Psychological Research (2014) 78:55–69
DOI 10.1007/s00426-012-0475-3

ORIGINAL ARTICLE

Beyond words: evidence for automatic language–gesture integration of symbolic gestures but not dynamic landscapes

Dana Vainiger · Ludovica Labruna · Richard B. Ivry · Michal Lavidor

Received: 16 May 2012 / Accepted: 27 December 2012 / Published online: 10 January 2013

Abstract Understanding actions based on either language or action observation is presumed to involve the motor system, reflecting the engagement of an embodied conceptual network. We examined how linguistic and gestural information were integrated in a series of cross-domain priming studies. We varied the task demands across three experiments in which symbolic gestures served as primes for verbal targets. Primes were clips of symbolic gestures taken from a rich set of emblems. Participants responded by making a lexical decision to the target (Experiment 1), naming the target (Experiment 2), or performing a semantic relatedness judgment (Experiment 3). The magnitude of semantic priming was larger in the relatedness judgment and lexical decision tasks compared to the naming task. Priming was also observed in a control task in which the primes were pictures of landscapes with conceptually related verbal targets. However, for these stimuli, the amount of priming was similar across the three tasks. We propose that action observation triggers an automatic, pre-lexical spread of activation, consistent with the idea that language–gesture integration occurs in an obligatory and automatic fashion.

D. Vainiger · M. Lavidor (✉)
Department of Psychology, Bar-Ilan University,
Ramat Gan, Israel
e-mail: michal.lavidor@gmail.com

L. Labruna · R. B. Ivry
Department of Psychology, University of California,
Berkeley, CA, USA

M. Lavidor
Department of Psychology, University of Hull,
Cottingham, UK

Introduction

The transmission of verbal information is frequently accompanied by communicative hand gestures. These gestures can facilitate cross-modal semantic processing, as evidenced by the fact that the productions of speech and gesture are temporally synchronized (e.g., McNeill, 1992; Morrel-Samuels & Krasuss, 1992) and semantically related (Kita, 2000; Krauss & Hadar, 1999; McNeill, 1992). In the present study, we focus on a semantically rich form of gesture, emblems, that have been shown to be communicatively effective (Krauss & Hadar, 1999).

Emblems refer to gestures that have become established as a vocabulary item with a standardized form and a shared meaning within a culture (e.g., the 'Ok' and 'V' signs). Emblems frequently occur in the absence of speech, often in circumstances that make the verbal transference difficult (e.g., when gesturing from a distance), but they are also common during speech, carrying meanings that are congruent with the content of the speech (Kendon, 2004). Understanding the mechanisms underlying emblem comprehension is also important given their integral role in the development of symbolic communication during infancy (Bates & Dick, 2002; Goodwyn, Acredolo, & Brown, 2000). Unlike actions, emblems are codified signs, denoting their meaning in a symbolic manner, similar to words. It may be that similar processes of conventionalization underline the formation of both vocabularies. As such, understanding how emblems are processed may shed light on the study of language acquisition (Corballis, 1998; Rizzolati & Arbib, 1998). While there has been extensive research on the functional role of gesticulations (Kita, 2000), only a few studies have employed emblems, and this work has mainly addressed descriptive aspects of emblems (Brookes, 2005; Kendon, 2004). We are familiar with only three studies that have explored the interaction between emblems and words in comprehension (Bernardis & Gentilluci, 2006; Gentilucci, Bernardis, Crisi, & Volta, 2006; Gunter & Bach, 2004).

Different models have been proposed to account for the integration of gestural and verbal information. The "independent systems" hypothesis proposes that, while speech and gesture stem from the same source concept in working memory, they diverge at an early stage into separate pathways. Support for this idea comes from findings that gestures are initiated before the corresponding speech (e.g., Morrel-Samuels & Krasuss, 1992; Schegloff, 1984) and that gestural production is related to lexical retrieval failures (e.g., Frick-Horbury & Guttentag, 1998; Rauscher Krauss & Chen, 1996). The conceptualization of a communicative intent is done exclusively from propositional (symbolic) representations that feed into verbal pathways. However, if gestures correspond to the communicative intent, they can supply cues for the lexical selection and ease word retrieval (i.e., cross-modal priming). Accordingly, gestures primarily benefit the speaker and may have little effect on the perceiver who relies on the interpretation of meaning from the verbal signals (Krauss et al., 2000; Krauss & Hadar, 1999; Rimé & Schiaratura, 1991). In this view, semantic integration of gesture and speech is regarded as a post-lexical process, taking place only after semantic processing of the verbal message has occurred (Kelly, Creigh, & Bartolotti, 2010a). Note that this model is about the producer, with the underlying assumption that the perceiver is minimally aided by the gestural component.

The "integrated systems" hypothesis stands in contrast to the independent systems hypothesis. By this model, gesture is seen as an integral part of meaning construction, such that thinking combines two modes of operation, linguistic and imagistic (Kita, 2000; Kita & Özyürek, 2003; McNeill, 1992). In comprehension, gestures are thought to interact with speech in an obligatory manner, helping disambiguate the verbal message to enhance communication (Bernardis & Gentilucci, 2006; Holle & Gunter, 2007; Kelly, Özyürek, & Maris, 2010b). The communicative effectiveness of gesture is supported by many classic manipulations, showing behavioral differences between conditions in which listening to speech is accompanied by congruent gestures compared to when listening to speech alone, or to speech accompanied by incongruent gestural information (Kendon, 1994).

