OXFORD

doi: 10.1093/cercor/bhw026 Advance Access Publication Date: 22 February 2016 Original Article

ORIGINAL ARTICLE

Somatotopic Semantic Priming and Prediction in the Motor System

Luigi Grisoni¹, Felix R. Dreyer¹ and Friedemann Pulvermüller^{1,2}

¹Brain Language Laboratory, Department of Philosophy and Humanities, Freie Universtät Berlin, 14195 Berlin, Germany and ²Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, 10099 Berlin, Germany

Address correspondence to Prof. Friedemann Pulvermüller, Brain and Language Laboratory, Department of Philosophy and Humanities, WE4, Freie Universität Berlin, 19145 Berlin, Germany. Email: friedemann.pulvermuller@fu-berlin.de; Dr Luigi Grisoni, Brain and Language Laboratory, Department of Philosophy and Humanities, WE4, Freie Universität Berlin, 19145 Berlin, Germany. Email: grisoniluigi@zedat.fu-berlin.de

Abstract

The recognition of action-related sounds and words activates motor regions, reflecting the semantic grounding of these symbols in action information; in addition, motor cortex exerts causal influences on sound perception and language comprehension. However, proponents of classic symbolic theories still dispute the role of modality-preferential systems such as the motor cortex in the semantic processing of meaningful stimuli. To clarify whether the motor system carries semantic processes, we investigated neurophysiological indexes of semantic relationships between action-related sounds and words. Event-related potentials revealed that action-related words produced significantly larger stimulus-evoked (Mismatch Negativity-like) and predictive brain responses (Readiness Potentials) when presented in body-part-incongruent sound contexts (e.g., "kiss" in footstep sound context; "kick" in whistle context) than in body-part-congruent contexts, a pattern reminiscent of neurophysiological correlates of semantic priming. Cortical generators of the semantic relatedness effect were localized in areas traditionally associated with semantic memory, including left inferior frontal cortex and temporal pole, and, crucially, in motor areas, where body-part congruency of action sound-word relationships was indexed by a somatotopic pattern of activation. As our results show neurophysiological manifestations of action-semantic priming in the motor cortex, they prove semantic processing in the motor system and thus in a modality-preferential system of the human brain.

Key words: grounded cognition, mismatch negativity, readiness potentials, semantic processing

Introduction

Traditional cognitive theories conceive concepts and meanings as abstract mental entities represented in an "amodal" symbolic system functionally separate from other cognitive, perceptual, and motor modules (Anderson 1985). In contrast to this classical view, recent theories of semantic and conceptual grounding emphasize functional interaction between perceptual, motor, and conceptual knowledge (Barsalou 1999; Pulvermüller 1999; Glenberg and Kaschak 2002). In this latter "action-perception perspective," modality-preferential sensory and motor systems along with multimodal convergence zones are seen as genuine loci of semantic processing contributing to the referential semantic knowledge about word object relationships, the storage of aspects of abstract meaning and even the functional relationships between symbols and the actions and goals they may be used to speak about (Barsalou 2008; Binder and Desai 2011; Pulvermüller 2013; Glenberg 2015).

The conceptual grounding approach in cognitive psychology is paralleled, and fuelled, by congruent discoveries in cognitive neurophysiology and neuroscience according to which sensory and motor processes are not functionally separated from cognition, but rather integrated at the level of cortical areas, single

[©] The Author 2016. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/ licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

neurons, and neuronal ensembles (Pulvermüller and Fadiga 2010; Binder and Desai 2011). Of special importance here is the discovery of sensorimotor and mirror neurons (Di Pellegrino et al. 1992), cells that fire both during the execution of actions of a specific type and during the observation of another individual performing the same action; some of these cells respond to the sound of an action (e.g., peanut breaking, Kohler et al. 2002). At the same time, research in the neuroscience of language showed that circuits distributed over modality-preferential (sensory, motor) and multimodal areas become active when meaningful words are being processed (Pulvermüller and Fadiga 2010) and are important for, and causally involved in language understanding as well (Pulvermüller, Hauk, et al. 2005; Glenberg et al. 2008). Crucially, the perception of action sounds and of spoken words typically used to speak about human actions equally activates specific parts of the motor system; these semantically related motor activations can be focal and specific, for example reflecting the body parts with which the relevant actions are typically performed (Hauk et al. 2004, 2006).

Although motor system activation reflects facets of the meaning of action-related language-including its body-part relationship --and functionally influences the processing of action words, one may still suggest that motor areas—and modality-preferential cortices more generally-do not take a genuine role in semantic processing and that true language and concept understanding occurs elsewhere, for example in modality-independent neural systems (Mahon and Caramazza 2008; Bedny and Caramazza 2011; Caramazza et al. 2014) situated in multimodal association cortices. To decide this crucial issue, it is necessary to use an accepted neural index of semantic processing and find out whether this index can be generated in the motor system. As the processing of semantic relationships is, without any doubt, a genuinely semantic process, any neurobiological index of such semantic relationship processing is a good candidate for exploring the brain basis of meaning mechanisms. In behavior, the processing of semantic relatedness is manifest as "semantic priming," the faster and more efficient processing of a meaningful sign if it is placed in context of a semantically related "prime" stimulus (e.g., "apple"-"pear"), when compared with its response in neutral contexts ("castle"—"pear"). An established physiological correlate of semantic priming is the relative reduction of the neurophysiological response to the primed meaningful element (Bentin et al. 1985; Holcomb and Neville 1990; Shtyrov and Pulvermüller 2007; Kiefer and Martens 2010). Therefore, the modulation of the physiological response to a meaningful stimulus by its semantic context offers unique perspectives on localizing the cortical correlates of semantic processing (Kiefer et al. 2011). Following established terminology in cognitive neuroscience, we will speak of "semantic priming (at the neurophysiological/-biological level)" in this sense, to refer to brain correlates of semantic similarity processing (cf., e.g., Kiefer et al. 2011).

Arguably, semantic priming can be carried by both modalitypreferential motor systems and multimodal association cortex, especially in the anterior temporal cortex. Kiefer et al. (2011) found late semantic priming effects in event-related potential, or ERP (latency ca. 400 ms), recordings when cross-modal priming between object and word stimuli was investigated. When ERP priming effects were localized in a standard brain, the underlying generators appeared to spread across anterior temporal cortex and the motor system. However, as localization of cortical sources was performed using group-averaged ERP data, it could be argued that the inclusion of motor regions in the rather widespread set of active areas may reflect limitations of the localization method rather than a specific contribution of motor systems. Ulrich et al. (2013) performed an fMRI experiment on words in a lexical decision task and found semantic priming effects in a range of areas including anterior temporal lobe and posterior left inferior frontal cortex, suggesting a semantic role of Broca's area and/or the motor system. Crucially, this study ruled out the possibility that "epiphenomenal" second thought-related effects were reflected; because these priming effects were also observed when prime words were not consciously perceived (masked priming). Still, this fMRI study does not address the latency at which neurobiological manifestations of semantic relationship processing first occur and whether they are specific to eloquent left-perisylvian areas or also involve adjacent parts of motor cortex.

Classic amodal systems theory and action-perception models of semantics make specific contrarian predictions on the role of motor systems in word meaning and sound processing. According to the amodal approach, nonlinguistic sounds and words are seen as being processed in different systems, in perceptual and linguistic-phonological modules, and meaning processing takes place after information transduction into an abstract symbolic code, within a further encapsulated module for semantics and concepts. For actions indexed by a sound or symbolized by an action word, current theories assume a coding in terms of abstract semantic features (e.g., "kill" as CAUSE TO DIE), by which the part of the body with which an action is typically performed is normally omitted. In contrast, semantic action-perception theory states that the perception and understanding of action sounds and words related to actions both draw on the motor system, where information about the sound- and word-related actions is processed instantaneously as soon as the meaningful icons and symbols are recognized. Words and sounds semantically linked to similar action concepts would therefore have overlapping or adjacent motor circuits, thus predicting action-semantic relationship processing and priming at the physiological level when an action word is presented in the context of a sound indexing a related action. As earlier works showed semantic (including action-semantic) cortical activation within 100-200 ms after a critical spoken word can first be recognized (Pulvermüller 2005; Shtyrov et al. 2014), physiological manifestations of semantic relatedness and thus priming effects can already be expected quite early. If words and sounds relate to actions typically performed with the same part of the body, the somatotopic organization of the motor system predicts that they are processed in close-by cortical loci, thus implying functional interaction and priming between the body-part-congruent signs. In contrast, body-part-incongruent action information would be channeled to distant loci in the motor cortex so that any functional interaction between motor circuits will be reduced or absent. Therefore, the action-related sound of a footstep and the spoken word "kick"-2 acoustically very different stimuli not semantically related to each other, except that the effector of the action semantically linked to the word is the same part of the body as that implied by the action indexed by the sound (footstep),-may exhibit mutual neurophysiological functional interaction (or priming) mediated by the motor representation due to partly overlapping neuronal resources. Even if the action schemas of a kicking action and that of a footstep were entirely nonoverlapping, adjacency of their representations in the somatotopically organized motor system implies rapid interaction and physiological priming between body-part-congruent sounds and words. As sensorimotor features such as body-part congruency are not standard semantic features and therefore are lost when symbols are transduced into an abstract conceptual-semantic code, cognitive-symbolic accounts would not predict or explain any related neurophysiological

priming unless standard criteria for semantic relatedness were met by the stimuli selected (see Materials and Methods and Results sections). In case of any physiological indexes of semantic similarity, the symbolic system account would only allow them in "amodal" systems in multimodal association cortex, but not in sensory or motor systems.

