Linguistic Representations of Motion Do Not Depend on the Visual Motion System

Andrea Pavan[1,2] and Giosuè Baggio[1]

[1]International School for Advanced Studies (SISSA) and [2]Institut für Psychologie, Universität Regensburg

Abstract

Embodied semantics proposes that constructing the meaning of motion verb phrases relies on representations of motion in sensory cortex. However, the data reported by earlier studies as evidence for this claim are also explained by a symbolic-semantics view proposing interactions between dissociable systems. In the experiments reported here, participants were visually adapted to real and implied leftward or rightward motion, which produced a motion aftereffect opposite to the direction of the adapting stimulus. Participants then decided whether a directionally ambiguous or a leftward- or rightward-directional verb phrase implied leftward or rightward motion. Because the visual system is engaged in the motion aftereffect, embodied semantics predicts that responses in the motion-aftereffect direction (opposite to the direction of the adapting stimulus) are facilitated, whereas symbolic semantics predicts response facilitation in the direction of the adapting stimulus (opposite to the direction of the motion aftereffect). We found response facilitation in the direction of real- and implied-motion adapting stimuli in ambiguous and directional verb phrases. These results suggest that visual and linguistic representations of motion can be dissociated.

The meanings of complex linguistic expressions are seldom fixed, given, or stored in memory and must therefore be constructed using contextual cues and elementary representations for words or other linguistic units of suitable size. Two families of theories have proposed contrasting ways of how meanings are computed in the brain in relation to sensorimotor systems. In the embodied-semantics view, the neural representations for words and sentences are shared with sensory systems: For example, the meaning of a verb denoting a motion event, such as "to slide," is represented within the neuronal networks responsible for visual motion perception (Barsalou, 1999). In contrast, symbolic semantics proposes that meanings are instantiated as patterns of brain activity that can be dissociated from modality-specific structures: The semantics of "to slide" may be represented outside sensorimotor systems in supramodal cortical networks (Binder & Desai, 2011).

However, this opposition seems to fail to capture the full range of empirical data, as neither theory clarifies the exact nature of the interplay between semantics and sensory domains. The symbolic-semantics account is consistent with information flow between brain systems, provided that the representations are dissociable. Pervasive phenomena, such as spreading activation (Mahon & Caramazza, 2008), could explain the kind of sensorimotor responses observed during language-comprehension tasks, which have been taken as evidence for the embodied-semantics theory.

In a number of experiments (reviewed by Zwaan & Madden, 2005), participants were shown a sentence followed by pictures. The content of the sentence and of the pictures could match or mismatch. For example, in the mismatching condition, participants would hear "The pitcher hurled the softball to you," and would view two pictures of a softball whose size decreased from one frame to the next, as though it were moving away from the observer (Zwaan, Madden, Yaxley, & Aveyard, 2004). In the matching condition, participants might hear the same phrase but see a series of pictures in which the size of the softball increased. Reaction times were shorter in the matching than in the mismatching condition. Such a *match advantage* was taken to suggest that sensory representations are routinely activated during verbal comprehension and are

Corresponding Authors:
Andrea Pavan, International School for Advanced Studies (SISSA), via Bonomea 265, Trieste, 34136, Italy
E-mail: apavan@sissa.it

Giosuè Baggio, International School for Advanced Studies (SISSA), via Bonomea 265, Trieste, 34136, Italy
E-mail: gbaggio@sissa.it

made available for direct perceptual comparison with subsequent or concurrent visual stimuli (Zwaan & Madden, 2005). However, these data seem equally consistent with the possibility that linguistic stimuli largely activate conceptual knowledge, as required by symbolic views of meaning, and that the pictures are rapidly analyzed also at the semantic level, where a comparison or an integration with the verbal material is carried out. Another possibility is that priming occurs across systems enjoying some degree of functional and anatomical segregation, with no need for semantic mediation (Mahon & Caramazza, 2008).

The aim of the present experiments was to answer two questions. First, do visual and linguistic representations of motion overlap within the same sensory systems, or do they interact while remaining dissociable? And, second, does interaction occur at the perceptual or at the semantic level?

To answer the first question, we conceived an experimental paradigm in which the embodied-semantics theory (shared representations) and the symbolic-semantics account (dissociable interacting representations) predict effects in opposite directions, and no interaction between systems was the null hypothesis. Following prolonged exposure to unidirectional motion, participants perceived a stationary or directionally ambiguous test pattern as moving in the opposite direction to that of adaptation. This illusion of motion is known as a *motion aftereffect* (for a review, see Mather, Pavan, Campana, & Casco, 2008). Motion adaptation and the motion aftereffect have been used extensively to examine the spatiotemporal tuning of mechanisms operating during the early stages of visual motion processing (Bex, Verstraten, & Mareschal, 1996; Graham, 1989; Ledgeway & Hutchinson, 2009; Mareschal, Ashida, Bex, Nishida, & Verstraten, 1997).

