Psychological Research
DOI 10.1007/s00426-011-0390-z

ORIGINAL ARTICLE

Bidirectional semantic interference between action and speech

Roman Liepelt · Thomas Dolk · Wolfgang Prinz

Received: 14 February 2011 / Accepted: 21 October 2011

Abstract Research on embodied cognition assumes that language processing involves modal simulations that recruit the same neural systems that are usually used for action execution. If this is true, one should find evidence for bidirectional crosstalk between action and language. Using a direct matching paradigm, this study tested if action–languages interactions are bidirectional (Experiments 1 and 2), and whether the effect of crosstalk between action perception and language production is due to facilitation or interference (Experiment 3). Replicating previous findings, we found evidence for crosstalk when manual actions had to be performed simultaneously to action–word perception (Experiment 1) and also when language had to be produced during simultaneous perception of hand actions (Experiment 2). These findings suggest a clear bidirectional relationship between action and language. The latter crosstalk effect was due to interference between action and language (Experiment 3). By extending previous research of embodied cognition, the present findings provide novel evidence suggesting that bidirectional functional relations between action and language are based on similar conceptual–semantic representations.

The authors wish to thank Patricia Grocke for help with data acquisition.

R. Liepelt (✉)
Junior Group "Neurocognition of Joint Action",
Department of Psychology, Westfälische Wilhelms-University,
Fliednerstrasse 21, 48149 Muenster, Germany
e-mail: roman.liepelt@uni-muenster.de

R. Liepelt · T. Dolk · W. Prinz
Department of Psychology, Max Planck Institute for Human
Cognitive and Brain Sciences, Leipzig, Germany

Published online: 11 November 2011

Introduction

Recent findings from experimental psychology and cognitive neuroscience investigating language comprehension challenged the traditional view according to which language processing is mediated by an amodal symbolic system (Fodor, 1983). More recent research on embodied cognition assumes that language comprehension involves modal simulations that recruit the same neural systems that are usually used for action execution (Barsalou, 2007, 2008; Barsalou & Wiemer-Hastings, 2005; Pulvermüller, 2005; Gentilucci & Dalla Volta, 2008). Accordingly, one's own representations that are captured during experiences of action execution can be reactivated to internally simulate a specific action or associations with that action (Barsalou, 2007, 2008; Barsalou & Wiemer-Hastings, 2005). Most evidence for embodied and action-grounded cognition comes from brain research suggesting that modal simulation processes are mediated by the mirror-neuron system (D'Ausilio, Pulvermüller, & Salmas, 2009; Mahon and Caramazza, 2008; Pulvermüller, 2005; Pulvermüller & Fadiga, 2010; Rizzolatti & Arbib, 1998, 1999; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996; Rizzolatti, Fogassi, & Gallese, 2001).

One core region of the mirror-neuron system in monkeys is the premotor cortex (area F5), which is considered to be the homolog of Broca's area in the human brain (Rizzolatti & Arbib, 1998), the main region for language production. A number of brain imaging studies have shown that action observation (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Iacoboni et al., 1999; Rizzolatti & Craighero, 2004) and action imitation (Grèzes, Armony, Rowe, & Passingham, 2003; Iacoboni et al., 1999; Koski et al., 2002) also involves the Broca's area. Since the Broca's area seems to be involved not only in language production, but was also found to be activated in action execution, the assumption of a close functional connection between action and language mediated by the mirror-neuron system seems likely (Iacoboni & Wilson, 2006; Nishitani, Schürmann, Amunts, & Hari, 2005).

Although there is increasing evidence for a close functional relation between action and language, the precise nature of this relation is still poorly understood. While many previous studies have provided evidence for action–language interactions in one direction or another (Buccino et al., 2005; Glenberg & Kaschak, 2002; Pulvermüller, 2005), few have attempted to determine if these interactions are clearly bidirectional. Further, it is unclear whether such interactions lead to reciprocal facilitation or interference effects.

Some previous studies provided evidence for facilitation effects between action and language (Glenberg & Kaschak, 2002; Pulvermüller, 2005) by showing facilitation of action execution when sentences were shown describing an action in a congruent direction. Accordingly, facilitation was suggested to result from priming effects of action words on the motor production system (Boulenger et al., 2006). However, other studies, as for example a study of Buccino and colleagues, showed that congruent lexical–semantic action sentence pairings lead to effector-specific interference (Buccino et al., 2005) and not facilitation, as would be suggested by the study of Glenberg & Kaschak (2002). In the Buccino study, participants were slower in hand responses when sentences involved semantically congruent hand actions suggesting that interference may result from competition for common resources when both domains use the same representations (Chersi, Thill, Ziemke, & Borghi, 2010; Sato, Mengarelli, Riggio, Gallese, & Buccino, 2008). As the timing between action and language was not identical in previous studies, a promising account for bridging the gap between these competing findings was recently provided by Chersi et al. (2010). These authors highlighted the temporal relationship between action execution and language processes as a relevant factor for the polarity of the interaction effect. Following this line, the computational model by Chersi et al. (2010) predicts interference for simultaneous processing of action and language, while delayed processing is thought to produce facilitation.