We here presented the critical words "kiss" and "kick" in the context of action (whistling, footstep) and nonaction (water drop) sounds. As dependent neurophysiological measures, we chose the Mismatch Negativity (MMN), and the Readiness Potential (RP) components. The MMN (Näätänen and Winkler 1999) reflects automatic change detection along with the activation of memory circuits for sounds (Frangos et al. 2005), phonemes (Näätänen et al. 1997), and meaningful words (Pulvermüller, Hummel; et al. 2001). The MMN also reveals activation of specific parts of the motor system elicited by both the sounds of actions performed with different parts of the body (Hauk et al. 2006) and similarly by words semantically related to body-part-specific actions (Pulvermüller, Hauk; et al. 2005; Shtyrov et al. 2014). As the MMN is typically elicited while subjects passively process and are even distracted from the eliciting critical stimuli, it reflects cognitive processes that are outside the focus of attention and in this sense automatic. Note, furthermore, that the MMN is an early brain response (latency: 100-250 ms) appropriate to investigate the earliest indexes of semantic processing (latency <200 ms) which are of greatest interest in the debate on grounded semantic processes. Processes following upon early semantic effects and therefore indexed by longer latency event-related potentials may be considered epiphenomenal to language perception and comprehension (see Kiefer and Pulvermüller 2012; Kemmerer 2014). A further brain response of interest was the RP, a negative deflection which typically builds up before the execution of an action (Kornhuber and Deecke 1965) and shows the somatotopic motor localization of the extremities involved in the upcoming action (Cheyne et al. 1991). As there is evidence that the RP also reflects the anticipation of the perception of an action in the visual modality (Kilner et al. 2004), a reliable RP may also be predicted when action sounds are being anticipated. Therefore, we asked whether body-part congruency between action sounds and action words is manifest in both the MMN and the RP and whether any observed neurophysiological indexes of action-semantic relationship and priming are generated in the cortical motor system.

Using an advanced oddball design (Näätänen et al. 2004), action sounds were presented as frequently repeated standard stimuli and action-related spoken words as rare deviant stimuli (see Fig. 1*a*-*b* and Materials and Methods for explanation). Action-perception theory predicts that MMN responses to action words are reduced if they are presented in body-part-congruent contexts of action sounds, but are larger in context of action sounds without relationship to the word. As action sounds were the frequently repeated standard stimulus their occurrence in the experiment could be reliably predicted by experiment participants and an anticipatory RP was therefore expected prior to their onset. We hypothesized that also the RP preceding action sounds may index body-part congruency with regard to the preceding action word.

Materials and Methods

Participants

Twenty-one healthy adults (mean age 25.6 years, ±4.9 SD; 13 females) participated after giving informed written consent. Participants were monolingual English native speakers with normal hearing, normal or corrected-to-normal visual acuity, and



Figure 1. Stimuli and stimulus properties. (a) Spectrograms of the action and nonaction sound stimuli, which were used as frequent standard stimuli (frequency is plotted against time, the color code gives spectral power), (b) Spectrograms for 2 of the rare deviant stimuli (critical words "kick" and "kiss"). (c) Results of behavioral ratings of the strength of the semantic relationships between critical words ("kick" left, "kiss" right) and each of the sounds used as standard stimuli. Data for footstep sounds are in black, for whistle sound in red, and for water drop sound in blue; means and standard errors of the mean (SEM) of ratings on a Visual Analog Scale (range 0–100) are displayed.

motor control and had no record of neurological or psychiatric disease. Datasets from 5 participants were excluded, 2 due to incomplete data acquisition caused by technical problems and 3 because of low signal-to-noise ratios (SNR <2). Therefore, data from 16 participants (mean age 25.4 years, ±4.5 SD; 9 females), all of them strongly right-handed as determined by the Oldfield handedness inventory (Oldfield 1971) (mean laterality quotient 82.3 ± 14.9 SD), entered the electroencephalography (EEG) analysis. Procedures were approved by the Ethics Committee of the Charité Universitätsmedizin, Campus Benjamin Franklin, Berlin, Germany.

Apparatus

The experiment was conducted in the electrically and acoustically shielded chamber of the Brain Language Laboratory at the Freie Universität Berlin. Outside the chamber, a personal computer (PC) controlled stimulus presentation, timing, and pseudo-randomization by using E-Prime 2.0.8.90 software (Psychology Software Tools, Inc., Pittsburg, PA, USA). All sounds were delivered binaurally via headphones (Ultrasone PRO 450, S-LOGIC[™]). Inside the chamber, a separate PC was used to show the distracting movie ("Shaun the Sheep," Aardman Animations, UK) to the participants, who were seated 1 m from the monitor (see Stimuli and Procedure section).

Stimuli and Procedures

For obtaining the MMN and RP, we used the so-called optimal paradigm (Näätänen et al. 2004; Pakarinen et al. 2009), which is known to elicit reliable MMN responses and can accommodate multiple types of infrequent random critical sounds (also called "deviant stimuli") randomly appearing among the frequently presented standard stimuli. As standard stimuli, we used pairs of brief acoustic events, which were easily identifiable as either 2 consecutive footsteps, whistles, or water drops sounds. Stimulus pairs were used as standard stimuli, because preliminary screening had revealed that pilot participants did not reliably recognize single familiar sounds presented out of context. Correct stimulus classification of the finally accepted paired standard stimuli was ascertained by asking participants to verbally label each of the 3 standard stimulus sounds. All EEG participants correctly identified the sound pairs. Double stimuli were presented with a stimulus onset asynchrony (SOA) of 500 ms and together were considered the frequently repeated standard stimulus, which was kept constant across one entire experimental block (Fig. 1a). Standard and deviant stimuli followed each other with an SOA of 1000 ms. The critical target stimuli were the 2 action-related words, "kick" and "kiss," which were presented as rare deviant stimuli (Fig. 1b). As required by the "optimal design," additional deviant stimuli were included to reduce the predictability of each deviant stimulus. These additional deviant stimulus words and the target stimuli were all phonological minimal pairs differing only in their last phoneme ("kill," "king," "kit," and "kish" [pronounced similarly to "quiche"]). However, theory-driven data analysis focused on the 2 "critical" words (i.e., "kick" and "kiss") that were selected according to semantic criteria and were matched for psycholinguistic features, including absolute lexical frequency ("kick" = 2270; "kiss" = 2484 per 100 million words; British National Corpus: http://www.natcorp.ox.ac.uk/) and lexical status (both being lexically ambiguous as noun and verb).

To avoid that participants could recognize words from idiosyncratic features of one specific recording token, 2 versions of each word, spoken by a male native speaker of English, matched for their fundamental frequency (F0) and sound energy (root mean square [RMS] of their acoustic signal) were used. Stimulus word length, F0 frequency, and sound energy were also matched between deviant stimulus words. Two raters, both native speakers of English not partaking in the physiological experiments, were presented with gates of increasing length to obtain the first point in time when acoustic signals allowed for unique identification of the critical words (recognition points, see Warren and Marslen-Wilson 1987). Results revealed word recognition points for the critical words at approximately 180 ms after their onsets.

The experiment consisted of 3 experimental blocks, whose order was counterbalanced over subjects. In each block, one of the nonlinguistic stimuli (water drop, whistle, or footstep sound) was the standard stimulus and all word stimuli were equiprobable deviants. For each block, the RMS of the acoustic signal amplitude of the standard sounds was adjusted to that of the deviants. The stimulus sequence in each block alternated between the unique standard stimulus and the 12 deviant stimuli (2 tokens of 6 different word types), whereby every second sound was (predictably) a standard stimulus and every other stimulus was an unpredictable deviant (P = 0.082). The occurrence of the stimuli was pseudo-randomized so that each word token was presented once in an array of 12 successive deviants and that 2 successive deviants always were different words. Each of the 3 blocks included 1200 stimuli (20 min), 600 standard and 600 deviant stimuli, and was split in 2 halves with a brief break inbetween. All acoustic stimuli were presented binaurally, through high-quality headphones (Ultrasone HFI-450 S-Logic, Wielenbach, Germany), at a comfortable hearing level. As it is standard in MMN research, subjects were instructed to focus their attention on a silent movie presented throughout the EEG recording and to ignore all sound stimuli.

Before the study, the entire stimulus set was evaluated by 10 English native speaker participants (mean age 28.3 years, ±5.19 SD; 6 female), who did not take part in the EEG experiment. These were presented with the question: "How strong is the semantic relationship between these pairs of sounds and words?" and listened to all pairs of sounds and words later used in the experiment. Subjects expressed their judgment by mouse button click on a continuous scale ranging from 0 to 100.