One prominent account of the motion aftereffect is the opponent-process-coding model (Mather, Verstraten, & Anstis, 1998). In its simplest form, this model is a network of two layers of units. Motion sensors in the first layer provide an initial estimate of motion energy, and a pair of units in the second layer (the opponent-energy layer; Adelson & Bergen, 1985) receive opposed outputs from the first layer. One unit from the second layer codes, for instance, for leftward motion, and the other codes for rightward motion. Motion in one particular direction is perceived when the output of the opponent-energy unit selective for that direction exceeds a certain threshold. Before adaptation, all units are at the resting level. During adaptation to leftward motion, for instance, the leftward sensor in the first layer responds strongly, whereas the rightward sensor remains at resting level. This leads to an above-threshold response by the leftward opponent-energy unit, whereas responses are suppressed in the rightward unit. After adaptation, the resting level of the leftward sensor is strongly suppressed, whereas that of the rightward sensor is (nearly) unaffected. Their relative difference will be reflected in the opponent-energy layer (i.e., the network's second layer), with the leftward unit showing a suppressed response and the rightward unit producing an above-threshold response that leads to

the perception of a motion aftereffect (Mather et al., 1998). This dynamic of adaptation and suppression can be seen as a form of automatic gain control, in which a motion sensor selective for a particular motion direction attenuates its own response to continuous intense stimulation, thereby improving encoding efficiency by maximizing stimulus-related information conveyed in neural spike trains (Mather et al., 2008).

In our experiments, participants were exposed to unidirectional leftward- or rightward-moving adapting visual stimuli for 30 s. Adaptation was followed by a verb phrase presented binaurally. Verb phrases could be either directionally ambiguous (e.g., "To take a road at a T junction"), referring to a motion event that is equally likely to involve a rightward or leftward change of direction, or directional, involving either leftward motion ("To overtake a car driving too slowly") or rightward motion ("To write a sentence on a sheet"). The task was to indicate by means of a button press whether each verb phrase suggested rightward or leftward motion.

If embodied semantics is correct, that is, if interpreting motion verb phrases necessarily recruits the visual system, responses should be faster when the direction assigned to an ambiguous verb phrase matches the direction of the motion aftereffect and is opposite to the adapted direction than when it does not match the direction of the motion aftereffect and is the same as the adapted direction. The rationale is that if visual adaptation in direction *d* leads to an attenuation of responses in motion sensors coding for *d*, it should be more difficult to recruit those sensors (opponent-energy units) to represent any *d*-ward motion suggested by the verb phrase (as is required by embodied semantics), but it should be easier to recruit motion sensors at the resting level that code for the opposite direction (i.e., those involved in the perception of the motion aftereffect). Conversely, if the direction of adaptation is a more vivid and a longer-lasting visual cue than the motion-aftereffect direction, then interaction between dissociable systems predicts faster responses when the direction assigned to the verb phrase matches the direction of adaptation than when it does not.

To answer the second question—whether interaction between vision and language occurs at the perceptual or at the conceptual level—we used two types of visual stimuli that induced adaptation to real or implied motion, respectively. To induce adaptation to real motion, we used a Gabor patch drifting either leftward or rightward for 30 s (Fig. 1a). In implied-motion adaptation, we used 30-s sequences of pictures depicting a human or an animal moving either leftward or rightward (Pavan, Cuturi, Maniglia, Casco, & Campana, 2011; Fig. 1b). The two experiments used the same verbal stimuli, presented in the same manner. Implied-motion adaptation should affect mainly intermediate-level (e.g., medial temporal, or MT, in the human brain) and higher-level networks for biological motion perception or action representation (Proverbio, Riva, & Zani, 2009; Senior et al., 2000; Williams & Wright, 2010). Adaptation to real motion (i.e., translational motion)
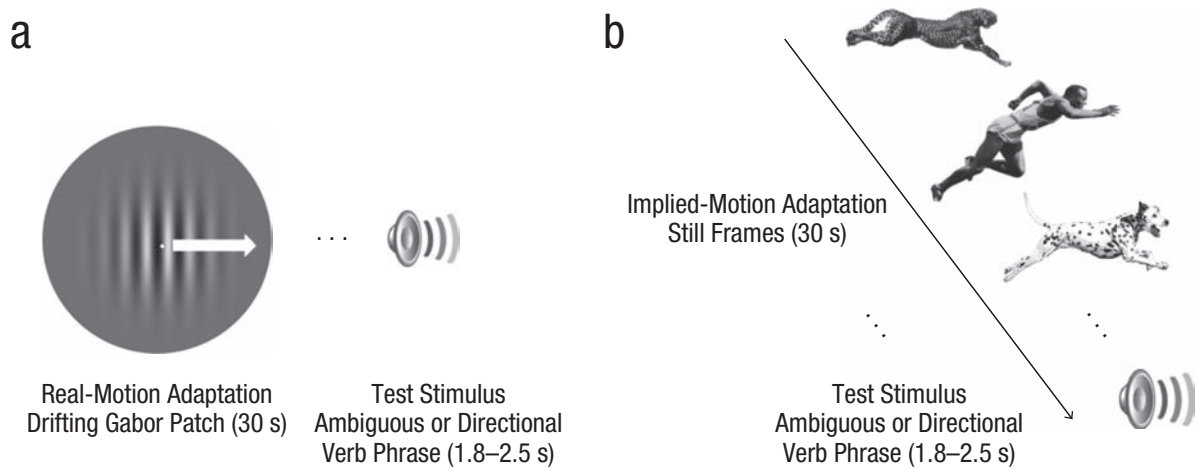
Fig. 1. Procedure for (a) the real-motion-adaptation experiment and (b) the implied-motion-adaptation experiment. For real-motion adaptation, participants saw a Gabor patch drifting leftward or rightward for 30 s. Immediately afterward, a verb phrase denoting an event was presented binaurally. Verb phrases could either be directionally ambiguous (e.g., "To take a road at a T junction") or suggest leftward or rightward movement (e.g., "To overtake a car driving too slowly"). For implied-motion adaptation, participants saw a sequence of photographs of moving people or animals for 30 s. Each picture lasted for 500 ms. The verbal test stimuli were the same as in the real-motion-adaptation experiment.