Even more in line with the embodied cognition theory (Barsalou, 2007), Gentilucci and colleagues provided evidence that effects of hand action on speech articulators may be of a much more direct nature as previously thought (Gentilucci & Dalla Volta, 2008; Gentilucci, Dalla Volta, & Gianelli, 2008). Instead of arguing that gesture and speech are two different communication systems with gesture only supporting verbal expression (Levelt, Richardson, & La Heij, 1985; Krauss, Morrel-Samuels, & Colasante, 1991; Krauss, Dushay, Chen, & Rauscher, 1995), the authors argue that action (gesture) and language (speech) form a single communication system linked to the same thought processes and differing only in the expression modality (Gentilucci et al., 2008; McNeill, 1992). According to Gentilucci and colleagues, action and language merge in Broca's area. Both the production and observation of speech (Rizzolatti et al., 1996), as well as the execution and observation of hand gestures (Decety et al., 1994), activate the Broca's area (Gentilucci et al., 2008). By showing that a) hand shapes of various kinds affect speech articulation and b) the stimulation of left-sided Broca's area with repetitive transcranial magnetic stimulation (rTMS) leads to suppression effects of gesture observation, Gentilucci and colleagues concluded that the two systems controlling hand and mouth action may directly interact in Broca's area. In line with this reasoning, the authors assume a direct effect of hand action on the speech articulators (Gentilucci & Dalla Volta, 2008). This perspective should not only predict the direct effects of action perception on language production (unidirectional crosstalk), but also the direct inverse effect of language processing on action execution (bidirectional crosstalk). This appears particularly tempting to investigate, as the strongest attention in the current research gained the effects of language comprehension on action execution using either concrete (Buccino et al., 2005; Glenberg & Kaschak, 2002; Pulvermüller, 2005) or more abstract language (Glenberg, Sato, Cattaneo, Riggio, et al., 2008). Clear behavioral evidence, testing effects of language comprehension on manual action execution, as well as effects of hand action perception on language production within the same study is still missing.

The present study

The present study is designed to specify the functional relation between action and language in a series of three behavioral experiments. Using a comparable paradigm, Experiments 1 and 2 are aimed to test if the functional interaction between action and language is bidirectional. In Experiment 3, we test if action–language interactions lead to facilitation (for corresponding action–word pairings), interference (for non-corresponding action–word pairings), or both.

Behaviorally, interactions between perception and action can be measured on a fine-grained level with a direct matching paradigm previously used for measuring perception–action links in imitation research, in which participants have to carry out an instructed response indicated by a symbolic stimulus while observing task-irrelevant congruent or incongruent actions (e.g., Blakemore & Frith, 2005; Brass, Bekkering, Wohlschläger, & Prinz, 2000). It has been demonstrated that the observation of a corresponding hand movement leads to facilitation of the same motor response, while a non-corresponding hand movement leads to interference (Bertenthal, Longo, & Kosobud, 2006; Brass et al., 2000; Press, Bird, Flach, & Heyes, 2005; Stürmer, Aschersleben, & Prinz, 2000). In the present study, we adapted this paradigm to measure the functional interactions between action and language replacing manual actions in the input modality by action words (Experiment 1) or in the output modality by semantically corresponding and non-corresponding vocal responses (Experiments 2 and 3).

Experiment 1

In Experiment 1, we aimed to create a behavioral setup to test possible crosstalk between language perception and action execution, extending previous findings showing priming effects of action perception on language processing (Pulvermüller, 2005; Gentilucci & Dalla Volta, 2008). Using a choice-reaction task, we tested whether perceived action words automatically produce differential crosstalk effects on semantically corresponding or non-corresponding hand actions.

If the systems that control action and language directly interact in the Broca's area, as suggested by the work of Gentilucci & Dalla Volta (2008), the mere observation of an action word should automatically activate a corresponding gesture production.

Methods

Participants

Twenty healthy undergraduate students (10 female; 20–28 years, mean age = 23.5) with no history of neurological disorders participated in Experiment 1. The participants were all right-handed as assessed by the Edinburgh Inventory (Oldfield, 1971), had normal or corrected-to-normal vision and normal hearing, and were naive with regard to the hypothesis of the experiment. They were paid €7 for participating.

Apparatus and stimuli

Stimuli were presented on a 17-in. color monitor that was connected to a Pentium I PC. Experiments were carried out using ERTS software (Experimental Runtime System; Beringer, 2000). As stimuli, we used action words. The first screen showed a neutral mask "XXXXXX" (see Fig. 1). This neutral mask was followed by a second picture showing either the German word "öffnen" (open) or "schließen" (close). Both words were displayed either in red or green. At a viewing distance of 80 cm, the words subtended a visual angle of $6.2° \times 0.6°$ (neutral), $2.5° \times 0.5°$ (open), and $3.9° \times 0.5°$ (close).