To ascertain that all word or sound stimuli were perceived as natural (and not, e.g., as synthetic or artificial), the entire stimulus set was once again evaluated by all EEG participants after the recordings. In this case, subjects were asked "How natural is this sound/word to you?"; they gave their response as described above, by mouse click on the continuous scale ranging from 0 to 100. In summary, participants participated in 3 blocks of an experiment following the MMN optimal design with different natural sounds (whistles, footsteps, or water drops) as frequent standard stimuli alternating with 6 unpredictable spoken word deviant stimuli. Sounds and word stimuli were well recognizable and matched for relevant acoustic and psycholinguistic features.

Electrophysiological Recordings

The EEG was recorded through 128 active electrodes embedded in a fabric cap (actiCAP 128Ch Standard-2; Brain Products GmbH, Munich, Germany) and arranged according to the international 10–5 system (Oostenveld and Praamstra 2001). Three electrodes (placed above and below the left eye and to the right outer canthus of the right eye) were used to measure, respectively, the vertical and horizontal electro-oculograms. All electrodes were referenced to an electrode placed on the tip of the nose. Data were amplified and recorded, using the BrainVision Recorder (version: 1.20.0003; Brain Products GmbH), with a passband of 0.1-250 Hz, sampled at 1000 Hz and stored on disk. Impedances of all active electrodes were kept below 10 K Ω . For offline analysis, which was carried out with Brain Products' Analyzer 2.0 (Brain Products GmbH), data were down-sampled to 250 Hz and a digital 1-Hz high-pass filter applied. Note that these filter settings are typical choices for MMN recordings (Näätänen et al. 1997). Trials were epoched to 2000 ms, including a 100-ms prestimulus baseline; the word onset was defined as zero. Therefore, each epoch contained the word-related MMN-like potential and the anticipatory RP in anticipation of the subsequent standard sound along with the sound-evoked activity. After data segmentation, independent component analysis with standard parameters for artifact removal, as implemented in EEGLAB 10 (Swartz Center for Computational Neuroscience, La Jolla, CA, USA; http://www. sccn.ucsd.edu/eeglab), was performed. A component was considered to be artifactual when its topography showed peak activity only over the horizontal or vertical eye electrodes and when it showed a smoothly decreasing power spectrum (typical for eye movement artifacts, Delorme and Makeig 2004). After calculating the independent components, eye blink and movement components were subtracted from the EEG data. On average, 2.3 (range 2-4) components out of 127 were removed from each participant's dataset. Subsequently, the data were filtered using a digital 20 Hz low-pass filter. On average, approximately 5% of the trials were rejected because voltage fluctuation exceeded $100 \,\mu V$ or they had artifacts due to amplifier clipping, bursts of electromyographic activity, or excessive alpha power. Finally, to exclude participants with ERPs with a low SNR, we defined the latency of the ERP responses from the grand average obtained by collapsing signals across all the conditions together. The signal was defined as the 100-ms window centered on the local amplitude maximum. The ratio between the root mean square of this signal and the same measure taken from the baseline (100 ms prior to stimulus onset) was then computed to obtain each individual participant's SNR. In summary, we used standard data recording and preprocessing methods for MMN recording. Since we were also interested in analyzing the anticipatory component prior to standard sounds (RP), we epoched trials of 2000 ms duration.

Data Analysis

Stimulus Ratings

The semantic relationship between natural sound and action-related words was assessed with a 3×2 repeated-measures ANOVA with the factors sound (water drop, whistle, footstep) and word ("kiss," "kick"). Potential differences in sound naturalness were assessed by means of a repeated-measures ANOVA with one threelevel factor sound (water, whistle, water drop), and naturalness of the critical words by a 2×2 repeated-measures ANOVA with the factors word type ("kick," "kiss") and token (comparing the 2 versions for each action-related words, see Stimuli and Procedure).

P50

As our results showed a pronounced P50 component, this ERP component was evaluated separately. For this set of statistical analyses, we focused on signals recorded from frontal electrodes, where the P50 is normally largest. To this end, the average of 5 frontal electrodes (F3, F1, Fz, F2, F4) was calculated. The P50 amplitude was defined as the average ERP in a time window of 40 ms width centered at the local positive maximum of the grand average word-elicited peak (latency <100 ms). P50 peak latencies were the same for the 2 critical words (54 ms for both "kick" and "kiss"). A 3×2 repeated-measures ANOVA with the factors

context (water drop, whistle, footstep) and word ("kiss," "kick") was used to assess significant differences.

Word-Elicited MMN-Like ERP

Because an MMN-like response was in the focus, we focused on fronto-central electrodes (F1, Fz, F2, FC1, FCz, FC2), where the MMN is known to be largest and, therefore, the best SNR can be expected (Pulvermüller and Shtyrov 2006). Because, the acoustic wave forms of the different word types were different, we expected different word-specific ERP waveforms and latencies of MMN-like responses contributing to these ERPs. Therefore, the time window for calculating the word-elicited MMN-like amplitude was adjusted to each critical word's ERP peak latency. The MMN-like response was calculated as the average ERP in a time window of 40 ms centered at the local maximum of the grand average word-elicited negative-going peak (within the interval 100-400 ms from word onset). The word-elicited peak latency for the word "kick" averaged across tokens in the 3 contexts was 306 ms from word onset, and that for the word "kiss" was 294 ms. Effects of word and context on the word-related ERP recorded from the 6 fronto-central electrodes were assessed with a 3×2 repeated-measures ANOVA with the factors Context (water drop, whistle, footstep) and Word ("kiss," "kick"). To test whether the topographic distributions of the word-elicited ERPs differed between the critical words, a further hypothesis-driven-repeated-measures ANOVA was carried out. The average signals recorded from 4 electrodes placed above the motor strip where our hypothesis predicted specific neurophysiological indexes of motor cortex activation (left: C5, TTP7h, CCP5h, CP5; central: Cz, CCP1h, CCP2h, CPz; right: C6, CCP6h, TTP8h, CP6) were submitted to a further 3 × 2 × 3 repeated-measures ANOVA with the factors context (water drop, whistle, footstep), word ("kiss," "kick"), and laterality (left, central, right) was carried out.

Readiness Potential

We used the RP to obtain clues about the brain's predictions of upcoming action- and nonaction sounds, and to investigate the dependence of such predictive brain activity on the semantic relationship between sounds and their preceding action word contexts. Because repetition suppression (Grill-Spector et al. 2006) and habituation (Groves and Thompson 1970) are known to dramatically decrease the cortical responses to frequently repeated stimuli, we evaluated the RP responses for the first half of the experiment.

Potential influences of semantic processing on action perception (i.e., RP) were analyzed on the average of fronto-central electrodes (F1, Fz, F2, FC1, FCz, FC2; same loci as for the MMN). Since the RP has been defined as a complex of several subcomponents (Shibasaki and Hallett 2006), we first analyzed the mean voltage of a large 100 ms-wide time window before standard stimulus onset and then performed a second analysis on its last 40 ms, where RP is known to shift toward body-part-specific motor cortex (Kristeva et al. 1991). To reduce as much as possible the effect of psycholinguistic variables on the RPs, we restricted RP analysis to epochs preceded by the 2 critical words (i.e., "kick" and "kiss"). Potential effects of semantic processing on action perception were assessed in a 3×2 repeated-measures ANOVA with the factors sound (water drop, whistle, footstep) and preceding word context ("kick," "kiss"). The final analysis of predictive motor cortex activation in the last 40-ms preceding sound onset used the same electrode groupings and factors as the topographical MMN analysis.

To further investigate significant main effects and interactions revealed by ANOVAs, F-tests were used for planned comparisons. All results reported survived Bonferroni correction. Partial eta-squared (η_p^2) is reported as index of effect size, assuming that effect sizes of 0.01–0.06 are small, those between 0.06–0.14 moderate and effects >0.14 large (Cohen 1988). When sphericity violations were found in the ANOVAs, Greenhouse-Geisser correction was applied (Greenhouse and Geisser 1959) and corrected P values are reported along with epsilon (ϵ) values. In order to test any functional relationship between the MMN-like and the RP responses, we performed Pearson correlation analyses between these 2 components. These analyses were performed on signals from the fronto-central electrodes (F1, Fz, F2, FC1, FC2, FC2). As 2 comparisons were performed for each action sound, the Bonferroni corrected threshold for significance of such correlation was P = 0.025.

fMRI and Source Localization

As the main prediction to be tested by this experiment addressed a function of the motor system, including primary and premotor cortex, in semantic priming, it was imperative to localize the cortical origin of the neurophysiological priming effects obtained. Significant interactions had revealed different ERP topographies for the MMN-like responses to the critical words along with significant neurophysiological priming effects, and, therefore, we performed distributed source localization on these responses at their respective peak latencies (kick: 306 ms., kiss: 294 ms). Statistical analyses focused on testing the difference between the sources of the primed and the unprimed conditions, the neurophysiological correlate of semantic priming, and the equally crucial comparison of the 2 body-part-incongruent conditions (i.e., "kick" in the context of whistle against "kiss" in the context of footstep) where negative ERP amplitudes stood out against all other conditions (see Results). The procedure adopted for source estimation was the standard method implemented in SPM8 (Litvak et al. 2011), which had previously been used in our lab (e.g., Hanna and Pulvermüller 2014). The template structural MRI included in SPM8 was used to create a cortical mesh of 8196 vertices, which was then co-registered with each subject's electrode cap space using 3 electrodes as fiducials: Fpz, TP9, and TP10. The volume conductors were constructed with an EEG (3-shell) boundary element model. The word-related MMN-like responses, within their respective time windows, were then inverted for each subject thereby constraining spatial source solutions uniformly across participants (Litvak and Friston 2008). This was done using the multiple sparse prior technique, specifically the "Greedy Search" algorithm. Activation maps were then smoothed using a Gaussian kernel of full-width halfmaximum (FWHM) 12 mm, resulting in 4 images per participant (i.e., "kick" and "kiss" in the 2 action-related contexts). Source averages and statistics were calculated at the group level.