was expected to recruit mostly lower (e.g., V1/V2, the V3 complex; Georgieva, Peeters, Kolster, Todd, & Orban, 2009) and intermediate levels of motion processing (e.g., V5/MT). Implied motion activates motion-sensitive areas. However, the motion aftereffect following adaptation to implied motion has not been clearly demonstrated yet. To investigate this effect, we conducted a control experiment in which we assessed the presence of an implied motion aftereffect (Winawer, Huk, & Boroditsky, 2008).

We hypothesized that if the visual and language systems communicate at the level of perceptual analysis, the main experiments should show an effect in real-motion adaptation only, because actual motion is perceived only in that case. However, if vision and language interact at the semantic level, there should be little or no difference between real and implied motion, because the adapting direction can be extracted from either stimulus type and used in a similar manner as a contextual cue for conceptually analyzing the verbal material.

Method

Real- and implied-motion-adaptation experiments

Participants. Forty right-handed native Italian speakers took part in the two main experiments. Twenty participants (mean age = 27.5 years, *SD* = 4.55; mean education = 15.4 years, *SD* = 3.89) were tested using real-motion adaptation, and 20 participants (mean age = 25.25 years, *SD* = 3.23; mean education = 16.5 years, *SD* = 3.36) were tested using implied-motion adaptation. Participants were naive to the purpose of the experiments and had normal or corrected-to-normal visual

acuity. All had binocular vision and binaural hearing. Participants enrolled voluntarily for compensation and gave informed consent to participate.

Apparatus. Visual stimuli were generated with Psychtoolbox (Brainard, 1997; Pelli, 1997) for MATLAB (The MathWorks, Natick, MA) and shown on a 19-in. CRT monitor (60-Hz refresh rate; 1,280 × 1,024 resolution). Luminance was measured with a Minolta LS-100 photometer. Adapting visual stimuli were created using a gamma-corrected lookup table to ensure display linearity.

Visual stimuli. The real-motion adapting stimulus was a vertically oriented Gabor patch consisting of a stationary Gaussian kernel function modulated by a sinusoidal luminance wave (Gabor width = 8°; spatial frequency = 1 cycle per degree; Fig. 1a). The sine-wave carrier drifted either leftward or rightward at 5.45°/s. The Michelson contrast for the adapting Gabor stimulus was 0.63. The implied-motion adapting stimuli were 36 gray-scale photographs of a human or an animal oriented leftward or rightward—opposite orientations were obtained by flipping the same photograph on its vertical axis (Fig. 1b). Each image was initially scaled to fit an area subtending 10.6° × 7.7°. Other objects were removed from the background.

Linguistic stimuli. The verbal stimuli were 10 directionally ambiguous verb phrases and 10 verb phrases suggesting leftward or rightward motion (all were presented in Italian). Both verb-phrase types described human-directed events involving either a leftward or a rightward change of direction. Ambiguous phrases referred to events that could unfold in either direction (e.g., "*Imboccare una strada a un bivio*," "To take a road

at a T junction"). Directional verb phrases denoted events that entailed motion in one preferred direction (e.g., "*Sorpassare un'automobile troppo lenta*," "To overtake a car driving too slowly," which suggests leftward motion). Among the directional verb phrases, five involved rightward motion and five involved leftward motion. The verbs' word-form frequency was matched between verb-phrase types (according to the CORIS/CODIS corpus; Rossini Favretti, Tamburini, & De Santis, 2002). Mean verb-phrase length was 6.4 words (*SD* = 0.843) for ambiguous phrases and 5.7 words (*SD* = 1.059) for directional phrases (*W* = 71.5, *p* = .0878).