The integration hypothesis is one variant of a general class of embodiment theories that are centered on the idea that language comprehension involves the engagement of neural systems for perception and action (e.g., Lakoff, 1987; Zwaan, 2004). Motor and premotor brain activities are observed during the processing of action-related language (Hauk, Johnsrude, & Pulvermüller, 2004; Oliveri et al. 2004; Pulvermüller, Shtyrov, & Ilmoniemi, 2005; Tettamanti et al., 2005), consistent with the idea that such concepts are linked to our bodily experience of the world (Lakoff & Johnson, 1999). Indeed, these observations are central to the idea that a human mirror neuron system (see Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010, for direct evidence of such system) helps mediate semantic understanding of actions by mapping input representations from observed actions onto already encoded motor schemes (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fogassi, & Gallese, 2001).

The presumed involvement of the motor system in both language and action semantic processing, suggests that the comprehension of words and gestures should share common conceptual representations. This hypothesis is supported by ERP evidence showing that the N400, a marker of semantic processing, is similarly distributed for gestures and words (Wu & Coulson, 2007). Similarly, comprehension of emblems and spoken words has been found to activate a largely overlapping network involving inferior and posterior temporal regions along the Sylvian fissure. This network is regarded as the core of the language system and may be seen as an amodal network supporting symbolic communication (Xu, Gannon, Emmorey, Smith, & Braun, 2009).

Semantic priming (SP) has served as an important behavioral tool in studies exploring links between linguistic and non-linguistic systems. Overall, those studies have shown priming effects between gestures and language, with behavioral and electrophysiological measures showing a differential response when the word and gesture are related, compared to when they are unrelated (e.g., Bernardis & Caramelli, 2009; Bernardis, Salillas, & Caramelli, 2008; Wu & Coulson, 2007; Yap, So, Melvin Yap, Tan, & Teoh, 2011). However, the relationship of gesture–language interactions to other forms of picture–word priming remains unclear, as does the interpretation of these interactions. In some studies, priming from gestures to words was interpreted to reflect facilitated processing of related meanings (Holle & Gunter, 2007; Kelly et al., 2010b; Yap et al., 2011). Alternatively, it has been claimed that the priming is actually the result of a lack of interference that arises when the prime and target are related (Bernardis et al., 2008). In comparison to a baseline, no-prime condition, naming latencies for words primed by unrelated pantomimes were slower. No facilitation was observed in a congruent condition in which the gestures and words were related. Interpreting priming as positive (facilitation) or negative (interference) is important for evaluating how gesture influences linguistic communication (Bernardis et al., 2008). Mechanisms underlying semantic priming are frequently divided into those associated with automatic and those associated with strategic processes. The effects of automatic processes tend to become manifest quickly, occurring without intention, need not be conscious, and generally produce facilitative, but not interfering priming. The effects of strategic processes take longer time to become manifest, are intentional, conscious, and can produce both facilitative and interfering priming (McNamara, 2005).

Our goal in the present study was to explore gesture–word priming in a systematic manner, with the expectation that this approach would shed light on the underlying mechanisms. We employed three different tasks to manipulate the influence of strategic processing on semantic priming. In each task, a target word was preceded by a gesture prime. Participants responded by making a lexical decision to the target (Experiment 1), naming the target (Experiment 2), or performing a semantic relatedness judgment of the prime–target pair (Experiment 3). If we assume that well-learned gestures provide an independent form for accessing meaning, then we would expect that the perception of these stimuli might trigger an automatic and pre-lexical spread of activation within a semantic network (Collins & Loftus, 1975). We would expect this to occur in all three experiments given that the prime and target share a related meaning. The activation should spread passively to related nodes in semantic and lexical networks, providing a mechanism that can facilitate the retrieval of an associated lexical unit.

Both lexical decision and naming tasks emphasize the analysis of the lexical form of the target (Andrews & Heathcote, 2001) and therefore enable an implicit examination of language–gesture integration. In contrast, semantic relatedness judgments emphasize controlled aspects of semantic processing (Neely, 1991) by virtue of requiring the explicit processing of the meaning of both the prime and target. Hence, we assume that naming is the most effortless task and should be driven by fast and automatic priming mechanisms. Lexical decision was assumed to sustain an intermediated level of controlled processing, as it involves some decisional processes (Duscherer & Holender, 2005; Neely & Keefe, 1989). Semantic relatedness judgments were considered the most demanding given that this task would entail prolonged and strategic semantic integration. Such judgments are assumed to require the explicit access of meanings (Kang, Blake & Woodman, 2011; Lovseth & Atchley, 2010; Wu & Thierry, 2010).

By comparing priming across the three tasks, we sought to determine the degree to which gesture–word integration could be attributed to fast, automatic and obligatory mechanisms of priming (Kelly et al. 2010a). We opted to include a relatively small proportion of related trials (40 % at most) and a short SOA, factors designated to emphasize automatic processing and minimize the contribution of strategic expectations (Neely & Keefe, 1989). A short SOA is also important for minimizing the opportunity to verbalize the gestures, since such verbalization might influence the results (Yap et al., 2011).

To allow better interpretation of the priming by symbolic gestures (SGs), we used three types of controls. First, in Experiment 1, we estimated facilitation and interference with respect to a baseline condition in which the primes were meaningless gestures. Second, in all three experiments, we included non-action conditions by creating short videos of landscape scenes. These allowed us to compare priming effects between two natural categories, actions and landscapes and to assess the specificity of the patterns of priming. Third, in Experiments 2 and 3, we included a condition with an identical meaningful priming context and unrelated targets from a distant semantic category (topic-unrelatedness). This provided an alternative way for estimating the degree of facilitation and interference from the primes.

In summary, across our three experiments, we tested the following predictions:

1. Facilitation for relatedness will always be obtained, as it is expected under both automatic and controlled mechanisms. However, interference will only be observed when controlled processing is required.
2. Semantic priming will be stronger between gesture and language compared to picture–word priming, given that the former entails embodied representations.
3. The magnitude of semantic priming will be positively related to the degree of controlled processing required by the task.