To test whether the primed and unprimed conditions differed in their activations, we first averaged the 2 primed conditions and the 2 unprimed conditions, respectively, carried out voxel-byvoxel paired t-Test on these images. This analysis led to an estimate of brain loci of general semantic priming effects. In addition, to test whether the brain activation patterns and priming effects elicited by the critical action-related words differed between each other within the sensorimotor cortices, we carried out paired t-tests on sources in predefined regions of interest (ROIs).

Two ROIs were defined based on the results of a separate fMRI localizer experiment, performed with different subjects. To this end, a group of eighteen participants (mean age 25.4 years, ±5.5 SD; 13 females, the same selection criteria as for the EEG experiment), who did not take part in the EEG study, performed foot and lip movements (Methods as in Hauk et al. 2004). Participants were scanned in a 3T Siemens Tim Trio system (Siemens, Erlangen). The brain regions were defined in relation to a baseline in

which the participants were resting. Participants had to perform toe movement with the right foot and lip movements avoiding contacts between lips. Each movement block was 15 s long and repeated 4 times, with 15 s of rest between blocks. The "peak activation voxel" (largest t value) in fronto-central cortex was selected per movement. ROIs were created with Marsbar 0.43 (MARSeille Boîte À Région d'Intérêt SPM toolbox) as 12-mm-radius spheres (i.e., matching the FWHM of the smoothing parameter) centered at the above-mentioned coordinates. These ROIs were then combined in a unique mask used as Explicit Mask in the paired t-test design. For fMRI and source analysis, P values were thresholded at P < 0.05 corrected for multiple comparisons using the family-wise error (FWE) procedure.

Results

Corpus Analysis and Stimulus Ratings

To estimate the semantic similarity between the critical words and the concepts underlying their context sounds, we followed established distributional semantic approaches (Landauer and Dumais 1997). Corpus statistics were performed using the British National Corpus or BNC (corpus size: 100 million words) by determining how frequently the action-related words appeared within 100-word neighborhoods of plausible sound designations and vice versa (e.g., "(foot)step . . . kiss" and "kiss . . . (foot)step"). Overall, a higher probability of co-occurrence for body-part-incongruent pairings (e.g., "step . . . kiss," "whistle . . . kick"), when compared with body-part-congruent ones ("step . . . kick" and "whistle . . . kick"), was observed (59 vs. 50 occurrences). Likewise, Latent Semantic Analysis metrics from http://swoogle.umbc.edu/ (Ding et al. 2005) confirmed greater similarity indexes for our body-part-incongruent stimulus pairs, when compared with the congruent pairs, when search was restricted to verbs. Therefore, taking contextual statistics of usage as criterion for semantic similarity (Landauer and Dumais 1997), these results suggest closer semantic relationships between our body-part-incongruent stimulus pairs than between body-part-congruent ones.

Statistics on semantic similarity ratings between sounds and action-related words revealed a main effect of Sound ($F_{2,18} = 9.26$, P = 0.002, η_p^2 = 0.51), with Bonferroni corrected post hoc t-tests confirming that action-related words were judged to be more similar semantically to the action sounds (whistle and footstep) when compared with the nonaction sound (water drop) (P = 0.022and P = 0.002, respectively). Crucially, the factors sound and word type revealed a highly significant interaction ($F_{2,18} = 16.53$, P = 0.00008, $\eta_p^2 = 0.65$); the word "kick" was assessed more similar to the footstep sound when compared with the whistle and water drop sounds (P = 0.0009 and P = 0.0002, respectively), and the word "kiss" was seen as more similar to the whistle sound when compared with the other 2 (P = 0.027 and P = 0.016, respectively). Finally, Bonferroni corrected post hoc tests did not reveal statistical difference between the 2 concordant pairs (i.e., "kick"-footstep, and "kiss"—whistle, P = 0.49) (Fig. 1c). Therefore, semantic ratings revealed semantic relationships explainable by body-part relationship, which are not easily explained by corpus statistics and objective distributional semantic methods. Stimulus ratings further revealed that our sounds and words were judged as similarly natural (F's < 1, P > 0.2).

Neurophysiological Results

P50 Is Larger in Nonaction than in Action Context

Inspection of waveforms suggested that word-elicited ERPs diverged from each other already at approximately 80-ms postword onsets. Indeed, a positive deflection prominent at frontal recording sites, which we classify as a P50, showed a main effect of Context ($F_{2,30} = 4.83$, $\varepsilon = 0.9$, adjusted P = 0.018, $\eta_p^2 = 0.24$). Planned comparisons revealed that the P50 elicited by action-related words, that is, by both mouth- and leg-words, were significantly larger in the nonaction sound, that is, water drop, context than in that of action sounds (whistle and footstep contexts, P = 0.043 and P = 0.028, respectively).

ERPs in the MMN Range Reflect Semantic Body-Part Congruency Between Action Sounds and Words

Figure 2a,b shows that the word-elicited ERP P50 was followed by a negative-going peak at approximately 180 ms (N100), a positivegoing wave maximal at approximately 240 ms and an additional negative-going deflection, which seemed to differ between words and contexts. The latter negativity showed fronto-central scalp distribution, negative polarity and early latency (~300 ms from word onset and ~120 ms from average word recognition points) consistently with the MMN profile. To investigate the influence of action-related sounds on action words, we first examined how the word-related early responses were modulated by the context. Fronto-central recordings revealed a main effect of context (F_{2,30} = 4.48, adjusted P = 0.028, ε = 0.82, η_p^2 = 0.23), due to larger responses in the whistle context than in the water drop context (P = 0.028), with footstep context responses in-between (marginally different from nonaction context, P = 0.078). Crucially, the factors context and word type revealed a highly significant interaction (F_{2,30} = 12.75, ε = 0.88, P = 0.0002, $\eta_{\rm p}^2$ = 0.46); the MMNlike component elicited by the word "kick" was greater in the context of whistle sounds than in the footstep and water drop contexts (P = 0.0032 and P = 0.01, respectively), with no significant difference between the latter 2 (P = 1.0). In the very same way, the word "kiss" showed a bigger negative-going ERP response in the incongruent action sound context, that is, the footstep context, than in the other 2, whistle and water drop contexts (P = 0.02and P = 0.002, respectively), again with similar amplitudes in the latter (P = 1.0). In sum, a cross-over double dissociation documents that an enhanced negativity emerged in action sound contexts only if deviant words and standard action sounds were body-part-incongruent (Fig. 2c).

The ANOVA with the additional factor laterality (see Materials and Methods) revealed a further main effect of Word ($F_{1,15} = 5.87$, P = 0.029, $\eta_p^2 = 0.28$), with larger ERPs to the word "kiss" as compare with the word "kick". Furthermore, we observed a significant effect of Laterality ($F_{2,30} = 8.87$, $\varepsilon = 0.62$, P = 0.0054, $\eta_p^2 = 0.37$), due to the well-known ERP maximum at central loci. Crucially, the factors word and laterality showed a significant interaction ($F_{2,30} = 3.79$, $\varepsilon = 0.96$, P = 0.036, $\eta_p^2 = 0.20$). Bonferroni corrected planned comparison tests investigating this latter interaction revealed that the words "kick" and "kiss" elicited more negative-going responses at the central electrodes than at left and right lateral recording sites (P < 0.00002 and P = 0.027, respectively). However, whereas the word "kick" did not show any significant laterality difference (left vs. right, P = 0.82), the word "kiss" led to more negative-going responses over the left than the right scalp (P = 0.000054) (Fig. 2d).

RP Reflects Body-Part Congruency of Action Sounds and Words

In a second evaluation, we focused on the neurophysiological activity preceding the standard sound stimuli. In a 100-ms time window preceding the onset of the standard stimuli, a slow negativity at fronto-central recording sites significantly diverged between action and nonaction sounds, as documented by a main effect of sound ($F_{2,30} = 7.34$, $\varepsilon = 0.95$, P = 0.003, $\eta_p^2 = 0.33$, Fig. 3a), which was due to larger RPs to mouth- and leg-related sounds

compared with the water drop sound (P = 0.011 and P = 0.0047, respectively). Crucially, the factors sound and context word showed a significant interaction ($F_{2,30} = 11.045$, adjusted P = 0.0004, $\varepsilon = 0.92$, $\eta_p^2 = 0.42$), revealing a cross-over dissociation pattern consistent with that of the MMN-like response. The RP was significantly enhanced when the action sound was preceded by a word semantically related to an action which was body-part-incongruent with the sound ("kick"-whistle sound, "kiss"-footstep sound) relative to congruent linguistic context ("kiss"—whistle sound, "kick"—footstep sound; P's = 0.023 and 0.011, respectively; Fig 3b,c). In contrast, the RP preceding the water drop sound did not show any modulation depending on the specific words presented. The additional analysis performed on data from the last 40 ms before sound onset in which the additional factor laterality was included, confirmed these results and revealed a main effect of laterality ($F_{2,30} = 5.61$, adjusted P = 0.012, $\varepsilon = 0.86, \eta_{\rm p}^2 = 0.27$).