Phrases were spoken by a trained female phonetician who was instructed to maintain her reading pace and pitch across the recording session. The samples (16-bit mono, 44.1-kHz rate) were denoised using a representative background noise profile, direct-current offset was removed, and maximum amplitudes were normalized to −3 dB. All audio samples were truncated at the nearest multiple of 50 ms following the offset of the speech waveform. The duration range of the resulting traces was 1.8 s to 2.5 s for both verb-phrase types (ambiguous phrases: *M* = 2.070 s, *SD* = 0.262; directional phrases: *M* = 2.170 s, *SD* = 0.249), *t*(17.953) = −0.8763, *p* = .3924.

Procedure. At the start of each experimental session, the Italian versions of the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973) and the Vividness of Movement Imagery Questionnaire (VMIQ; Isaacs, Marks, & Russell, 1986) were administered. Participants sat in a dark room 57 cm away from the screen. A chin rest was used to maintain head position. In the real-motion-adaptation experiment, the monitor was set at a mean luminance of 52.75 cd/m^2. On each trial, participants were instructed to fixate a white circular spot (diameter = 0.21°) at the center of the Gabor patch, which occupied the center of the screen. Adaptation lasted for 30 s with the Gabor patch drifting leftward or rightward (randomized within blocks). Adaptation was immediately followed by a phrase presented binaurally using circumaural headphones. Starting at the acoustic onset of the phrase, the fixation point turned red for 4 s. During this interval, participants indicated whether the verb phrase suggested either leftward or rightward motion by pressing one of two designated keys on a standard keyboard. A recovery period of 35 s of uniform (mean luminance) screen with a white fixation point followed the response.

The procedure for adaptation to implied motion was the same as that for real-motion adaptation, with the following exceptions. The screen was set at the maximum luminance (101.05 cd/m^2). On each trial, an adaptation sequence was generated by displaying a sample of 36 randomly shuffled photographs without repetition. No photo was repeated until each had been shown once. Pictures were shown every 500 ms for a total period of 30 s. During adaptation to implied motion, the fixation point was not present (Winawer et al., 2008). Adaptation was immediately followed by a verb phrase presented binaurally. Starting at the acoustic onset of the verb phrase, a red fixation point appeared for 4 s. During this

interval, participants had to indicate whether the verb phrase suggested either leftward or rightward movement. After the response, there was a recovery period of 35 s of uniform (white) screen with a black fixation point.

In the two experiments, each verb phrase was used twice, following leftward-moving and rightward-moving adapting stimuli on different trials. The order in which the verb phrases were presented was randomized in each session.

At the end of each experimental session, we debriefed participants to make sure they were not familiar with motion adaptation, the motion aftereffect, the waterfall illusion (a motion-aftereffect illusion), and similar phenomena. In particular, we asked which way, after seeing rightward or leftward motion for 30 s, they would expect a stationary pattern to appear to move (Dils & Boroditsky, 2010).

Data analysis. For real-motion adaptation, the mean frequency of correct responses and mean reaction times were calculated separately for cases in which the direction indicated by the participant's response matched the direction of the adapting stimulus (matching trials) and for cases in which the direction of the participant's response mismatched the direction of the adapting stimulus (mismatching trials). For implied-motion adaptation, the mean frequency of correct responses and mean reaction times of correct responses (rightward-directional verb phrases were given a right response; leftward-directional verb phrases were given a left response) were analyzed separately for cases in which the direction of the adapting stimulus and response matched and for cases in which they mismatched. In each condition and for differences between conditions, mean reaction times and correct-response frequencies were correlated with VVIQ and VMIQ scores using Pearson's correlation tests; *p* values were Bonferroni corrected.

Control experiments

To test whether the visual stimuli used in the main experiments induced reliable motion adaptation, we conducted two control studies, in which we attempted to quantify the relative strength of the motion aftereffect.

Participants. Four right-handed volunteers participated in the control experiments: 2 participants (mean age = 27 years, *SD* = 2) were tested with real-motion stimuli, and 2 participants were tested with implied-motion stimuli (mean age = 29.25 years, *SD* = 4.27). Participants were naive to the purpose of the experiments and had normal or corrected-to-normal visual acuity.

Stimuli and procedure. Participants sat in a dark room 57 cm away from the screen. Adapting stimuli were the same as those used in the main experiments. The test stimuli were a dense spatial array of 264 white dots (101.05 cd/m^2) for the real-motion experiment and 264 black dots (0.2 cd/m^2) for the implied-motion experiment. Dots were displayed within a

circular window (diameter = 8°) at the center of the screen (dot density = 5.25 dots per degree^2). Each moving dot had a width of 0.1°. Dots had a speed of 1.77°/s. Each dot had a limited lifetime. When the age of a dot reached 83.3 ms, the dot was replaced by a new dot at a random position within the circular window. Dots traveling outside the circular window wrapped around to the opposite point on the circumference. The duration of each test pattern was randomized trial by trial between 1.8 s and 2.5 s (matching the duration range of the verb phrases in the main experiment) in 50-ms steps to test whether visual adaptation was effective during the entire period over which the verb phrases were presented.