Procedure and design

Participants had to execute a hand opening or closing action in response to word color. Green required participants to open their right hand, whereas red required participants to close their right hand. We measured responses with an optical response board connected to the test computer. In congruent trials, the word and the required response were semantically identical (open–open; close–close). In incongruent trials, the word and the required response differed (open–close; close–open). Participants were instructed to respond to the color irrespective of the content of the word.

The experiment consisted of four blocks of 64 trials. All four combinations of perceived and to be executed movements were randomly presented within each block. Between blocks, a short break was given. In total, participants performed 256 trials. Each trial began with a picture showing the neutral mask lasting for 1,500 ms. The second picture showed either the German word "öffnen" (open) or "schließen" (close) for 1,500 ms, either colored in red or green. Participant had to respond as quickly and accurately as possible by producing an opening (green) or closing action (red) when perceiving the color of the word. After a response was given or 2,500 ms had passed, a blank screen was presented for 1,500 ms. In total, each trial lasted for 5,500 ms.

Results

Reaction times

Prior to statistical analyses, all trials in which responses were incorrect (0.8%) were excluded from statistical reaction time (RT) analyses. Mean RTs for the 20 participants were submitted to a repeated measure analysis of variance (ANOVA) including the two-level within-subjects factor *Congruency* (congruent, incongruent). The analysis (see Fig. 2) showed a significant main effect of *Congruency*, $F(1, 19) = 4.52$, $p < 0.05$, $\eta^2 = 0.19$, showing that responses were faster for congruent (451 ms) than for incongruent trials (461 ms).

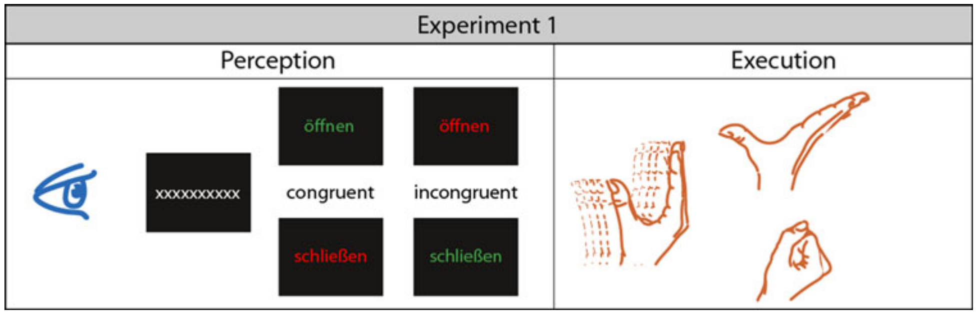

Fig. 1 Stimuli and design used in Experiment 1, showing the perceived word stimuli starting with the neutral letter stimulus, followed by congruent and incongruent color–word combinations (Perception, *left panel*) and required manual responses (Execution, *right panel*)

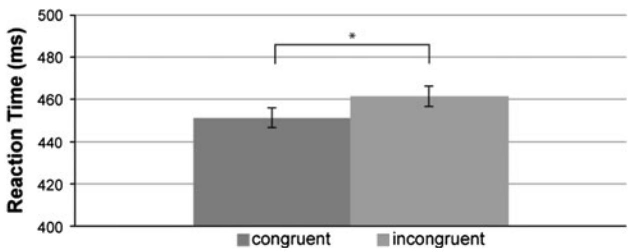

Fig. 2 Mean reaction time as a function of the induced type of congruency (congruent, incongruent) in Experiment 1, $*p < 0.05$. *Error bars* represent standard errors (SE) of the mean differences

Error rates

Errors did not differ significantly between congruent and incongruent trials, $F(1, 19) = 3.07$, $p > 0.05$, $\eta^2 = 0.14$.

Discussion

In Experiment 1, we found evidence for crosstalk between action and language when participants visually perceived task-irrelevant action words (open and close) and had to respond with a manual opening or closing gesture to word color only. In line with previous findings (Glenberg and Kaschak, 2002), the results provide evidence for automatic crosstalk between action and language. Yet, the mere observation of an action word seems to automatically activate a corresponding gesture in action production. In accordance with the work of Gentilucci & Dalla Volta (2008), the present findings suggest a direct interaction between the two systems that control action and language probably taking place in the Broca's area (Gentilucci et al., 2008).

Error analyses indicated that these findings were not due to a speed–accuracy tradeoff. However, it remains unclear in Experiment 1 whether crosstalk between action and language is functionally bidirectional.