Experiment 1: Lexical decision

Method

Participants

Twenty students from Bar-Ilan University (11 females), aged 21–32 years ($M = 23.2$, SD 2.5), took part in the lexical decision task. For this and subsequent experiments, participants were right handed, as evaluated with the Edinburgh Handedness Inventory (Oldfield, 1971), and had normal or corrected-to-normal vision. All were native speakers of Hebrew, had lived in Israel since infancy, and had no history of dyslexia or attention deficit diagnosis. The participants gave written informed consent.

Stimuli

Experimental stimuli were selected after an extensive pretesting phase. First, candidate symbolic gestures (SGs) and meaningless gestures (MGs) were videotaped and edited into 1,520 ms video clips. The first 240 ms presented the actress in a static posture with her palms on the table. During the remaining 1,280 ms, the actor produced a gesture. The actress wore a white mask to eliminate cues from facial expressions and gaze direction. Candidate landscape clips

(LSs) were taken from nature movies depicting a salient natural scene. These were edited into 1,520 ms clips. Examples of the video clips are presented in Fig. 1.

Sixteen healthy volunteers (9 women), aged 17–45 years ($M = 31.6$, SD 7.4), native Hebrew speakers who had lived in Israel since infancy, were recruited to judge the candidate materials. For the SG, the judges were told that some of the video clips depicted movements that conveyed a conventional meaning within the culture, while other clips depicted movements lacking conventional meaning. In each trial, they were asked to rate the clip on a 1–5 Likert scale for conventionality (1, totally meaningless; 5, highly conventional meaning). For items given a rating of 3 or greater, they were then asked to identify the gesture with 1–3 words, providing the best verbal description (for items rated 4 or 5) or a best guess (for items rated 3). The instructions stressed that the participant should respond with the first meaning that comes to mind. For the LS, the judges were told that the videos were of inanimate objects and they were asked to verbally describe the main object or scene with 1–3 words. If they did not recognize the object or scene, they should indicate this.

Gestures were analyzed for their conventionality and meaningless agreement scores, calculated as the percentage of judges who rated items as 4–5 or 1–2 on the conventionality scale, respectively. From the full set, 115 of the gestures were rated as symbolic and 73 gestures were rated as meaningless, based on an agreement score of at least 75 % of judges. Using the same criterion, 80 LSs were approved as having a clear and accepted meaning. The selected meaning was the one with the highest semantic agreement score, calculated as the percentage of judges who offered names corresponding to that meaning. A congruent name was selected out of those different names. This was chosen as the one with the highest lexical agreement score, calculated as the percentage of judges who offered this exact name or its variations (i.e., singular and plural forms, declensions of the same root).

Congruent names were characterized according to their linguistic category (e.g., noun, verb, adjective, and interjection), length, frequency (Log_{10} of the number of search results in the Google website; Blair, Urland, & Ma, 2002), and concreteness. For the latter, a group of six raters were asked to judge the items as either concrete or abstract. Semantically unrelated names were matched to each of the congruent names, preserving their linguistic categories, length, concreteness, as well as overarching topic. Unrelated names assigned to the SGs had a different social content or communicative intent (e.g., "don't know" gesture presented with "You exaggerate" text), while unrelated names to the LSs were of another natural object or scene (e.g., a cloudy storm clip presented with a: "Laguna" text).

For the priming task, two stimuli lists were generated, with 236 prime–target pairs in each. Primes were repeated in different conditions, but not within subjects. Each list contained 28 % congruent pairs (42 SG and 24 LS trials), 32 % unrelated pairs (28 SG, 24 MG and 32 LS trials), and 40 % in which the targets were pseudo-words (46 SG, 16

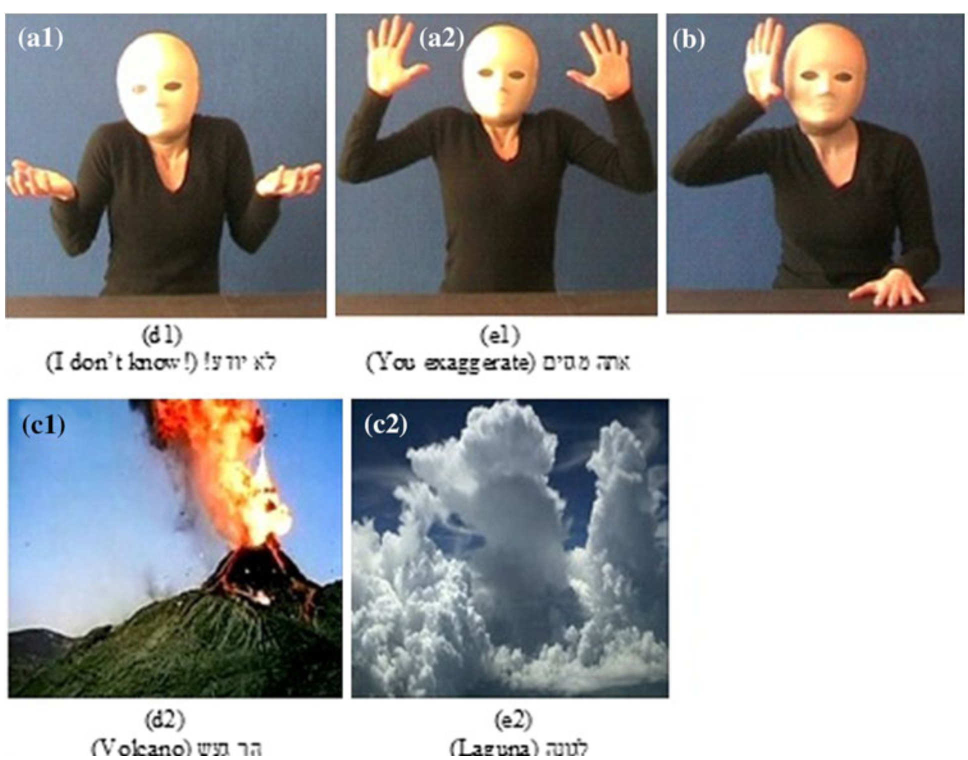


Fig. 1 Examples of the stimuli: symbolic gesture (**a1**) presented with its congruent name (*d1*); symbolic gesture (**a2**) presented with its unrelated name (*e1*); meaningless gesture (**b**); landscape (**c1**) presented with its congruent name (*d2*); landscape (**c2**) presented with its unrelated name (*e2*)