The strength of the motion aftereffect was estimated using a motion-nulling technique (Ledgeway, 1994). That is, the percentage of dots coherently moving in the same direction as the adapting stimulus was varied to null the perception of a unidirectional motion aftereffect (Hiris & Blake, 1992; Ledgeway & Hutchinson, 2009). The percentage of coherently moving dots was varied from trial to trial according to a simple up-down staircase procedure (Levitt, 1971). Leftward and rightward adaptations were shown in separate blocks. Each control experiment had two stages: (a) estimating baseline performance (no adaptation) and (b) estimating the percentage of coherently moving dots that nulled the motion aftereffect.

In the first stage of each control experiment, we estimated the percentage of coherent motion below which each participant was unable to perceive directional motion. Leftward and rightward motion were presented in separate blocks. Participants performed a direction-discrimination task. The initial step size of the staircase was set at 10% coherence. After each reversal, step size was halved until 0.15% coherence was attained. The adaptive procedure could stop either after a total of 100 trials or after 12 reversals, whichever came first. The mean percentage of coherently moving dots across the last 8 reversals was taken as a threshold estimate.

In a second stage, participants were adapted for 30 s to drifting Gabor patches (real motion) or to still frames of moving humans or animals (implied motion). For each participant, the staircases started at the threshold percentage for coherent motion estimated in the first stage. After adaptation, participants had to judge whether dots moved either leftward or rightward. The initial step size was 5% coherence. After each reversal, the step size was halved until 0.15% coherence was attained. Within a block, the direction of the adapting stimuli was constant. To avoid any cumulative effects of adaptation, we set the intertrial interval to 35 s.

Results

Main experiments

Following real-motion adaptation, the frequency with which participants' responses to ambiguous phrases matched the direction of the adapting stimulus was at chance level (M = 5.225 out of 10 trials, SD = 0.87; Wilcoxon signed-rank test: V = 54.5,[1] p = .232). The same finding applied to mismatching responses (M = 4.625, SD = 0.87; V = 20, p = .0764). No difference was observed between the two data series (Wilcoxon rank-sum test: V = 75.5, p = .1552; Fig. 2a). Participants were faster when the direction indicated by their response matched the direction of the adapting stimulus (M = 2.902 s, SD = 0.45) than when it mismatched the direction of the adapting stimulus (M = 3.090 s, SD = 0.49). The mean difference in reaction times between mismatching and matching trials was 0.188 s (SD = 0.24; V = 26, p = .0058; Fig. 2a).

The frequency of correct responses to directional phrases following real-motion adaptation was significantly above chance in trials in which the direction indicated by the participant's response matched the direction of the adapting stimulus (M = 7.450, SD = 1.76; V = 203, p = .0002), as well as in trials in which there was a mismatch between response direction and adapting-stimulus direction (M = 7.400, SD = 1.57; V = 207, p = .0001). No significant difference between matching and mismatching trials was found (V = 54.5, p = .9232; Fig. 2b). Correct responses were faster in trials in which the direction indicated by the response matched the direction of the adapting stimulus (M = 3.025 s, SD = 0.59) than when it mismatched the direction of the adapting stimulus (M = 3.217 s, SD = 0.72). The mean reaction-time difference was 0.192 s (SD = 0.36; V = 46, p = .0289; Fig. 2b).

Following implied-motion adaptation, the mean response frequency to ambiguous verb phrases was at chance level in matching trials (M = 4.925, SD = 0.95; V = 48, p = .5027) as well as in mismatching trials (M = 5.075, SD = 0.95; V = 72, p = .5027). No significant difference between the two trial types was observed (V = 48, p = .5027; Fig. 2c). Reaction times were shorter when the direction indicated by the response matched that of the adapting stimulus (M = 2.744 s, SD = 0.43) than when it did not (M = 2.844 s, SD = 0.39). The mean reaction time difference was 0.100 s (SD = 0.20; V = 48, p = .0349; Fig. 2c).