Experiment 2

In Experiment 2, we investigated whether crosstalk between action and language was functionally bidirectional. In particular, we tested whether perceived manual actions automatically produce crosstalk on speech production. Participants had to say open or close in response to a green or a red cue appearing above a task-irrelevant human hand model, which either performed an opening or closing action.

In Experiment 1, we showed that visually perceiving action words automatically leads to crosstalk with semantically corresponding or non-corresponding gesture production. If semantic crosstalk, as found in Experiment 1, is bidirectional, one should also find similar crosstalk effects when participants perceive gestures and have to produce semantically corresponding or non-corresponding verbal responses. Such an effect would even more strongly support the assumption of a direct effect of hand action on the speech articulators (Gentilucci & Dalla Volta, 2008).

Methods

Participants

Twenty new healthy undergraduate students (10 female; 21–27 years of age, mean age = 24.7) with no history of neurological disorders participated in Experiment 2. The participants were all right-handed as assessed by the Edinburgh Inventory (Oldfield, 1971), had normal or corrected-to-normal vision and normal hearing, and were naive with regard to the hypothesis of the experiment. They were paid €7 for participating.

Apparatus and stimuli

The Apparatus was similar to Experiment 1. However, the stimuli we used in Experiment 2 were hand opening and closing actions that were modeled after a study of Press et al., (2005). The sequence of hand movements consisted of two consecutively presented pictures of a right human hand positioned in an egocentric perspective (see Fig. 3), creating the impression of a human hand movement. At a viewing distance of 80 cm, the hands subtended a visual angle of 9.2° × 14.7° (neutral), 18.5° × 14.4° (open), and

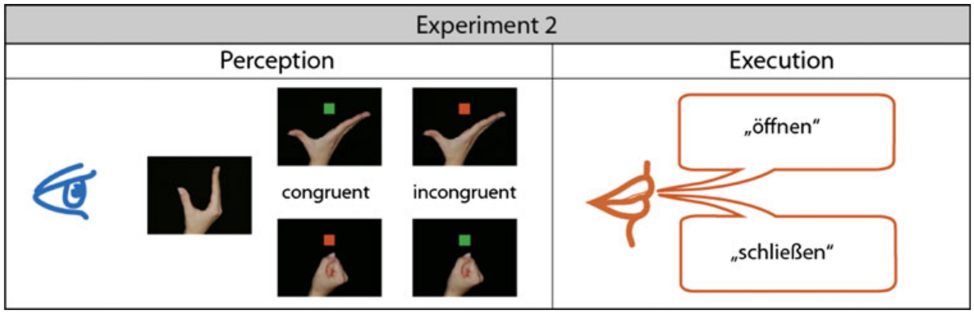

Fig. 3 Stimuli and design used in Experiment 2, showing perceived hand stimuli in a neutral starting position followed by congruent and incongruent cue color–hand position combinations (Perception, *left panel*) and required vocal responses (Execution, *right panel*)

9.4° × 12.0° (closed). A colored square (green or red; 2.5° × 2.5°), which always appeared at the same position above the hand, served as the response cue.

Procedure and design

Participants had to say either "öffnen" or "schließen" (German for "open" and "close") in response to a colored square (green, "open"; red, "close"). Simultaneous to the appearance of the colored square, either an opened or closed hand was presented which was either semantically corresponding or non-corresponding to the required response indicated by the color of the square (see Fig. 3). While in Experiment 1, the action-related stimulus (word) itself provided the critical color dimension for responding, in Experiment 2, the task-relevant color cue was presented closely above the hand. In congruent trials, the observed hand and the required vocal response were semantically identical (open–open; close–close). In incongruent trials, the observed hand and the required vocal response differed (open–close; close–open). Participants were instructed to respond to the color irrespective of the observed action. Vocal responses were recorded with a voice key connected to the computer.

The experiment consisted of four blocks of 64 trials. As in Experiment 1, all four combinations of perceived and to be executed movements were randomly presented within each block. Between blocks, a short break was given. In total, participants performed 256 trials. The trial timing was identical to Experiment 1.

Results

Reaction times

Prior to statistical analyses, all trials in which responses were incorrect (0.9%) were excluded from the statistical RT analyses. Mean RTs for the 20 participants were submitted to a repeated measure ANOVA including the two-level within-subjects factor *Congruency* (congruent, incongruent). The analysis (see Fig. 4) showed a significant main effect of *Congruency*, $F(1, 19) = 18.91$, $p < 0.001$, $\eta^2 = 0.50$, with faster responses for congruent (459 ms) than for incongruent trials (476 ms).

Error rates

Errors did not differ significantly between the different types of *Congruency*, $F(1, 19) = 2.37, p > 0.05, \eta^2 = 0.11$.