Correlation analysis between RP- and MMN-like ERPs showed significant positive correlations only for the incongruent wordsound pairs. The size of the MMN-like response to the word "kiss" in the footstep context correlated with the subsequent footstep-related RP (r = 0.69, P = 0.003) and the MMN-like response to "kick" in the whistle context correlated with the subsequent whistle-related RP (r = 0.61, P = 0.013) (see Fig 3*d*–*e*).

FMRI and Source Localization

First, we extracted data from the fMRI localizer task. Motor activations during foot and lip movements were each contrasted with the resting baseline, and for each comparison, the most pronounced fronto-central activation cluster was extracted. fMRI results for the first contrast revealed activity located in the dorsolateral precentral region (-6, -6, 68, P < 0.001, FWE corrected), whereas the second contrast revealed activity in the ventral motor region (-54, -12, 42, P < 0.001, FWE corrected) (see Fig. 2e and Table 1).

To evaluate potential differences in source space between the primed and unprimed conditions, we performed paired t-tests comparing the average of 2 body-part-congruent and the 2 body-part-incongruent conditions. Results of this first evaluation showed, for the contrast unprimed > primed, significant activation clusters in left inferior frontal areas (-34, 36, -2, P < 0.001, FWE corrected), left anterior temporal pole (-32, 12, -40, P = 0.007, FWE corrected), and dorsal motor cortex (12, -26, 68, P = 0.029, FWE corrected) (see Fig. 2f). The opposite contrast (i.e., primed > unprimed) failed to reveal significant effects after FWE correction.

To test our main hypothesis of early motor activation in semantic processing, we carried out source analysis in ROIs defined with the coordinates extracted from the localizer task (see Materials and Methods). First, we compared the cortical loci of the priming effects elicited by critical words in the sensorimotor ROIs. Comparison of unprimed versus primed leg-word-elicited activation ("kick" in whistle context >"kick" in footstep context) showed the dorsal activation sensorimotor cluster active (-6, -18, 68; -12, -14, 74; 2, -14, 72, P=0.037, FWE corrected) (see Fig. 2q), whereas the semantic priming effect for the face-related word ("kiss" in footstep context > "kiss" in whistle context) significantly activated the inferior sensorimotor ROI (-52, -6, 32; -54, -22, 36; -62, -20, 38, P = 0.038, FWE corrected) (see Fig. 2g). Second, we compared the motor activations for the crucial words in their respective body-part-incongruent action sound contexts, testing for areas with greater responses to the processing of "kick" compared with "kiss" ["kick" > "kiss"] and vice versa ["kiss" > "kick"]. Source estimation results from the first contrast showed a significant cluster located in the left dorso-medial fronto-central region (-12, -16, 70, P = 0.047, FWE corrected) consistent



Figure 2. MMN-like ERP responses to action words. (*a*, *b*) Event-related potential elicited by the 2 critical words in the 3 contexts (footstep context in black, whistle in red, water drop in blue). The top panel (*a*) shows ERPs to "kick" and the bottom panel (*b*) those to "kiss" with their respective current source density, CSD, maps, which provide an estimate of cortical surface activity after removal of volume-conduction effects. (*c*, *d*) The statistically significant interactions between the factors context and word (means and SEM) (*c*) and the factors centrality and word (*d*) are shown. (*e*) Results of the fMRI localizer task. The activations show significant activation clusters (FWE corrected) in the left motor regions for the contrasts foot movements > rest (violet) and lip movements > resting (yellow). (*f*) Significant activation clusters (FWE corrected) for the cortical sources of the general semantic relationship effect obtained by comparing the unprimed conditions (i.e., average of the body-part-incongruent conditions (i.e., average of the specific semantic relationship effects obtained by comparing "kick" unprimed against the "kics" primed conditions (vellow). (*h*), source estimations for the activation contrast between ERPs elicited by words presented in body-part-incongruent context ("kick" > "kics" in violet, "kics" > "kick" in yellow, FWE corrected).



Figure 3. RPs predicting sounds in the context of action words. (a) Event-related potentials in anticipation of the 3 different sounds (footstep sound in black, whistle sound in red, water drop in blue) with their respective CSD maps. The last 100 ms before sound onset are highlighted and used for statistics in figure parts (c-e). (b) Event-related potentials in anticipation of the action sounds in body-part-incongruent (solid line) and in body-part-congruent context (dotted line). (c) Display of the significant interaction between sound and word context. Values (means, SEMs) are presented on the left and right for the 2 critical words and in different colors for the contexts. (d, e) Correlations between the MMN and RP components elicited in incongruent contexts. Correlation plots for the action word "kics" and the subsequent footstep RP (d, black) and for the word "kick" and its subsequent whistle RP (e, red) are shown.

with early involvement of lower limb motor representation. The opposite contrast (i.e., "kiss" > "kick") was significant for a cluster located in the ventral ROI area (-50, -4, 34, P = 0.044, FWE corrected) coherent with mouth motor representation (see Fig. 2h and Table 1). When face- and leg-related action words were presented in body-part-congruent contexts, word-elicited cortical sources did not show significant differences.

Discussion

Our results show that the brain processes elicited by words are modified by the context of action sounds in which they appear, and that, vice versa, the anticipatory brain wave when expecting predictable sounds depends on word context. Crucially, dependent on the relationship between the words' meanings and the nature of the action indexed by action sounds, there was substantial enhancement or reduction of ERPs, including both the word-evoked MMN-like and the sound-expectancy-related RP components. The crucial factors for modulating these ERPs seem to be 2-fold: 1) In the context of sounds not related to actions, MMNs elicited by action words are small and the RP to these nonaction sounds is small too. 2) In the context of action sounds, the MMN responses to action words depends on body-

part congruency between sound and word: if sound and word relate semantically to the same part of the body, the MMN is small, if they index actions performed with different parts of the body, the MMN is substantial. The word-related MMN-like response thus revealed the typical neurophysiological pattern indicative of semantic priming (see Bentin et al. 1985; Holcomb and Neville 1990; Shtyrov and Pulvermüller 2007; Kiefer and Martens 2010). The RP anticipating action sounds follow the same pattern of physiological priming reflecting semantic relationship: the RP predicting action sound is bigger when the preceding word is body-part-incongruent with the sound. Furthermore, correlation analyses showed that the context-dependent dynamics of the RP corresponds to those of the MMN. Crucially, source localization analyses revealed a general priming effect in inferior frontal and anterior temporal areas, whereas specific physiological semantic priming effects for action words with specific body-part relationship appeared, respectively, in dorsal and ventral motor areas, consistent with semantic relationship processing carried by body-part-specific action representations in somatotopically organized sensorimotor cortex.

The time course of the functional influence of action sounds on meaningful words was rapid, as the first sound-dependent neurophysiological changes in word-evoked potentials occurred

Table 1 fMRI and source localization data

	х	у	Z	t-Values	Number of voxels	P-values FWE corrected	Brodmann areas	Cortical areas
Source localization: unprimed > primed	-34	36	-2	8.13	2014	<0.001	47	Inferior frontal
	-32	12	-40	6.08	288	0.007	20	anterior temporal
	12	-26	68	6.12	45	0.029	4	Motor cortex
Source localization: kick unprimed > kick primed	-6	-18	68	3.82	100	0.037	4	Motor cortex
	-12	-14	74	3.63			6	Premotor cortex
	2	-14	72	3.60			6	Premotor cortex
Source localization: kiss unprimed > kiss primed	-52	-6	32	4.87	657	0.038	4	Motor cortex
	-54	-22	36	3.35			3	Somatosensory cortex
	-62	-20	38	3.33			1	Somatosensory cortex
Source localization: kick > kiss	-12	-16	70	3.43	86	0.047	6	Premotor cortex
Source localization: kiss > kick	-50	-4	34	4.04	195	0.044	4	Motor cortex
fMRI localizer task: foot movement > baseline	-6	-6	68	10.97	676	<0.001	6	Premotor cortex
fMRI localizer task: lip movement > baseline	-54	-12	42	11.63	363	<0.001	4	Motor cortex

The table displays the MNI coordinates, t-values, number of voxels per each significant cluster, FWE-corrected P-values, Brodmann labels, and cortical areas for all significant contrasts between the fMRI localizer tasks and the source localizations.

already approximately 300 ms after word onset. Note that the deviant stimulus words could first be recognized approximately 180 ms after their onset and, therefore, the delay of approximately 300 ms upon word onset appeared approximately 120 ms after the word recognition point (Marslen-Wilson 1987), thus implying an early brain index of understanding (Pulvermüller et al. 2009). Still, given that the N400 brain response is known to reflect semantic priming (Bentin et al. 1985; Holcomb and Neville 1990; Kiefer and Martens 2010; Kiefer et al. 2011) and, similar to our present response, appears around 300–500 ms after critical word onset, the present results seem equally consistent with an interpretation in terms of N400 dynamics. However, although both latency and polarity of the semantic priming effect are open to both views, the fronto-central topography of the component argues in favor of an MMN and against an N400 (whose distribution is normally maximal at parietal sites). We therefore call it a "word-evoked MMN-like" ERP. Early semantic priming effects within 200 ms after recognition have earlier been reported by Kiefer et al. (2011) between pictures of object with similar function (pliers-nutcracker) at central ERP recording sites. However, this early central effect was only found for congruent action-related object pictures, not when linguistic stimuli were used as primes. Shtyrov and Pulvermüller (2007) had reported early MMN semantic priming between word pairs, although no unambiguous evidence for motor cortex involvement was obtained.