For implied-motion adaptation, the frequency of correct responses to directional verb phrases was above chance level in matching trials (M = 8.100, SD = 0.79; V = 54, p = .0001) and in mismatching trials (M = 8.650, SD = 1.04; V = 210, p = .0001). The frequency of correct responses in mismatching trials was slightly higher than in matching trials (V = 26, p = .0430; Fig. 2d). Participants responded more quickly in trials in which the direction of the adapting stimulus and response matched (M = 2.680 s, SD = 0.37), compared with trials in which they did not (M = 2.810 s, SD = 0.45). The mean difference in reaction times was 0.130 s (SD = 0.22; V = 38.5, p = .0137; Fig. 2d). A higher percentage of correct responses in the slower condition (here, on mismatching trials) was found only in this comparison. This could reflect a speed-accuracy trade-off, which makes this result only prima facie consistent with the predictions of the embodied-semantics account (i.e., more correct responses in the direction of the motion aftereffect) and does not therefore compromise our main findings.
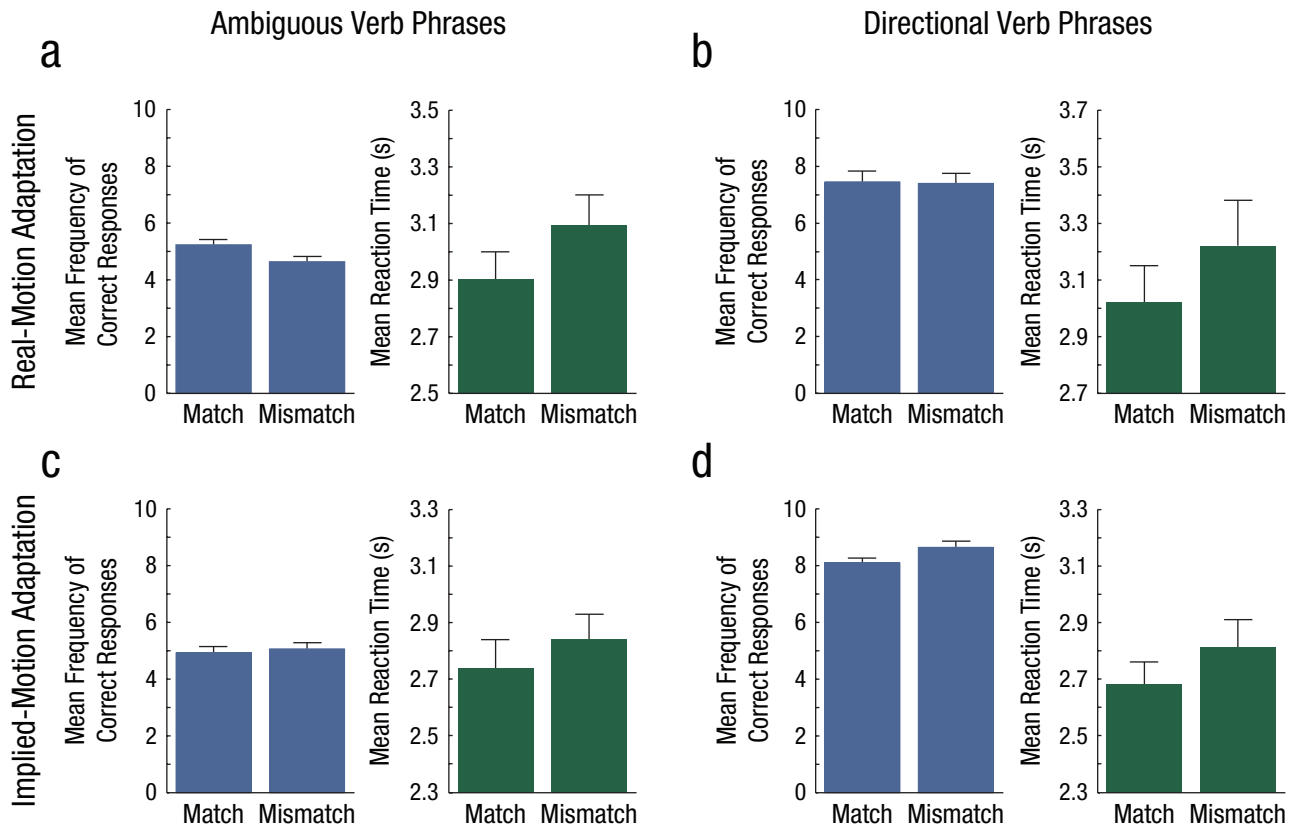
Fig. 2. Results of the real-motion-adaptation experiment (top row) and the implied-motion-adaptation experiment (bottom row). For each experiment, results are shown separately for ambiguous verb phrases (a, c) and directional verb phrases (b, d). The mean frequency of correct responses and the mean reaction time are shown for each type of verb phrase as a function of whether the direction indicated by the participants' responses matched or mismatched the direction of the adapting stimuli. Error bars indicate standard errors of the mean.

No correlation between VVIQ or VMIQ scores and reaction times or response frequencies was found. During debriefing, we asked participants what would be the perceived direction of a stationary pattern presented following adaptation to rightward or leftward motion, and we found that most of them were unable to predict the actual motion-aftereffect direction. This shows that participants were naive with respect to the aims of our experiment.

Control experiments

In the real-motion control experiment, 30 s of exposure to a drifting Gabor patch (5.45°/s) was sufficient to produce reliable motion adaptation. A high percentage of dots moving coherently in the same direction as the adapting stimulus was necessary to null the motion aftereffect and cause participants to perceive random motion (no-adaptation condition: $M = 3.35\%$, $SEM = 1.56$; adaptation condition: $M = 21.02\%$, $SEM = 6.21$), paired-samples $t(3) = -3.77$, $p = .033$.