Discussion

The results of Experiment 2 were quite clear. Visually perceiving a hand opening or closing action automatically produced crosstalk on a vocal motor response (saying open or close). Semantic correspondence between the observed hand movement and the vocal response produced better performance than non-correspondence. Our findings are in line with previous work of Glenberg & Kaschak (2002), providing evidence for crosstalk between action and language. The findings of Experiment 2 show a direct effect of hand action on speech articulators (Gentilucci & Dalla Volta, 2008). Interestingly, our findings are also in line with recent findings from imitation research showing evidence for crosstalk between a visually perceived hand action and manual action production (Bertenthal et al., 2006; Brass et al., 2000; Press et al., 2005). However, the present findings extend such crosstalk effects to the language domain. Taken together, the findings of Experiments 1 and 2 suggest that crosstalk between action and language is functionally bidirectional.

On a more theoretical level, our findings are in good accordance with embodied cognition theories (Barsalou, 2007, 2008; Gentilucci & Dalla Volta, 2008). Language comprehension seems to involve modal simulations that recruit the same neural systems that are usually used for action execution (Barsalou, 2007, 2008; Barsalou & Wiemer-Hastings, 2005; Pulvermüller, 2005; Gentilucci & Dalla Volta, 2008). One's own representations that are captured during previous action experiences seem to be reactivated to internally simulate a specific action or conceptual associations with that action. Yet, the present

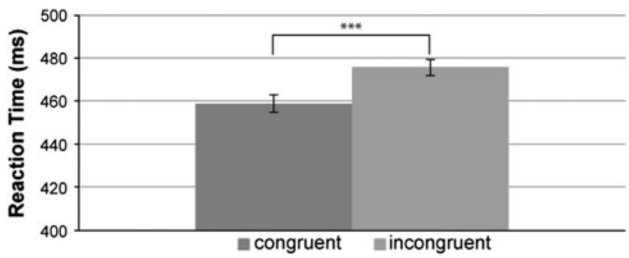

Fig. 4 Mean reaction time as a function of the induced type of congruency (congruent, incongruent) in Experiment 2, ***$p < 0.001$. *Error bars* represent standard errors (SE) of the mean differences

findings provide convergent evidence that action and language make use of the same motor representations considered to be part of the human mirror-neuron system (Grèzes et al. 2003; Iacoboni et al., 1999; Rizzolatti & Arbib, 1998).

However, the previous experiments do not allow conclusions about the question of whether crosstalk between action and language is due to facilitation for corresponding action–word pairings and/or interference for non-corresponding action–word pairings. To directly test if crosstalk effects observed in the present study are due to facilitation or interference, we performed a final experiment using a similar paradigm as in Experiment 2.

Experiment 3

Experiment 3 aimed to test whether crosstalk between action and language results from facilitation priming or interference between the perceived stimuli and the to be executed response. Under congruent conditions one may assume that the transformation of the stimulus color in a verbal motor representation takes place automatically with no need for further response selection, because the corresponding motor representation is already pre-activated by the prime. However, if an incongruent motor representation is externally triggered by the perceived action, a time-consuming reevaluation process may be required to separate the intended verbal response from the externally triggered response (Liepelt, von Cramon, & Brass, 2008a). Facilitation priming would be expected when perceived actions correspond to the action word that has to be produced in response to the color cue (Glenberg & Kaschak, 2002). Alternatively, crosstalk may result from direct semantic interference for incongruent action–word pairings (Gentilucci & Dalla Volta, 2008). To test these predictions we added a neutral control condition to the design of Experiment 2, in which the perceived hand remained stationary in a neutral position during the entire trial. As in congruent and incongruent conditions, participants had to select a verbal response based on cue color only.

If crosstalk is due to facilitation, then we should find significantly faster RTs for congruent as for neutral conditions. If crosstalk is due to direct interference for incongruent action–word pairings, then we should find significantly slower RTs for incongruent as for neutral conditions.

Methods

Participants

Twenty new healthy undergraduate students (10 female; 20–27 years, mean age = 24.9) with no history of neurological disorders participated in Experiment 3. The participants were all right-handed as assessed by the Edinburgh Inventory (Oldfield, 1971), had normal or corrected-to-normal vision and normal hearing, and were naive with regard to the hypothesis of the experiment. They were paid €7 for participating.

Apparatus and stimuli

Apparatus and stimuli were identical to those used in Experiment 2, except that we added a control condition: The neutral hand position, presented as the first picture in Experiment 2, was now once again presented simultaneously to the appearance of the colored square (red or green).

Procedure and design

The procedure and design were identical to Experiment 2. In Experiment 3, we used four blocks of 96 trials separated by short breaks. Thus, participants performed 384 trials in total.