Enhanced P50 Responses to Words in Nonaction Sound Context

Although our experimental hypotheses focused on MMN-like and RP responses, we observed modulation of an early wordevoked P50 component by context sounds. Substantially, before the deviant words could first be recognized, the word-evoked brain response was enhanced in the nonaction sound context (water drop). In action sound context (whistle and footstep), a relatively small P50 was elicited. Because of its occurrence before the word recognition point, this brain response cannot reflect language understanding at the semantic level. However, it is possible to perceive elementary features of speech sounds early-on, so that the upcoming sounds could be recognized as phonetic in nature and thus as related to articulatory action. Taking an action-oriented perspective, the language stimuli or articulatory action sounds would stand out more in the nonaction context (i.e., water drop) as compare with that of other action sounds. Therefore, already the earliest component of the word-evoked potential may reflect the expectation of the stimuli's action-relatedness manifest at the phonetic (but not at the semantic) level.

Body-Part Congruency Drives Semantic Relationship Processing: MMN Results

Sounds and words of human actions activate different parts of the motor system, depending on the part of the body commonly used to produce the sound or to perform the action the word is typically used to speak about (Pulvermüller, Kujala; et al. 2001; Hauk et al. 2004, 2006; Shtyrov et al. 2004; Pulvermüller, Shtyrov, et al. 2005; Klepp et al. 2014; Shtyrov et al. 2014). These activations emerge rapidly, within approximately 200 ms after stimulus information first allows for identifying the meaningful stimuli. The observation that functional changes in motor systems exert a causal effect on action word processing (Pulvermüller, Hauk, et al. 2005; Glenberg et al. 2008; Liuzzi et al. 2010; Shebani and Pulvermüller 2013; Schomers et al. 2015; Dreyer et al. 2015) have been taken as evidence for an intrinsic functional link between linguistic and motor brain systems in semantic processing (Kiefer and Pulvermüller 2012). However, an alternative view proposes that information in the motor system is not semantic in nature, although it may still be used to supplement semantic representations and, in very special conditions, influence semantics (Mahon and Caramazza 2008).

To evaluate and empirically test these positions against each other, we presented action words in body-part-congruent and -incongruent action contexts. If the motor system can carry a semantic role, this predicts a pattern of physiological semantic priming, that is, that body-part-(in-)-congruent context would reduce (increase) the brain response elicited by the action words. This pattern was revealed by the MMN-like brain response. Major sources for discriminating between body-part-incongruent word stimuli were localized in inferior frontal, anterior temporal and motor areas. Crucially, we observed stronger somatotopic activations within the sensorimotor cortex for the unprimed as compare with primed conditions. This pattern was consistent with specific loci in the mouth- and leg-regions for the processing of semantic relationships between face- and leg-related words/sounds, respectively (see Discussion).

The cell assembly model of language explains this pattern of results by action word representations, conceived as circuits for word forms in perisylvian language cortex heavily interlinked with circuits for semantically related action schemas reaching into primary and premotor cortex and action sound mirror representations binding motor programs to acoustic information about these actions (Pulvermüller 2005). Neurophysiological priming and relationship processing is mediated by overlap or proximity of the cortical circuits for word- and sound-related actions performed with the same or closely adjacent parts of the body. Note again that the somatotopic organization of the motor system implies that body-part-congruent action sound/word information ("kiss"—whistle, "kick"—footstep) is mapped to closely adjacent loci in cortex. In the case of body-part-incongruent words and sounds, motor representations are in different parts of the motor strip, thus failing to achieve functional cross-talk and priming.

To explain the generally reduced activation seen to words presented in context of nonaction sounds, this model can be extended. Action sound and action word each activate their respective neuronal population in the motor cortex, so that in the body-part-incongruent action sound context, the understanding of an (unprimed) action word yields a second activation in motor systems. The net ERP response is therefore large. However, nonaction sounds (water drop) do not activate motor circuits, so that there is overall less activity in motor systems, surfacing as reduced ERP. Therefore, single-action words presented in nonaction sound context would activate the motor system only once and therefore produce less motor activity than the action words in inconsistent action sound context, which activate 2 distinct action circuits along the motor strip. In contrast, MMN-like reduction to words in body-part-congruent action sound context is explained differently, by action-mediated priming. Consistent with this proposal, only the body-part-incongruent sound/word pairs elicited enhanced MMN-like responses.

We note that this model needs extension to account for additional semantic effects. For example, body-part congruency may, in certain contexts, be a necessary condition for action-semantic priming, but it is certainly not a sufficient one. Many experiments show that compatibility of actions performed with the same body part influences semantic and cognitive processing and some results even demonstrate dramatic performance reduction if motor actions and action-related semantic meaning of words and sentences are incompatible (as is the case, e.g., for "push" and "pull") (Glenberg and Kaschak 2002; Glenberg et al. 2008 and many others, see above). However, such inhibitory effects between body-partcongruent action and semantic representations are only present if clearly antagonistic complex motor movements are involved, which cannot be executed together. Therefore, we included in our present experiment actions that do not exclude each other, as, for example, footsteps and kicking, which are both automatized and do not strictly exclude each other in direct succession and even in one coherent action. What our data therefore show is that for not strictly incompatible body-part-congruent action concepts indexed by consecutively presented sounds and words, there is a priming effect in the motor system.

Source Estimation

ERPs showed different topographical patterns for the MMN-like response to the words "kick" and "kiss" presented in body-partincongruent sound contexts, suggesting different neuronal generator distributions. This suggestion was confirmed by neurophysiological source analysis, which revealed that generators in specific cortical areas of interest, predefined based on an fMRI

experiment using a motor localizer tasks, were differentially activated by these action-related words. In particular, the contrast of "kick" versus "kiss" produced activation in left centro-dorsolateral areas, consistent with main cortical sources in leg-related sections of primary and premotor cortices. The opposite contrast was associated with activation in inferior-fronto-central cortical areas in and adjacent to the cortical motor representations of the face and articulators. Comparing the unprimed versus primed conditions for the words with different action-related meanings --- "kick" and "kiss"-the same somatotopic pattern was obtained, thus demonstrating that semantic priming for action words involves the motor system specifically, in a somatotopic manner. Therefore, these localization results provide additional support for the hypothesis that the motor system did in fact house part of the neuronal machinery for the neurophysiological congruency effects observed. As one may argue that neurophysiological source estimation does not provide a firm proof of generator localization, due to the well-known Helmholtz Inverse Problem (von Helmholtz 1853), it is particularly important to point to the fact that our present localizations are consistent with pre-existing data and theory. Body-part-incongruent compared with -congruent contexts elicited stronger activation at key sites known to be of great relevance for semantic memory and semantic priming (inferior frontal and anterior temporal area; for Discussion, see Pulvermüller 2013; Kiefer et al. 2011, see also next section). Likewise, the sources of the action-semantic effects seen here are consistent with a range of previous fMRI studies of word and sentence processing (Hauk et al. 2004; Pulvermüller 2013).

Functional Interaction in the Motor System: Is It Semantic?

A range of previous neuroimaging and neurophysiological studies revealed that semantic priming is physiologically manifest as reduces activation in the anterior inferior-temporal cortex and in more posterior inferior- and middle-temporal areas and fusiform gyrus, with spoken words also eliciting effects in superior temporal cortex (Crinion et al. 2006; Kiefer et al. 2011; Hulrich et al. 2013; Sim et al. 2015). Subcortical structures, including basal ganglia and thalamus, have also been reported to index semantic priming-at least when no switching between languages is involved (Crinion et al. 2006; Ulrich et al. 2013). Neurobiological correlates of semantic priming in inferior and middle frontal gyri and in the precentral motor system were found in some semantic tasks where overt responses had to be performed (Kiefer et al. 2011; Ulrich et al. 2013), but these activations are normally interpreted as correlates of decision, motor preparation, control and execution processes. In contrast, the temporal, especially lefttemporal activity modulations with semantic priming are interpreted as direct evidence for semantic processing in the temporal cortex. In line with this previous evidence, our results show enhanced activations for the unprimed as compare with primed conditions in inferior frontal and anterior temporal areas. However, on the background of the grounded cognition debate, it was necessary to directly address the question whether the motor system, and therefore modality-preferential cortical systems in general, can contribute to semantic processing. This made it necessary to rule out motor and decision confounds and to investigate semantically related functional interaction and putative priming effects specifically supported by circumscribed motor sites. Here, we avoided motor and decision confounds by presenting semantically related and unrelated stimuli "out of the focus on attention" in an oddball-like distraction paradigm (Näätänen et al. 2007). We targeted specific priming effects

supported by different parts of the motor system taking advantage of the somatotopy of the motor system and exploiting body-part congruency and a feature known as "semantic somatotopy" (Pulvermüller 2005; Kiefer and Pulvermüller 2012). In our study, semantic priming effects across modalities (action sounds -action words) were clearly evident, significant and localized to different motor areas using statistical analyses of sources obtained from individual participants not engaging in motor responses. This result shows that the motor system reflects semantic relationship processing at earliest latencies, at the same time when the first indexes of priming emerge in multimodal areas. Clearly, our results do not imply that the motor system is the only seat of semantics, and are in line with current views that semantic meaning is carried by distributed neuronal circuits, whose cortical distributions differ between the meaning types represented. Additional involvement of frontal, parietal and temporal convergence zones in semantic processing and priming is therefore consistent with our current results (Barsalou 2008; Binder and Desai 2011; Pulvermüller 2013; Glenberg 2015), which in fact provide direct evidence for general semantic priming in 2 convergence zones, that is, inferior frontal cortex and temporal pole. Notably, however, localizations of semantic relationship processing were most specific in the motor system, where relationships within different semantic subcategories (face- and leg-related items) were mapped onto different somatotopic areas.