In the implied-motion control experiment, 30 s of exposure to photos of leftward- and rightward-oriented humans or animals induced reliable adaptation, albeit weaker than the adaptation obtained using real-motion stimuli. A paired-samples

t test revealed a difference between the no-adaptation condition ($M = 1.63\%$, $SEM = 1.2$) and the adaptation condition ($M = 5.55\%$, $SEM = 1.01$), $t(3) = -4.74$, $p = .018$. In brief, in both the main and the control experiments, using real and implied motion, reliable adaptation was obtained, which covered the entire duration of the verb phrases used in the main experiments.

Discussion

In the experiments reported here, we investigated the effects of adaptation to real and implied visual motion on processing auditorily presented linguistic phrases suggesting leftward or rightward motion. Visual adaptation in one direction induced a motion aftereffect in the opposite direction. If language relies on visual systems for the interpretation of motion phrases, the motion aftereffect should facilitate the comprehension of a phrase that suggests movement in the same direction as the motion aftereffect and opposite to the direction of the adapting stimulus. However, we found clear-cut evidence to the contrary. In a task in which, following visual adaptation, participants decided whether a phrase suggested leftward or rightward movement, responses were faster if the phrase direction

was the same as the direction of the adapting stimulus than if it was in the opposite direction. This should not be mistaken with the match advantage reported in earlier studies (Zwaan & Madden, 2005).

Previous research was unable to disentangle shared representations between systems from dissociable interacting representations linked by priming or cuing (Mahon & Caramazza, 2008). We conducted our experiments precisely to distinguish these two possibilities. If representations are shared, as posited by the embodied-semantics hypothesis, then responses should be faster in the direction of the motion aftereffect than in the opposite direction. In contrast, if representations interact, responses should be faster in the direction opposite to rather than toward the motion aftereffect, because the direction of the adapting stimulus provides a longer-lasting and more vivid visual cue than does the direction of the motion aftereffect. In two experiments tapping into different levels of the visual-processing hierarchy, we demonstrated that the motion aftereffect has no influence on the comprehension of motion phrases and, thus, that representations of motion induced by visual and linguistic stimuli can be dissociated, though they may not be functionally unconnected.

Our findings are consistent with dissociations between language and visuospatial processing described in patients (Bek, Blades, Siegal, & Varley, 2010; Kemmerer & Tranel, 2000) and with experiments emphasizing the role of the task and context in producing sensorimotor activation during verbal processing (Papeo, Vallesi, Isaja, & Rumiati, 2009). Taken together with recent experimental work, our data point to a more complex picture than is proposed by either embodied cognition or its opposite, a modular view featuring strict informational encapsulation between systems (Fodor, 1983), which served as the null hypothesis in our experiments.

We found a match advantage in two visual-adaptation paradigms. Nonetheless, other studies with different tasks and stimulation conditions reported a match disadvantage instead (Kaschak et al., 2005). This latter finding, however, would seem consistent with an effect of the motion aftereffect on verb-phrase processing only if adaptation had been obtained, for which these studies provided no evidence. It could be that the lack of a motion aftereffect on verb-phrase comprehension in our experiments may be due to the fact that the direction of the adapting stimulus is a suprathreshold task-irrelevant signal that is attentionally filtered out. Tsushima, Sasaki, and Watanabe (2006) showed that suprathreshold task-irrelevant motion induces activity in medial temporal visual area but does not interfere with a concomitant visual task, which suggests that suprathreshold task-irrelevant signals may be suppressed: Their functional MRI results indicated that the lateral prefrontal cortex inhibits extrastriate visual areas. In our experiments, however, it is quite unlikely that the direction of the adapting stimulus was taken as a task-irrelevant signal and that its effect was filtered out. In that case, one would expect neither an effect of adaptation, nor a motion aftereffect, though we provided evidence for each in the main and control experiments.

Earlier research has reported facilitation during language processing in cases in which a visual stimulus matches the content of preceding verbal material. Match advantages of this kind are consistent with two possibilities: Either the visual and linguistic stimuli are compared or integrated at the perceptual level, as befits the embodied-semantics view, or comparison, or integration, occurs at the conceptual level. In our study, we used real- and implied-motion adaptation to determine the level at which the visual and language systems interact. If interaction occurs at the level of perceptual analysis, given that motion was perceived only in the real-motion-adaptation condition, there should be an effect only for real and not for implied motion. Because we found the same qualitative pattern of effects for real and implied motion, we are inclined to conclude that the two systems interact at the conceptual level. Our data suggest that in both cases, the adapting direction was extracted from the visual stimulus and used as a contextual cue for the semantic analysis of the verbal material.

The results of the present experiments do not necessarily provide support for formal accounts of meaning, which are more consistent with modularity or informational encapsulation (Baggio, van Lambalgen, & Hagoort, 2012) than with the interactive view that we favor. Our results suggest instead that, in processing conditions such as those used in our experiments, the role of sensory systems in language processing is to extract contextual cues that can be integrated into a cross-modal representation of discourse (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) rather than to provide a modality-specific substrate for that representation.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Note

1. All *V* statistics were derived from Wilcoxon signed-rank tests, except where indicated.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*, 284–299.