MG and 32 LS trials). Pseudo-words were created by switching positions of 2–3 letters of items in a word list that did not share meanings with any of the word targets, but were matched in terms of lexical characteristics and overarching topic. The selection and assignment of items were controlled through the continuous calculation of descriptive statistics to maximize the match across different lists and conditions for background variables (i.e., conventionality, semantic and lexical agreement scores, names' length, frequency, concreteness and lexical categories).

Design

A 3 × 2 × 2 within-subjects design was employed, with prime type (symbolic gestures, SGs; meaningless gestures, MGs; landscapes, LSs), semantic relation (congruent, unrelated) and target lexicality (words, pseudo-words) as factors. Note that congruency is nonsensical for the meaningless gestures, based on our arbitrary assignment of prime–target pairs to congruent and unrelated categories. The meaningless gestures were only employed in the first analysis. RT and accuracy were recorded as dependent variables. Prime–target pairs included SG or LS priming over congruent or unrelated names, and MG priming over unrelated names.

Procedure and apparatus

Subjects were seated with their eyes approximately 60 cm from a computer monitor (75 Hz refresh rate). For each subject, one of the two lists was used. The stimuli were tested in two blocks, with one block involving gesture primes (SG + MG) and the other landscape primes. The order of specific trials was randomized. Presentation was controlled by the software package, E-prime 2.0.

Subjects were informed that a short video clip of an actress or a landscape will be shown, followed by the brief presentation of a string of letters. Their task was to make a lexical decision, pressing one of two keys with the right index finger to indicate if the letter string formed a word or a pseudo-word. Emphasis was given that the responses should be made as quickly and accurately as possible. There were 20 practice trials. The response buttons, blocks order and lists were counterbalanced.

Each trial began with the presentation of a fixation marker for 1,000 ms, followed by the 1,520 ms video clip (the prime). The fixation marker was then re-presented for 30 ms and replaced by the string of letters (the target). The target remained on the screen for only 180 ms. The screen was then blanked for 2,000 ms. Responses were recorded only if they followed target onset. All stimuli were displayed on a black background. The video clips occupied a 197 × 152 mm central rectangle and the words were presented in the center in a 22-point courier new, bold, white font.

Results and discussion

Participants were very accurate on the LDT, with accuracy above 94 % across all conditions. The primary analyses focus was on the reaction times (RT) of correct trials. We excluded RTs were outside *a* ± 2.5 SD range around each participant's mean in each condition. This criterion excluded 5.9 % of the correct trials. Following this filtering, RT distribution was normal (with kurtosis value of 0.79).

Mean RTs and SDs in each condition and experiment are provided in Table 1. The initial analysis examined the priming effects for gestures. There was a significant effect in both analyses based on participants and items [$F_1(2,38) = 17.80$, $p < 0.001$, $F_2(2,185) = 28.14$, $p < 0.001$, respectively]. For this and subsequent analyses and unless otherwise stated, post hoc tests were conducted on the participants' means using the Bonferroni adjustment method. This analysis revealed that RTs were faster when the prime and target were congruent compared to when the primes were meaningless gestures, our baseline condition ($p < 0.001$). Thus, we observed a facilitative priming effect. The difference in RT between the unrelated symbolic gesture (SG) and baseline meaningless gesture (MG) conditions did not differ ($p > 0.5$), suggesting there was no interference when the prime and target were not congruent in meaning.

We next compared the priming effects for gestures and landscapes videos. For this analysis, we did not include the MG primes since these do not exist for the LS condition. There was a main effect of semantic relation [$F_1(1,19) = 35.44$, $p < 0.001$; $F_2(1,232) = 49.24$, $p < 0.001$], resulting from faster RTs when the prime–target were congruent ($M = 714$ ms, SEM 25) compared to when they were unrelated ($M = 776$ ms, SEM 27). There was also a significant prime type × semantic relation interaction [$F_1(1,19) = 5.34$, $p < 0.05$; $F_2(1,232) = 5.1$, $p < 0.05$; see Fig. 2]. RTs on congruent trials were faster for the SG condition compared to the landscape (LS) condition. The two prime types did not differ on unrelated trials ($F < 1$).

RTs on trials in which the target was a pseudo-word did not differ across conditions (all *p*s > 0.3).

To summarize, in a lexical decision task with a relatively short SOA and a small percentage of congruent trials, the SG primes facilitated the processing of congruent lexical targets. This finding is consistent with the first two predictions presented in "Introduction". First, the priming effects were asymmetric. When the prime and target were related, RTs were faster compared to baseline (facilitation). In contrast, when the prime and target were unrelated, RTs were not slower relative to baseline (no interference). Second, there was a significant interaction of relatedness and prime type reflecting the fact that SP was stronger for the gestures compared to the landscapes.