Results

Reaction times

Prior to statistical analyses, all trials in which responses were incorrect (0.6%) were excluded from statistical RT analyses. Mean RTs for the 20 participants were submitted to a repeated measure ANOVA including the three-level within-subjects factor *Congruency* (congruent, neutral, incongruent). The analysis showed a significant main effect of *Congruency*, $F(2, 38) = 6.30, p < 0.01, \eta^2 = 0.25$, showing that responses differed between the three conditions (see Fig. 5). Responses were significantly slower for incongruent (465 ms) than for neutral (455 ms), $F(1, 19) = 6.86$, $p < 0.05$, $\eta^2 = 0.27$ and congruent trials

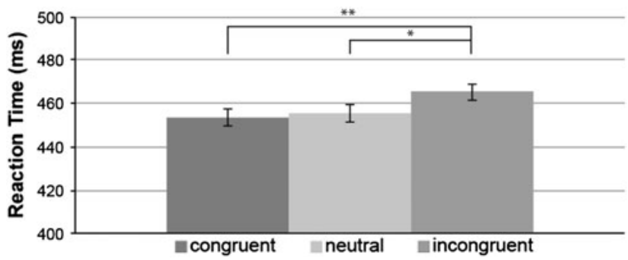

Fig. 5 Mean reaction time as a function of the induced type of congruency (congruent, neutral, incongruent) in Experiment 3, *$p < 0.05$, **$p < 0.01$. *Error bars* represent standard errors (SE) of the mean differences

(453 ms), $F(1, 19) = 9.25$, $p < 0.01$, $\eta^2 = 0.33$. However, responses for congruent compared to neutral trials did not differ significantly, $F(1, 19) = 0.35$, $p > 0.05$, $\eta^2 = 0.02$.

Error rates

Errors did not differ significantly between the different types of *Congruency*, $F(2, 38) = 0.13$, $p > 0.05$, $\eta^2 = 0.06$.

Discussion

In Experiment 3, we replicated the finding of crosstalk between action and language found in Experiment 2, as indicated by a significant difference between incongruent and congruent conditions. Further, we found significant differences in response times between incongruent and neutral conditions, but not between congruent and neutral conditions. These findings are in line with the assumption that crosstalk between action and language results from direct interference for incongruent action–word pairings. Perceived hand actions seem to directly affect the speech articulators (Gentilucci & Dalla Volta, 2008). Thus, when an incongruent motor representation is externally triggered by the perceived manual action, a time-consuming reevaluation process seems to be required to separate the intended verbal response from the externally triggered response. In contrast, our findings do not support the assumption of facilitation priming due to pre-activated speech representations by the prime, when action words describe an action in a congruent direction to the action word that has to be produced. However, further research has to be performed testing that the lack of facilitation effects in Experiment 3 is not due to generally fast response times, so that no further facilitation can be reached by the congruent prime.

In summary, perceiving an action seems to automatically activate a conceptual–semantic representation, which is not be restricted to a specific output modality. Similar conceptual–semantic representations seem to be used for action execution and language production (Gentilucci & Dalla Volta, 2008; Bernadis & Gentilucci, 2005).

General discussion

Previous studies provided evidence for a close functional relation between action and language. The aim of the present study was to specify the precise nature of this relationship. In a series of three behavioral experiments, we measured if the functional interactions between action and language were bidirectional (Experiments 1 and 2). Further, we tested if crosstalk effects resulting from these interactions were due to facilitation or interference (Experiment 3). Our findings suggest that the motor and the language system share conceptual–semantic representations.

In Experiment 1, we extended previous findings showing evidence for crosstalk between action and language (Buccino et al., 2005; Glenberg & Kaschak, 2002; Pulvermüller, 2005; Rueschemeyer, Lindemann, van Rooij, van Dam & Bekkering, 2010; Springer & Prinz, 2010). Visually perceiving an action word (open or close) automatically produced crosstalk on manual action execution by means of opening or closing a hand in response to the color of the word. Experiment 2 showed the inverse effect. Visually perceiving a hand opening or closing action automatically produced crosstalk on a vocal response saying open or close. Together, the results of Experiments 1 and 2 clearly suggest that crosstalk between action and language is functionally bidirectional (Rueschemeyer et al., 2010; Tettamanti et al., 2005). Action execution interacts automatically with language perception and the perception of action modulates language production. While in Experiment 1, the action-related stimulus (word) itself showed the critical color dimension for responding, the response dimension in Experiment 2 was located in an external, not action-related stimulus (square). While we do not think that this difference is critical in terms of the underlying mechanism producing this interaction, placing the critical response dimension directly in the action-related stimulus may strengthen the underlying crosstalk effect; an issue that should be investigated more closely in future research.

In Experiment 3, we tested whether crosstalk between action and language was due to facilitation of semantically corresponding or interference of semantically non-corresponding action word pairings. While previous studies found either evidence for facilitation (Rueschemeyer et al., 2010) or interference (Buccino et al., 2005) for semantically corresponding action–word pairings, we found evidence for interference for non-corresponding action–word pairings. Our findings seem to be in contrast with the view that general competition for common resources is the main source of crosstalk, at least when action and language have to be processed simultaneously (Buccino et al., 2005), as we found no evidence for interference when actions and words comprised a congruent semantic relation competing for the same representation. However, the present results are in line with the assumption that crosstalk under simultaneous action–language processing requirements results from direct lexical–semantic interference (Chersi et al., 2010) for incongruent action–word pairings (Gentilucci et al., 2008; Pulvermüller, 2005).

