nicoletti, R., Riggio, L., and Borghi, A. M. (2011). Abstract and concrete sentences, embodiment, and language. *Front. Psychol.* 2:227. doi: 10.3389/fpsyg.2011.00227

Scorolli, C., Jacquet, P. O., Binkofski, F., Nicoletti, R., Tessari, A., and Borghi, A. M. (2012). Abstract and concrete phrases processing differentially modulates corticospinal excitability. *Brain Res.* 1488, 60–71.

Semin, G. R., and Smith, E. R. (2008). *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches.* New York, NY: Cambridge University Press.

Sergent, J., Ohta, S., and MacDonald, B. (1992). Neuroanatomy of face and object processing. *Brain* 115(Pt 1), 15–36.

Smith, L. B. (2005). Cognition as a dynamic system: principles from embodiment. *Dev. Rev.* 25, 278–298.

Toni, I., de Lange, F. P., Noordzij, M. L., and Hagoort, P. (2008). Language beyond action. *J. Physiol. Paris* 102, 71–79.

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., and Jefferies, E. (2011). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cereb. Cortex* 21, 1066–1075.

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., and Jefferies, E. (2012). Executive semantic processing is underpinned by a large-scale neural network: revealing the contribution of left prefrontal, posterior temporal, and parietal cortex to controlled retrieval and selection using TMS. *J. Cogn. Neurosci.* 24, 133–147.

Ziemke, T. (2002). Situated and embodied cognition (special issue). *Cogn. Syst. Res.* 3, 271–554.

Zwaan, R. A. (2004). "The immersed experiencer: toward an embodied theory of language comprehension," in *The Psychology of Learning and Motivation*, Vol. 44, ed B. H. Ross (New York, NY: Academic Press), 35–62.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

PROCEDURE

Besides selecting the linguistic material with regard to familiarity, probability of use, and lexical frequencies, the noun-verb combinations were standardized concerning imageability, literality, quantity of motion and age of acquisition. Thirty German students of the University of Hamburg were asked to rate on a continuous line scale (scores ranging from 0 to 100) the ease or difficulty with which each multi-word expression evoked mental images (imageability: low imagery rate—high imagery rate), how literally they would take each pair (literality: literal—no literal) as well as whether and to what extent each pair elicited movement information (quantity of motion: not much movement—much movement). Additionally, 10 German students were asked to rate at which age they approximately had learned to use each noun-verb combination (age of acquisition). For each rating, scores' averages (M), and scores' standard deviations (SD) were calculated for each condition: noun referring to a graspable object/motor verb (CC), noun referring to a graspable object/non-motor verb (CA), noun referring to a non-graspable entity/motor verb (AC), noun referring to a non-graspable entity/non-motor verb (AA).

RESULTS: IMAGEABILITY

The Concrete noun—Concrete verb combinations were judged as the easiest to imagine ($M = 69.10$, $SD = 12.76$), followed by the Concrete noun—Abstract verb combinations ($M = 52.72$, $SD = 15.80$), the Abstract noun—Concrete verb combinations ($M = 48.53$, $SD = 12.92$), and finally by the Abstract noun—Abstract verb combinations ($M = 45.56$, $SD = 14.51$). Therefore, results found the noun to be stronger than the verb in determining the imageability of the whole multi-word expression (see Paivio, 1965).

RESULTS: LITERALITY

The Concrete noun—Abstract Verb combinations were rated as to take most literally ($M = 18.89$, $SD = 13.72$), followed by the Concrete noun—Concrete verb combinations ($M = 20.22$, $SD = 18.12$), the Abstract noun—Abstract verb combinations ($M = 31.23$, $SD = 19.59$), and finally by the Abstract noun—Concrete verb combinations ($M = 56.95$, $SD = 19.01$). Thus, participants judged the multi-word expression of condition CA

as most literal and the multi-word expression of condition AC as most metaphorical. Note that the meaning of the concrete verb may be different across condition CC and condition AC depending on the context, e.g., the meaning of the verb "schmieden" (to forge) within the expression "einen Ring schmieden" (to forge a ring) differs from its meaning within the expression "einen Plan schmieden" (to forge a plan). In contrast, the meaning of the concrete noun remains the same across condition CC and condition CA, e.g., "einen Ring schmieden" (to forge a ring)—"einen Ring auswählen" (to select a ring).

RESULTS: QUANTITY OF MOTION

The Concrete noun—Concrete verb combinations were rated as eliciting most movement information ($M = 34.29$, $SD = 13.95$), followed by the Abstract noun—Concrete verb combinations ($M = 27.22$, $SD = 12.82$), the Abstract noun—Abstract verb combinations ($M = 17.98$, $SD = 13.87$), and finally by the Concrete Noun—Abstract verb combinations ($M = 13.99$, $SD = 7.39$). Thus, multi-word expression of condition CA elicits less movement.

RESULTS: AGE OF ACQUISITION

The age rated by adults has been demonstrated to be the major independent predictor of the objective age of acquisition indices (Gilhooly and Gilhooly, 1980; Zevin and Seidenberg, 2002). Participants rated the Concrete noun—Concrete verb combinations as the ones they learnt first ($M = 7.82$ years old, $SD = 2.21$), followed by the Concrete noun—Abstract verb combinations ($M = 8.64$ years old, $SD = 2.55$), and finally by both Abstract noun—Concrete verb combinations ($M = 10.74$ years old, $SD = 1.95$) and Abstract noun—Abstract verb combinations ($M = 10.24$ years old, $SD = 2.35$). Hence, based on concrete nouns were learned earlier than abstract nouns results suggest the noun to be critical in language acquisition.

REFERENCES

Gilhooly, K. J., and Gilhooly, M. L. (1980). The validity of age-of-acquisition ratings. *Br. J. Psychol.* 71, 105–110.

Paivio, A. (1965). Abstractness, imagery, and meaningfulness in paired-associate learning. *J. Verbal Learn. Verbal Behav.* 4, 32–38.

Zevin, J. D., and Seidenberg, M. S. (2002). Age of acquisition effects in reading and other tasks. *J. Mem. Lang.* 47, 1–29.