Table 1 Mean RT for correct responses (±SD) and percent correct for each experiment (Exp. 1: lexical decision; Exp. 2: naming, Exp. 3: semantic relatedness judgment) and according to the prime–target pairing

Prime type	Target	Exp. 1 ($n = 20$)	Exp. 2 ($n = 20$)	Exp. 3 ($n = 23$)
RT				
MG	Unrelated	768 (136)		
SG	Congruent	699 (103)	680 (114)	764 (147)
	Unrelated-TR	781 (127)	708 (111)	882 (135)
	Unrelated-TU		719 (115)	782 (114)
LS	Congruent	729 (129)	704 (111)	807 (156)
	Unrelated-TR	770 (122)	729 (116)	844 (125)
	Unrelated-TU		716 (119)	814 (141)
MG	Pseudo-words	835 (136)		
SG	Pseudo-words	830 (135)		
LS	Pseudo-words	850 (138)		
Accuracy				
MG	Unrelated	96.4 (0.03)		
SG	Congruent	97.7 (0.03)	96.6 (0.06)	97.0 (0.04)
	Unrelated-TR	95.3 (0.06)	95.9 (0.05)	93.1 (0.08)
	Unrelated-TU		96.4 (0.05)	98.1 (0.03)
LS	Congruent	97.4 (0.03)	97.9 (0.03)	94.5 (0.06)
	Unrelated-TR	96.8 (0.05)	96.1 (0.06)	97.5 (0.04)
	Unrelated-TU		95.0 (0.05)	97.3 (0.05)
MG	Pseudo-words	94.3 (0.04)		
SG	Pseudo-words	98.5 (0.03)		
LS	Pseudo-words	94.5 (0.11)		

MG meaningless gesture, *SG* symbolic gesture, *LS* landscape scenes, *TR* topic related, *TU* topic-unrelatedness

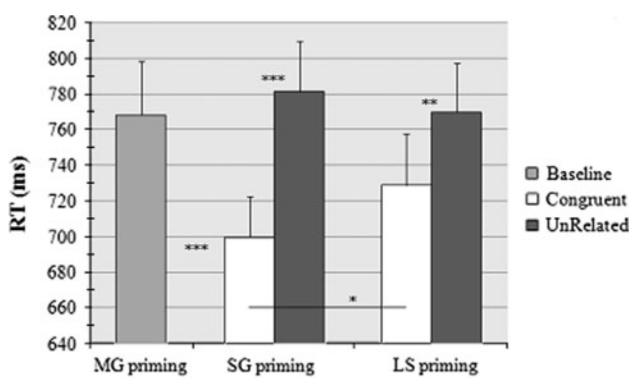


Fig. 2 Experiment 1: mean RTs in ms (+SEM) of correct lexical decisions, according to the prime type (*MGs* meaningless gestures, *SGs* symbolic gestures, *LSs* landscape scenes) and semantic relation (congruent, unrelated) (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$)

Experiment 2: Naming

Method

Participants

Twenty healthy subjects took part in the naming experiment (11 females), aged 20–35 years ($M = 25.45$, SD 4.2).

Stimuli

84 SGs and 72 LSs, with their paired word targets from Experiment 1, were used in this study. The stimuli were divided into three lists, with each list composed of 130 prime–target pairs. Primes repeated across the different semantic levels, but did not appear in the same list. In addition to the congruent and unrelated items, we formed a new category of unrelated prime–target pairs. As in Experiment 1, the unrelated items were composed on the basis of topic-relatedness (TR) to the congruent targets while the new unrelated condition involved the pairing of the meaningful primes with unrelated targets from a distant semantic category (topic-unrelatedness; TU, e.g., a clip of a farewell gesture and the word “fire”). This provides an alternative way to evaluate potential facilitation and interference from the primes. Targets in the third semantic level (unrelated-TU) were the congruent targets of the other prime type in the other lists. 40 % of the trials included prime–target pairs that were congruent (28 SG and 24 LS trials). The other 60 % included pairs that were unrelated (42 SG and 36 LS trials), divided equally between the TR and TU conditions. The different conditions were matched for relevant background variables, as in the previous design.

Design

A 2 × 3 within-subjects design was employed, with prime type (SGs, symbolic gestures; LS, landscape scenes) and semantic relation (congruent, unrelated-TR, unrelated-TU) as factors, and RTs and accuracy as the depended variables.

Procedure and apparatus

Each subject was tested with one of the three stimulus lists, with the selected list divided into separate blocks for the SG and LS primes. On each trial, a short video clip of an actress or landscape was presented, followed by a brief presentation of 1–3 words. Participants were instructed to read the word(s) out loud as quickly and accurately as possible. A microphone, connected to the computer, was used to record the responses. The test blocks were preceded by 10 practice trials. The blocks order and stimuli lists were counterbalanced.

The trial time line was similar to Experiment 1, except that the re-presentation of the fixation marker between the prime and target was now extended to 100 ms. Vocal responses were analyzed manually for naming latency (i.e., the time interval between target onset and naming onset) and accuracy, using Praat software (http://www.fon.hum.uva.nl/praat/).

Results and discussion

Correct naming responses with RTs within ±3 SDs around each participant's mean in each condition (95.73 % of the data, ranging between 389 and 1,438 ms) were included in the analyses. Following this filtering, RT distribution was normal (with kurtosis value of 0.64). Accuracy was above 95 % in all conditions.

The effect of prime type was significant [$F_1(1,19) = 5.53$, $p < 0.05$; $F_2(1,384) = 4.52$, $p < 0.05$], with naming RTs faster in the SG condition ($M = 702$, SEM 25) compared to the LS condition ($M = 716$, SEM 26). There was also a main effect of semantic relation [$F_1(2,38) = 14.29$, $p < 0.001$; $F_2(2,384) = 8.62$, $p < 0.001$], with faster RTs on congruent trials ($M = 692$, SEM 25) compared to the unrelated conditions, TR ($M = 718$, SEM 25; $p < 0.001$) and TU ($M = 717$, SEM 26; $p < 0.01$).