The use of simultaneous processing of perception and action execution in the present study had the advantage that temporal dynamics of crosstalk were kept constant across all experiments. However, temporal dynamics were recently discussed as possible factors for the polarity of emerging interaction effects between action and language (Chersi et al., 2010). A computational model by Chersi and colleagues predicts interference for simultaneous processing of action and language, while delayed processing is thought to produce facilitation (Chersi et al., 2010). The former prediction is in line with the finding of interference that we observed in Experiment 3 under simultaneous processing conditions of action and language. The latter prediction represents an interesting extension for future research using a similar direct matching paradigm as it was used in the present study, but adding a temporal delay between perception and action execution.

Further, our findings support the assumption that action and language are mediated by higher-level semantic representations (Meltzoff and Moore, 1997) that are not restricted to a specific modality. Action (gesture) and language (speech) seem to form a single communication system linked to the same thought processes differing only in expression modality (Gentilucci et al., 2008; McNeill, 1992). In this respect, our findings are also in line with the assumption that the two systems controlling hand and mouth actions directly interact in one brain region, probably the Broca's area (Gentilucci & Dalla Volta, 2008). While Experiment 1 showed a direct effect of action–word processing on hand action, Experiments 2 and 3 showed the inverse, a direct effect of perceived hand action on the speech articulators (Gentilucci & Dalla Volta, 2008). Taken together, the findings of all three experiments support embodied cognition theories (Barsalou, 2007, 2008) assuming that language comprehension involves modal simulations that recruit the same neural systems usually used for action execution.

The question if motor activity actually plays a constitutive role for language meaning cannot be fully answered with the present data set. Rather than arguing that movement itself is meaning, one might also argue that movement is associated with meaning. Perceiving action words, as well as perceiving actions, might through association automatically generate motor codes that interfere with action execution (Heyes, 2001). The latter view would be in line with the recent work of Glenberg and colleagues showing that use-induced motor plasticity affects the processing of abstract and concrete language (Glenberg, Sato, & Cattaneo, 2008).

The present work has also some interesting implications for the closely related field of imitation research (Bertenthal et al., 2006; Brass et al., 2000; Liepelt et al., 2008a, Liepelt, von Cramon, & Brass, 2008b; Press et al., 2005; Stürmer et al., 2000). One possible way to think about automatic imitation effects may also be in terms of conceptual equivalence representations that mediate between perception and action. What is needed for such an extension is the general idea that a stimulus currently available and the response that has to be executed (and that can only be anticipated during the time of stimulation) are represented at the same representational level. The present experiments may suggest that this representational level is of a more general conceptual nature than previously assumed.

Moreover, the present findings seem to be in line with recent findings from transcranial magnetic stimulation and lesion studies. These studies show that real and virtual lesions affecting motor regions lead to a disruption of phoneme comprehension and the understanding of semantic categories and grammar (D'Ausilio et al., 2009; Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Pulvermüller, 2005; Pulvermüller & Fadiga, 2010). Further support for this view is provided by behavioral studies showing that participants have difficulties making a sensible judgment when a response is required in the opposite direction (toward or away from the body) than implied by the meaning of a sentence (Glenberg & Kaschak, 2002). These authors proposed a set of three processes in which words and syntax were transformed into action-based meaning. First, words are indexed to analogical perceptual symbols based on the brain states underlying the perception of the referent. Second, affordances are derived from the perceptual symbols. Finally, affordances are meshed under the guidance of syntactic constructions (Kaschak & Glenberg, 2000). The present study seems to suggest that, at least under the present task conditions, a transformation of language into action-based meaning may also be achieved in a much more direct and automatic way (Bernadis & Gentilucci, 2005; Gentilucci & Dalla Volta, 2008; Tettamanti et al., 2005) via conceptual equivalence representations (Meltzoff & Moore, 1997).

Altogether, the findings from the present study suggest that the motor system contributes critically to conceptual–semantic processing. Our findings seem to support the assumption that the development of conceptual meaning through gestures may have paved the way for a more flexible vocalization system, a system for language (Rizzolatti & Arbib, 1998).

References

Barsalou, L. W. (2007). Grounding symbolic operations in the brain's modal systems. In G. R. Semin & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 9–42). New York: Cambridge University Press.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645.

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher & R. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thought* (pp. 129–163). New York: Cambridge University Press.

Beringer, J. (2000). Experimental Runtime System (Version 3.32). [Computer software]. Frankfurt am Main, Germany: BeriSoft Corporation.