If semantics is seen as the product of a symbolic system in "amodal"-or, more realistically, multimodal-cortex, and motor cortex is at best admitted a supplementary role of semantic enrichment (Mahon and Caramazza 2008), the observed priming effects mediated by the motor system and the behavioral ratings of the strength of the semantic relationships between the 2 critical words and each of the natural sounds (Fig. 1c) appear difficult to explain. Distributional semantic theory construes semantic relationship in terms of co-occurrence statistics of words with each other (Landauer and Dumais 1997). Assuming that word co-occurrences reveal information about conceptual similarity, the co-occurrence statistics between critical words and the verbal labels of the actions indexed by the sounds provide clues about sound/word semantic relationships. Therefore, we performed a corpus search for the 8 possible pairings of the verbal labels of the critical action-related concepts of our experiment (for "step-kick," "step-kiss," "whistle-kick," and "whistle-kiss" and the same pairings in the opposite order) and, in addition, used latent semantic analysis to assess semantic similarity between our action words stimuli and the action sound concepts. As outlined in the Results section, these corpus criteria revealed similar semantic relationships between body-part-congruent and -incongruent stimulus pairs, or even a bias in favor of the bodypart-incongruent pairings. This result suggests either no differences in semantic relationships or a stronger semantic relationship between those items that relate to different parts of the body. Contrasting with these corpus' statistical results, our behavioral data obtained from semantic similarity ratings show that bodypart congruency is reflected in judgments of greater semantic similarity (Fig. 1c). These latter semantic rating results are also consistent with the physiological priming pattern seen in the MMN-like responses, and with the RP pattern as well. Why, then, should concepts semantically unrelated according to standard text-statistical criteria, prime each other at the physiological level and should be judged as semantically close? An obvious explanation is offered by shared or adjacent neuronal substrates of semantic entries in somatotopically organized sensorimotor cortex. As the sensorimotor system mediates semantic priming at the neurophysiological level and concordant semantic

relationships are revealed by semantic ratings, this system should therefore be seen as a genuine contributor to semantic processing.

Action Sound-Predictive RP Indexes Body-Part Congruency of Action Word Contexts

We also investigated the effect of context words on the brain processes related to the expectation and prediction of the upcoming action sounds. As mentioned, the relevant neurophysiological index here is the RP, which normally reflects motor preparation (Walter et al. 1964; Kornhuber and Deecke 1965), but is known to indicate action-relatedness of expected visual percepts (i.e., action perception) too (Kilner et al. 2004). Here participants were exposed to action sounds without being engaged in tasks requiring the understanding of the intentions underlying a specific action. Furthermore, the RP is closely related to the late part of the so-called contingent negative variation, which precedes expected stimulation and reflects sensorimotor integration (Walter et al. 1964). The RP preceding the onset of action sounds was compared between word contexts. As a control which was missing in the previous work (Kilner et al. 2004), a predictable stimulus unrelated to action was investigated. Our results revealed a clear RP only for action-related sounds, whereas for the not-action-related water drop stimulus, the RP was minimal (Fig. 3a). Importantly, larger RPs emerged when the preceding action word and predicted action sound were body-part-incongruent. There was RP reduction for body-part-congruent action sounds and words, and, thus, again, semantic congruency processing consistent with a priming pattern. Although RPs did not reveal similarly significant topographical differences and no significantly different cortical sources between word- or sound-related activities, it is well known that the main cortical generators of the RP are localized in motor and premotor cortex (Shibasaki and Hallett 2006). Therefore, it seems plausible that also the context-specific modulation of the RP has its antecedents in the motor system.

Because the physiological semantic priming effect of bodypart-congruent context appeared to be similar for the RP- and the MMN-like responses, we performed correlation analyses to investigate possible relationship between these temporally separated components. Indeed, in body-part-incongruent conditions, where MMNs and RPs were largest, there were significant correlations between the MMNs and their subsequent RPs. Therefore, although MMN and RP were separated in time and normally have nonidentical cortical generator constellations (Praamstra et al. 1996; Rinne et al. 2000), this correlational result supports the interpretation of a common cortical mechanism for action-related semantic processes reflected by both RPs and MMNs.

Conclusions

Body-part congruency between action sounds and action words was manifest in neurophysiological brain responses, both MMNlike and RP components. The context of (in)congruent action sounds lead to reduced (enhanced) MMN responses to action words in somatotopically structured motor cortex, and the context of (in)congruent action words reduced (enhanced) the RP in anticipation of a predictable upcoming action sound. These neurophysiological signatures are consistent with a neurophysiological pattern of semantic priming between action sounds and action words related to the same part of the body. Body-part congruency was also a factor reflected in semantic judgments but not in distributional semantic analyses obtained using standard corpora and statistical criteria. Source localization results indicated that the motor and premotor cortices contributed to the physiological indexes of semantic congruency between action sounds and words. In proving a role of the motor system in generating the physiological indexes of semantic congruency, our results are evidence for the early involvement of the sensorimotor system in perceiving, understanding, and semantically relating sounds and words. Our results argue against an epiphenomenal role of motor systems in semantic comprehension and higher cognitive processing but rather suggest that most specific semantic processes are located in modality-specific systems.

Funding

This work was supported by the Freie Universität Berlin (startup grant to FP), the Deutsche Forschungsgemeinschaft (Excellence Cluster Languages of Emotion, Pu 97/15-1, Pu 97/16-1), the Medical Research Council (UK) (MC_US_A060_0034, U1055.04.003.00001.01 to FP), the Engineering and Physical Sciences Research Council and the Behavioural and Brain Sciences Research Council, UK (BABEL grant, EP/J004561/1). Funding to pay the Open Access publication charges for this article was provided by the Deutsche Forschungsgemeinschaft.

Notes

We thank Verena Büscher, Mauro Cantino, Prof. Luciano Fadiga and Dr. Jeff Hanna, Prof. Markus Kiefer and Prof. Yury Shtyrov for their help and input at different stages of this work. *Conflict* of *Interest*: None declared.

References

- Anderson JR. 1985. Cognitive psychology and its implications. New York: Freeman & Co.
- Barsalou LW. 2008. Grounded cognition. Annu Rev Psychol. 59:617–645.
- Barsalou LW. 1999. Perceptual symbol systems. Behav Brain Sci. 22:577–609; discussion 610–560.
- Bedny M, Caramazza A. 2011. Perception, action, and word meanings in the human brain: the case from action verbs. Ann N Y Acad Sci. 1224:81–95.
- Bentin S, McCarthy G, Wood CC. 1985. Event-related potentials, lexical decision and semantic priming. Electroencephalogr Clin Neurophysiol. 60:343–355.
- Binder JR, Desai RH. 2011. The neurobiology of semantic memory. Trends Cogn Sci. 15:527–536.
- Caramazza A, Anzellotti S, Strnad L, Lingnau A. 2014. Annu Rev Neurosci. 37:1–15.
- Cheyne D, Kristeva R, Deecke L. 1991. Homuncular organization of human motor cortex as indicated by neuromagnetic recordings. Neurosci Lett. 122:17–20.
- Cohen J. 1988. Statistical power analysis for the behavioral sciences. Hillsdale: Erlbaum.
- Crinion J, Turner R, Grogan A, Hanakawa T, Noppeney U, Devlin JT, Aso T, Urayama S, Fukuyama H, Stockton K, et al. 2006. Language control in the bilingual brain. Science. 312:1537–1540.
- Delorme A, Makeig S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 134:9–21.
- Ding L, Finin T, Joshi A, Peng Y, Pan R, Reddivari P. 2005. Search on the semantic web. IEEE Comput 38:62–69.
- Di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G. 1992. Understanding motor events: a neurophysiological study. Exp Brain Res. 91:176–180.