The prime type × semantic relation interaction was significant in the participants' analysis [$F_1(2,38) = 5.14$, $p = 0.01$; $F_2(2,384) = 2.26$, $p = 0.10$, see Fig. 3]. For SG priming [$F(2,18) = 10.24$, $p = 0.001$], RTs to congruent targets were faster than RTs in both unrelated conditions, TR ($p < 0.01$) and TU-baseline ($p = 0.001$). The two unrelated conditions did not differ from one another ($p > 0.40$). For LS priming [$F(2,18) = 14.27$, $p < 0.001$], RTs to congruent targets were faster than unrelated-TR targets ($p < 0.001$), but did not differ when compared to the TU-baseline condition ($p > 0.2$).

In a second post hoc analysis, the prime-type effect was found to be significant only for the congruent [$F(1,19) = 9.48$, $p < 0.01$] and unrelated-TR conditions [$F(1,19) = 8.89$, $p < 0.01$], with RTs faster in the SG priming compared to the LS priming. There was no difference on the baseline trials (TU and TR, $F < 1$).

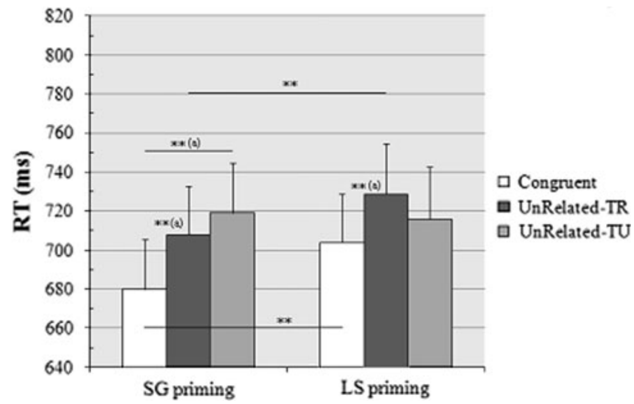


Fig. 3 Experiment 2: mean RTs in ms (+SEM) of correct naming responses, according to the prime type (*SGs* symbolic gestures, *LSs* landscape scenes) and semantic relation [congruent, unrelated-TR (topic related), unrelated-TU (topic unrelated)] [**$p < 0.01$, (*a*) Bonferroni adjusted]

To summarize, the results for the naming task replicated, for the most part, those obtained in the lexical decision task of Experiment 1. Semantic priming was again obtained for both the symbolic gesture and landscape primes. With symbolic gestures, RTs were faster when the prime and targets were congruent compared to the baseline condition. Interference was again not found when the primes were incongruent with the target, an effect observed for both the symbolic and landscape primes. While we defer a formal comparison across experiments until after Experiment 3 (see below), the priming effects here are generally smaller than in Experiment 1. Whereas the lexical decision task of that experiment showed a larger priming effect for SG compared to LS (82 vs. 41 ms), the magnitude of the priming effects were similar for SG and LS in the current experiment (28 and 25 ms, respectively). Reduced priming has previously been reported with naming compared to LD (Duscherer & Holender, 2005; Perea & Rosa, 2002), a result interpreted to indicate that naming is more dependent on the fast and automatic activation of lexical nodes, whereas lexical decision may entail some degree of strategic processing.

Experiment 3: Semantic relatedness judgment

Method

Twenty-three healthy participants took part in the relatedness judgment experiment (14 females), aged 20–36 years ($M = 26.13$, SD 4.73). The procedure was identical to Experiment 2, except that participants were now required to explicitly judge whether or not the target word was related to the prime. This binary decision was made by pressing one of two keys with the right index finger.

Results and discussion

Trials in which the RTs were within ±3 SDs around each participant's mean in each condition (95.21 % of the data, ranging between 374 and 2,168 ms) were included in the analyses. Following this filtering, RT distribution was normal (with kurtosis value of 0.70). Accuracy was above 93 % in all conditions.

A 2 (prime type) × 3 (semantic relation) two-way ANOVA revealed a main effect of semantic relation [$F_1(2,44) = 19.04$, $p < 0.001$; $F_2(2,384) = 21.42$, $p < 0.001$] and a significant prime type × semantic relation interaction [$F_1(2,44) = 9.06, p < 0.001$; $F_2(2,384) = 5.62$, $p < 0.01$, see Fig. 4]. Across the two prime types, RTs were faster on congruent trials ($M = 786$, SEM 32) and TU-baseline trials ($M = 798$, SEM 25) compared to the unrelated-TR trials ($M = 863$, SEM 28), both *p*s < 0.001. However, the interaction reflected the fact that, while RTs for congruent trials were faster than the unrelated-TR baseline condition, the congruency effect was marginally larger under SG priming [$F(1,22) = 4.27$, $p = 0.051$].

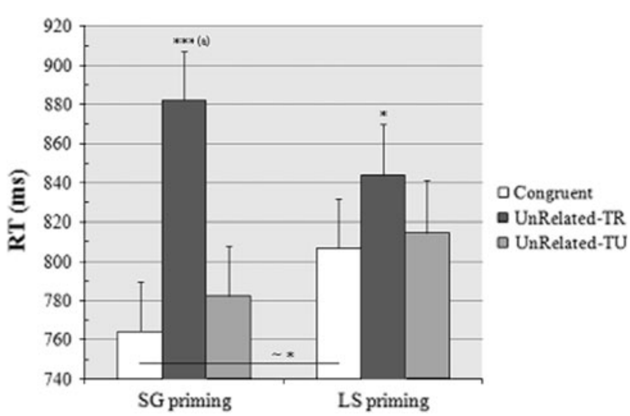


Fig. 4 Experiment 3: mean RTs in ms (+SEM) of correct semantic relatedness judgments, according to the prime type (*SGs* symbolic gestures, *LSs* landscape scenes) and semantic relation [congruent, unrelated-TR (topic related), unrelated-TU (topic unrelated)] [*$p = 0.051$, *$p < 0.05$, ***$p < 0.01$, (*a*) Bonferroni adjusted]

Comparisons across experiments

Table 2 presents the difference between congruent and unrelated targets in the three experiments. Statistical comparisons were performed using independent *t* tests.