Baggio, G., van Lambalgen, M., & Hagoort, P. (2012). The processing consequences of compositionality. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 657–674). Oxford, England: Oxford University Press.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660.

Bek, J., Blades, M., Siegal, M., & Varley, R. (2010). Language and spatial reorientation: Evidence from severe aphasia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 646–658.

Bex, P. J., Verstraten, F. A., & Mareschal, I. (1996). Temporal and spatial frequency tuning of the flicker motion aftereffect. *Vision Research*, *36*, 2721–2727.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *11*, 527–536.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Dils, A. T., & Boroditsky, L. (2010). Visual motion aftereffect from understanding motion language. *Proceedings of the National Academy of Sciences, USA*, *107*, 16396–16400.

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Georgieva, S., Peeters, R., Kolster, H., Todd, J. T., & Orban, G. A. (2009). The processing of three-dimensional shape from disparity in the human brain. *Journal of Neuroscience*, *29*, 727–742.

Graham, N. V. S. (1989). *Visual pattern analyzers.* Oxford, England: Oxford University Press.

Hiris, E., & Blake, R. (1992). Another perspective on the visual motion aftereffect. *Proceedings of the National Academy of Sciences, USA*, *89*, 9025–9028.

Isaacs, A., Marks, D. F., & Russell, D. G. (1986). An instrument for assessing imagery of movement: The Vividness of Movement Imagery Questionnaire (VMIQ). *Journal of Mental Imagery*, *10*, 23–30.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition*, *94*, B79–B89.

Kemmerer, D., & Tranel, D. (2000). A double dissociation between linguistic and perceptual representations of spatial relationships. *Cognitive Neuropsychology*, *17*, 393–414.

Ledgeway, T. (1994). Adaptation to second-order motion results in a motion after-effect for directionally-ambiguous test stimuli. *Vision Research*, *34*, 2879–2889.

Ledgeway, T., & Hutchinson, C. V. (2009). Visual adaptation reveals asymmetric spatial frequency tuning for motion. *Journal of Vision*, *9*(1), Article 4. Retrieved from http://www.journalofvision.org/content/9/1/4.abstract

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, *49*, 467–477.

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.

Mareschal, I., Ashida, H., Bex, P. J., Nishida, S., & Verstraten, F. A. (1997). Linking lower and higher stages of motion processing? *Vision Research*, *37*, 1755–1759.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*, 17–24.

Mather, G., Pavan, A., Campana, G., & Casco, C. (2008). The motion aftereffect reloaded. *Trends in Cognitive Sciences*, *12*, 481–487.

Mather, G., Verstraten, F., & Anstis, S. (1998). *The motion aftereffect: A modern perspective*. Cambridge, MA: MIT Press.

Papeo, L., Vallesi, A., Isaja, A., & Rumiati, R. I. (2009). Effects of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PLoS ONE*, *4*(2), e4508. Retrieved from http://www.plosone.org/article/info:doi/10.1371/journal.pone.0004508

Pavan, A., Cuturi, L. F., Maniglia, M., Casco, C., & Campana, G. (2011). Implied motion from static photographs influences the perceived position of stationary objects. *Vision Research*, *51*, 187–194.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Proverbio, A. M., Riva, F., & Zani, A. (2009). Observation of static pictures of dynamic actions enhances the activity of movement-related brain areas. *PLoS ONE*, *4*(5), e5389. Retrieved from http://www.plosone.org/article/info:doi/10.1371/journal.pone.00053890005389

Rossini Favretti, R., Tamburini, F., & De Santis, C. (2002). A corpus of written Italian: A defined and a dynamic model. In A. Wilson, P. Rayson, & T. McEnery (Eds.), *A rainbow of corpora: Corpus linguistics and the languages of the world* (pp. 27–38). Munich, Germany: Lincom-Europa.

Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E. T., Brammer, M., & David, A. S. (2000). The functional neuroanatomy of implicit-motion perception or representational momentum. *Current Biology*, *10*, 16–22.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.

Tsushima, Y., Sasaki, Y., & Watanabe, T. (2006). Greater disruption due to failure of inhibitory control on an ambiguous distractor. *Science*, *314*, 1786–1788.

Williams, A. L., & Wright, M. J. (2010). Static representations of speed and their neural correlates in human area MT/V5. *NeuroReport*, *20*, 1466–1470.

Winawer, J., Huk, A. C., & Boroditsky, L. (2008). A motion aftereffect from still photographs depicting motion. *Psychological Science*, *19*, 276–283.

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 224–245). Cambridge, England: Cambridge University Press.

Zwaan, R. A., Madden, C. J., Yaxley, R. H., & Aveyard, M. E. (2004). Moving words: Dynamic representations in language comprehension. *Cognitive Science*, *28*, 611–619.