- Dreyer FR, Frey D, Arana S, von Saldern S, Picht T, Vajkoczy P, Pulvermüller F. 2015. Is the motor system necessary for processing action and abstract emotion words? Evidence from focal brain lesions. Front Psychol. 6:1661.
- Frangos J, Ritter W, Friedman D. 2005. Brain potentials to sexually suggestive whistles show meaning modulates the mismatch negativity. Neuroreport. 16:1313–1317.
- Glenberg AM. 2015. Few believe the world is flat: how embodiment is changing the scientific understanding of cognition. Can J Exp Psychol. 69(2):165–171.
- Glenberg AM, Kaschak MP. 2002. Grounding language in action. Psychon Bull Rev. 9:558–565.
- Glenberg AM, Sato M, Cattaneo L. 2008. Use-induced motor plasticity affects the processing of abstract and concrete language. Curr Biol. 18(7):290–291.
- Greenhouse SW, Geisser S. 1959. On methods in the analysis of profile data. Psychometrika. 24:95–112.
- Grill-Spector K, Henson R, Martin A. 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci. 10:14–23.
- Groves PM, Thompson RF. 1970. Habituation: a dual-process theory. Psychol Rev. 77:419.
- Hanna J, Pulvermüller F. 2014. Neurophysiological evidence for whole form retrieval of complex derived words: a mismatch negativity study. Front Hum Neurosci. 8:886.
- Hauk O, Johnsrude I, Pulvermüller F. 2004. Somatotopic representation of action words in the motor and premotor cortex. Neuron. 41:301–307.
- Hauk O, Shtyrov Y, Pulvermüller F. 2006. The sound of actions as reflected by mismatch negativity: rapid activation of cortical sensory-motor networks by sounds associated with finger and tongue movements. Eur J Neurosci. 23:811–821.
- Holcomb PJ, Neville HJ. 1990. Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. Lang Cognit Process. 5:281–312.
- Kemmerer D. 2014. Are the motor features of verb meanings represented in the precentral motor cortices? Yes, but within the context of a flexible, multilevel architecture for conceptual knowledge. Psychon Bull Rev. 22:1068–1075.
- Kiefer M, Martens U. 2010. Attentional sensitization of unconscious cognition: task sets modulate subsequent masked semantic priming. J Exp Psychol Gen. 139:464–489.
- Kiefer M, Pulvermüller F. 2012. Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. Cortex. 48:805–825.
- Kiefer M, Sim EJ, Helbig HB, Graf M. 2011. Tracking the time course of action priming on object recognition: evidence for fast and slow influences of action on perception. J Cogn Neurosci. 23:1864–1874.
- Kilner JM, Vargas C, Duval S, Blakemore SJ, Sirigu A. 2004. Motor activation prior to observation of a predicted movement. Nat Neurosci. 7:1299–1301.
- Klepp A, Weissler H, Niccolai V, Terhalle A, Geisler H, Schnitzler A, Biermann-Ruben K. 2014. Neuromagnetic hand and foot motor sources recruited during action verb processing. Brain Lang. 128:41–52.
- Kohler E, Keysers C, Umilta MA, Fogassi L, Gallese V, Rizzolatti G. 2002. Hearing sounds, understanding actions: action representation in mirror neurons. Science. 297:846–848.
- Kornhuber HH, Deecke L. 1965. Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. Pflügers Arch für die Gesamte Neurol. 284:1–17.

- Kristeva R, Cheyne D, Deecke L. 1991. Neuromagnetic fields accompanying unilateral and bilateral voluntary movements: topography and analysis of cortical sources. Electroencephalogr Clin Neurophysiol. 81:284–298.
- Landauer TK, Dumais ST. 1997. A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychol Rev. 104:211–240.
- Litvak V, Friston K. 2008. Electromagnetic source reconstruction for group studies. Neuroimage. 42:1490–1498.
- Litvak V, Mattout J, Kiebel S, Phillips C, Henson R, Kilner J, Barnes G, Oostenveld R, Daunizeau J, Flandin G, et al. 2011. EEG and MEG data analysis in SPM8. Comput Intell Neurosci. 2011:852961.
- Liuzzi G, Freundlieb N, Ridder V, Hoppe J, Heise K, Zimerman M, Dobel C, Enriquez-Geppert S, Gerloff C, Zwitserlood P, et al. 2010. The involvement of the left motor cortex in learning of a novel action word lexicon. Curr Biol. 20:1745–1751.
- Mahon BZ, 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.
- Marslen-Wilson WD. 1987. Functional parallelism in spoken word-recognition. Cognition. 25:71–102.
- Näätänen R, Lehtikoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio A, Alku P, Ilmoniemi RJ, Luuk A, et al. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature. 385: 432–434.
- Näätänen R, Pakarinen S, Rinne T, Takegata R. 2004. The mismatch negativity (MMN): towards the optimal paradigm. Clin Neurophysiol. 115:140–144.
- Näätänen R, Winkler I. 1999. The concept of auditory stimulus representation in cognitive neuroscience. Psychol Bull. 12:826–859.
- Oldfield RC. 1971. The assessment and analysis of handedness: the Edinburgh Inventory. Neuropsychologia. 9:97–113.
- Oostenveld R, Praamstra P. 2001. The five percent electrode system for high-resolution EEG and ERP measurements. Clin Neurophysiol. 112:713–719.
- Pakarinen S, Lovio R, Huotilainen M, Alku P, Näätänen R, Kujala T. 2009. Fast multi-feature paradigm for recording several mismatch negativities (MMNs) to phonetic and acoustic changes in speech sounds. Biol Psychol. 82:219–226.
- Praamstra P, Stegeman D, Horstink M, Cools A. 1996. Dipole source analysis suggests selective modulation of the supplementary motor area contribution to the readiness potential. Electroencephalogr Clin Neurophysiol. 98:468–477.
- Pulvermüller F. 2005. Brain mechanisms linking language and action. Nat Rev Neurosci. 6:576–582.
- Pulvermüller F. 2013. How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. Trends Cognit Sci. 17:458–470.
- Pulvermüller F. 1999. Words in the brain's language. Behav Brain Sci. 22:253–336.
- Pulvermüller F, Fadiga L. 2010. Active perception: sensorimotor circuits as a cortical basis for language. Nat Rev Neurosci. 11:351–360.
- Pulvermüller F, Hauk O, Nikulin VV, Ilmoniemi RJ. 2005. Functional links between motor and language systems. Eur J Neurosci. 21:793–797.

- Pulvermüller F, Hummel F, Härle M. 2001. Walking or Talking? Behavioral and neurophysiological correlates of action verb processing. Brain Lang. 78:143–168.
- Pulvermüller F, Kujala T, Shtyrov Y, Simola J, Tiitinen H, Alku P, Alho K, Martinkauppi S, Ilmoniemi RJ, Näätänen R. 2001. Memory traces for words as revealed by the mismatch negativity. Neuroimage. 14:607–616.
- Pulvermüller F, Shtyrov Y. 2006. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. Prog Neurobiol. 79:49–71.
- Pulvermüller F, Shtyrov Y, Hauk O. 2009. Understanding in an instant: neurophysiological evidence for mechanistic language circuits in the brain. Brain Lang. 110:81–94.
- Pulvermüller F, Shtyrov Y, Ilmoniemi RJ. 2005. Brain signatures of meaning access in action word recognition. J Cogn Neurosci. 17:884–892.
- Rinne T, Alho K, Ilmoniemi RJ, Virtanen J, Naatanen R. 2000. Separate time behaviors of the temporal and frontal mismatch negativity sources. Neuroimage. 12:14–19.
- Schomers M, Kirilina E, Weigand A, Bajbouj M, Pulvermüller F. 2015. Causal influence of articulatory motor cortex on comprehending single spoken words: TMS evidence. Cereb Cortex. 25:3894–3902.
- Shebani Z, Pulvermüller F. 2013. Moving the hands and feet specifically impairs working memory for arm- and leg-related action words. Cortex. 49:222–231.
- Shibasaki H, Hallett M. 2006. What is the Bereitschaftspotential? Clin Neurophysiol. 117:2341–2356.
- Shtyrov Y, Butorina A, Nikolaeva A, Stroganova T. 2014. Automatic ultrarapid activation and inhibition of cortical motor systems in spoken word comprehension. Proceedings of the National Academy of Sciences of the United States of America. 111:1918–1923.
- Shtyrov Y, Hauk O, Pulvermüller F. 2004. Distributed neuronal networks for encoding category-specific semantic information: the mismatch negativity to action words. Eur J Neurosci. 19:1083–1092.
- Shtyrov Y, Pulvermüller F. 2007. Early MEG activation dynamics in the left temporal and inferior frontal cortex reflect semantic context integration. J Cogn Neurosci. 19:1633–1642.
- Sim EJ, Helbig HB, Graf M, Kiefer M. 2015. When action observation facilitates visual perception: activation in visuo-motor areas contributes to object recognition. Cereb Cortex. 25: 2907–2918.
- Ulrich M, Hoenig K, Grön G, Kiefer M. 2013. Brain activation during masked and unmasked semantic priming: commonalities and differences. J Cogn Neurosci. 25:2216–2229.
- von Helmholtz H. 1853. Über einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern, mit Anwendung auf die thierisch-elektrischen Versuche. Ann Phys Chem. 89:211–233, 353–377.
- Walter WG, Cooper R, Aldridge VJ, McCallum WC, Winter AL. 1964. Contingent negative variation: an electric sign of sensorimotor association and expectancy in the human brain. Nature. 203:380–384.
- Warren P, Marslen-Wilson W. 1987. Continuous uptake of acoustic cues in spoken word recognition. Percept Psychophys. 41:262–275.