To evaluate our first prediction, we considered facilitation and interference effects for the three types of tasks. The priming magnitude and direction were evaluated with respect to a baseline condition. Facilitative priming for congruent targets was expected in all tasks, given that this can arise from both automatic and controlled mechanisms. Congruent primes led to faster RTs in the lexical decision task of Experiment 1 and the naming task of Experiment 2. Unexpectedly, facilitation was not reliable in the semantic relatedness judgment task of Experiment 3, the task in which we would expect a stronger contribution from strategic processes. However, unlike the lexical decision task in which congruent and baseline stimuli led to the same response (e.g., both are words), semantic judgments result in the response "related" on congruent trials and "unrelated" on baseline trials. As such, the tasks likely involve different decision-related processes, making comparisons between the two conditions problematic. We do note that facilitation was clearly evident in the semantic relatedness task for the gesture primes in comparison to the landscapes priming.

Interference was expected only when task demands require controlled processing. This prediction was confirmed: RTs on unrelated trials were significantly slower than baseline trials in the semantic relatedness task of Experiment 3. This effect was especially pronounced for the gesture stimuli (100 ms) compared to the landscape stimuli (30 ms). We did not observe interference for the lexical decision and naming tasks.

The second prediction, derived from embodied cognition theory (e.g., Lakoff, 1987; Zwaan, 2004), was that semantic priming would be larger for concepts that are related to a bodily schema (e.g., gestures) compared to concepts that lack such body-oriented representations (e.g., landscapes). We recognize that comparisons across stimulus categories are problematic; nonetheless, the results of all three experiments were consistent with this prediction.

The third prediction addressed the magnitude of priming within a stimulus category, with the expectation that priming would be larger for tasks that were relatively demanding on controlled processes. The results for the semantic gestures were generally consistent with this prediction, given the assumption that the lexical decision and semantic relatedness tasks entail greater control than naming. In comparison to the naming task, priming was larger for the lexical decision [$t(38) = 3.4$, $p < 0.01$] and semantic relatedness [$t(41) = 4.05$, $p < 0.001$] tasks. Interestingly, the magnitude of landscape priming did not change significantly across tasks (all *p*s > 0.2).

We note that there were considerable differences in RTs between the three experiments and this might influence the magnitude of priming. To address this concern, we performed a between-experiment comparison using a normalized measure of priming (percent change with respect to baseline), focusing on whether there were differences

Table 2 Between experiment comparison

Exp.	Task	RT SG congruent	RT SG unrelated	RT LS congruent	RT LS unrelated	SG prime	LS prime	% SG prime	% LS prime
1	Lexical decision	700	782	727	768	82 (14)	41 (13)	11.7 (0.8)	5.6 (0.6)
2	Naming	680	708	702	727	28 (7)	25 (5)	4.1 (0.5)	3.5 (0.5)
3	Semantic judgment	764	882	805	842	118 (20)	37 (16)	15.4 (0.8)	4.5 (0.60)

Mean RTs, absolute differences, and proportional changes (±SE) between congruent and unrelated targets

SG symbolic gesture, *LS* landscape scenes

between the symbolic and landscape conditions. The effect of stimulus type was reliable [$F(1,60) = 80.7$, $p < 0.00001$], with more facilitation observed for the symbolic gesture primes (mean 10.4 %) compared to the landscape primes (4.5 %) (see Table 2).

General discussion

In this study, we set out to carefully examine the relationship between symbolic gestures and verbal language. We considered different ways in which these two modes of communication could interact. The "independent systems" hypothesis proposes that, while speech and gesture stem from the same source concept in working memory, they diverge at an early stage into separate pathways (Morrel-Samuels & Krasuss, 1992; Schegloff, 1984), with communicative intent done exclusively from propositional (symbolic) representations that feed into verbal pathways. In this view, semantic integration of gesture and speech is regarded as a post-lexical process, taking place only after semantic processing of the verbal message has occurred (Kelly et al. 2010a). Note that this model is about the communicator, with the underlying assumption that the perceiver is minimally aided by the gestural component.

An alternative hypothesis is the "integrated systems" where a gesture is seen as an integral part of meaning construction, such that thinking combines two modes of operation, linguistic and imagistic (Kita, 2000; Kita & Özyürek, 2003; McNeill, 1992). In comprehension, gestures are thought to interact with speech in an obligatory manner, helping disambiguate the verbal massage to enhance communication (Bernardis & Gentilucci, 2006; Holle & Gunter, 2007; Kelly et al. 2010b).

Considering the contrasting theories, the present research aimed to investigate the interaction of non-verbal and verbal language comprehension under differing levels of controlled processing. We used three tasks—naming, lexical decision, and semantic relatedness judgments—that differ in terms of the demands they place on controlled processing. We used a relatively small proportion of trials in which the prime and target were related, as well as a short SOA because these conditions have been shown to emphasize automatic processing while minimizing the contribution of strategic expectations (Neely & Keefe, 1989).

Our first goal was to assess whether gesture–word priming would be obtained under all three conditions. Consistent with the prior work (e.g., Holle & Gunter, 2007; Kelly et al., 2010b), RTs in all three tasks were facilitated when the primes were congruent with the target.

We also predicted that semantic priming of words would be stronger with gestural primes compared to landscape primes given that the former entail embodied representations (see Yap et al., 2011; Wu & Coulson, 2007). The magnitude of priming in the landscape conditions was less than that observed in the gesture conditions, especially for the lexical decision and semantic relatedness tasks. This suggests that the benefit of embodied representations may be modulated by the degree of controlled processing.