Bernardis, P., & Gentilucci, M. (2005). Speech and gesture share the same communication system. *Neuropsychologia, 44*, 178–190.

Bertenthal, B. I., Longo, M. R., & Kosobud, A. (2006). Imitative response tendencies following observation of intransitive actions. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 210–225.

Blakemore, S. J., & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia, 43*, 260–267.

Boulenger, V., Roy, A. C., Paulignan, Y., Déprez, V., Jeannerod, M., & Nazir, T. A. (2006). Cross-talk between language processes and overt motor behaviour in the first 200 ms of processing. *Journal of Cognitive Neuroscience, 18*, 1607–1615.

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain Cognition, 44*, 124–143.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, B., & Rizzolatti, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research, 24*, 355–363.

Chersi, F., Thill, S., Ziemke, T., & Borghi, A. M. (2010). Sentence processing: linking language to motor chains. *Frontiers in neurorobotics, 4*, 4.

D'Ausilio, A., Pulvermüller, F., Salmas, P., et al. (2009). The motor somatotopy of speech perception. *Current Biology, 19*, 381–385.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadardy, B., Woods, R., et al. (1994). Mapping motor representations with positron emission tomography. *Nature, 371*, 600–602.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European Journal of Neuroscience, 15*, 399–402.

Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology*. Cambridge: MIT Press.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593–609.

Gentilucci, M., & Dalla Volta, R. (2008). Spoken language and arm gesture are controlled by the same motor control system. *Quarterly Journal of Experimental Psychology, 61*, 944–957.

Gentilucci, M., Dalla Volta, R., & Gianelli, C. (2008). When the hand speak. *Journal of Physiology–Paris, 102*, 21–30.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*, 558–565.

Glenberg, A. M., Sato, M., & Cattaneo, L. (2008a). Use-induced motor plasticity affects the processing of abstract and concrete language. *Current Biology, 18*, R290–R291.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008b). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology, 61*, 905–919.

Grèzes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurones in the human brain: An fMRI study. *Neuroimage, 18*, 928–937.

Heyes, C. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences, 5*, 253–261.

Iacoboni, M., & Wilson, S. M. (2006). Beyond a single area: Motor control and language within a neural architecture encompassing Broca's area. *Cortex, 42*, 503–506.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science, 286*, 2526–2528.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language, 43*, 508–529.

Koski, L., Wohlschläger, A., Bekkering, H., Woods, R. P., Dubeau, M.-C., Mazziotta, J. C., et al. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cerebral Cortex, 12*, 847–855.

Krauss, R. M., Dushay, R. A., Chen, Y. S., & Rauscher, F. (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology, 31*(6), 533–552.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology, 61*(5), 743–754.

Levelt, W. J., Richardson, G., & La Heij, W. (1985). Pointing and voicing in deictic expressions. *Journal of Memory and Language, 24*, 133–164.

Liepelt, R., von Cramon, D. Y., & Brass, M. (2008a). What is matched in direct matching? Intention attribution modulates motor priming. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 578–591.

Liepelt, R., von Cramon, D. Y., & Brass, M. (2008b). How do we infer others' goals from non-stereotypic actions? The outcome of context-sensitive inferential processing in right inferior parietal and posterior temporal cortex. *NeuroImage, 43*, 784–792.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology–Paris, 102*, 59–70.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.

Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting, 6*, 179–192.

Nishitani, N., Schürmann, M., Amunts, K., & Hari, R. (2005). Broca's region: From action to language. *Physiology, 20*, 60–69.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia, 9*, 97–113.

Press, C., Bird, G., Flach, R., & Heyes, C. (2005). Robotic movement elicits automatic imitation. *Cognitive Brain Research, 25*, 632–640.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience, 6*, 576–582.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience, 11*, 351–360.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences, 21*, 188–194.

Rizzolatti, G., & Arbib, M. A. (1999). From grasping to speech: Imitation might provide a missing link: Reply. *Trends in Neurosciences, 22*, 152.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169–192.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. Brain Research. *Cognitive Brain Research, 3*, 131–141.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience, 2*, 661–670.

Rueschemeyer, S.-A., Lindemann, O., van Rooij, D., van Dam, W., & Bekkering, H. (2010). Effects of intentional motor actions on embodied language processing. *Experimental Psychology, 57*, 260–266.

Sato, M., Mengarelli, M., Riggio, L., Gallese, V., & Buccino, G. (2008). Task related modulation of the motor system during language processing. *Brain and Language, 105*, 83–90.

Springer, A., & Prinz, W. (2010). Action semantics modulate action prediction. *The Quarterly Journal of Experimental Psychology, 63*, 2141–2158.

Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: A study of imitation. *Journal of Experimental Psychology, 26*, 1746–1759.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., et al. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience, 17*, 273–281.