The modulatory effect of controlled processing was evident in the comparison between the three experimental tasks. We assume that naming is the most effortless task, placing little demands on post-lexical semantic analysis. The fact that we observed some degree of semantic priming would indicate that there is some obligatory linkage between gestures and their linguistic referents (Kelly et al. 2010a). However, we observed larger facilitation in the lexical decision task compared to naming. This has been attributed to the additional contribution of post-lexical semantic matching (Duscherer & Holender, 2005; Neely & Keefe, 1989).

Negative priming (interference) was expected only when task demands require controlled processing. This prediction was confirmed: RTs on unrelated trials were significantly slower than baseline trials in the semantic relatedness task. This effect was especially pronounced for the gesture stimuli compared to the landscape stimuli, and there was no interference for the lexical decision and naming tasks.

The lack of interference in our naming results was in contrast to the results reported by Bernardis et al. (2008). In Bernardis et al. (2008), negative priming was observed with gestures. However, the baseline in that study was a condition in which the prime was excluded. In contrast, we estimated facilitation and interference from the primes by including conditions in which the primes were meaningless gestures (in the lexical decision) or from a semantically unrelated category (in the naming and semantic judgment tasks). Therefore, the lexical processing in the experimental and baseline conditions always included priming, unlike the experimental design in Bernardis et al.'s (2008) study.

For naming, the observation of communicative hand gestures, whether congruent or not, might be expected to increase the communicational attunement of the participants. This hypothesis is motivated by the idea that the manual and oral control systems might provide separate means to action (Bernardis & Gentilucci, 2006). Therefore, naming of words that carry a social content was overall facilitated under gestures priming, in both congruent and unrelated (topic related) conditions, compared to the control priming by landscapes. Hence, the naming and lexical decision demonstrated that implicit processing of language–gesture relationship does not lead to a cost (i.e., interference) following a semantic mismatch. This can be considered an advantage, as the verbal processing continues without interference. On the other hand, this can be taken as a disadvantage, since the recipient might miss a chance to learn valuable information.

It is possible that the priming effects observed with our visual clips may be verbally mediated given that the clips were presented for over a second. Indeed, it would seem unlikely that verbal processes would not have been engaged as the participants process the primes. Nonetheless, the differences between the gesture and landscape primes suggest that gesture–language associations provide a unique link compared to other semantic categories. While we recognize that there may be other factors that differ between the gestures and landscapes, the targets for the two domains were selected to be similar on a number of categories (e.g., lexical frequency). Thus, there is no a priori reason to expect that the priming effects should differ for gestures and landscapes, especially on the lexical decision task. The fact that we did observe greater facilitation for the gestures in all of the tasks points to a gesture–language association, one consistent with the ideas of cognitive embodiment.

What implications may our results offer to a more complete understanding of gesture–language integration? Kelly et al. (2010b) proposed that gesture and speech interact to enhance language comprehension in a reciprocal and obligatory manner. Our findings are consistent with this hypothesis as we observed priming effects even in conditions that emphasized fast and automatic processing. This suggests that integration may occur at a pre-lexical stage of processing, with gestures conveying meaning independent of language (Kendon, 1994; Kita, 2000; McNeill, 1992). We suggest that gestures may influence our linguistic representation of actions in two distinct ways. First, there may be a passive form of interaction, one that is evident during automatic processing (Obermier, Holle, & Gunter, 2011). Indeed, the form of interaction here may reflect the operation of general mechanisms for cross-modal priming one that entails the automatic spread of semantic activation. The larger priming effects for gestures compared to landscapes in the more automatized tasks (i.e., lexical decision and naming) may reflect stronger associations between gesture primes and targets. The stronger gestures' associations (compared to landscapes) are predicted by the embodiment theory, due to the large overlap between action semantics and language semantics, shaped by the experience of the individual.

The second form of interaction between gesture and language involves a more "active" form of processing, one that is evident under controlled and attentive processing, and operates at a post-lexical level. Interaction at this level would promote both facilitation and interference effects. This hypothesis is consistent with recent evidence showing that when the synchrony between gesture and speech is disrupted, effortful processing is required to combine the contents across these two domains (Obermeier et al., 2011). These interactions would enable the interplay of analytic and motoric thinking, an idea promoted by Kita (2000) in the study of production. Here we reach a similar conclusion with respect to language comprehension, one that can be supported through either speech or, as shown in the present study, with emblematic actions.

Conclusions

The current experiments were designed to address and contrast two models of how symbolic gestures and verbal language are integrated, one that assumes that speech and gesture diverge at an early stage (independent systems), and one where gesture is seen as an integral part of meaning construction (integrated systems). We employed three tasks—naming, lexical decision, and semantic relatedness judgments—that require different levels of controlled processing. We observed greater facilitation for gesture–language pairs compared to control conditions that involved the pairing of landscapes with words. This pattern is consistent with the hypothesis of stronger gesture–language associations, an idea captured by the idea of cognitive embodiment due to the overlap between action semantics and language semantics. Moreover, consideration of how performance was influenced by the demands on controlled processing suggests that gesture may influence linguistic representation of actions in two distinct ways. First, there may be a passive form of interaction, one that is evident in tasks that can be performed via automatic processing. The second form of interaction is evident under controlled and attentive processing, operating at a post-lexical level and associated with both facilitation and interference effects.

Acknowledgments This study was supported by the BSF Grant 2007184 awarded to R. Ivry and M. Lavidor.

Appendix 1

See Table 3.

Table 3 Prime–target pairs examples—original stimuli were presented in Hebrew, English translations were added for clarification

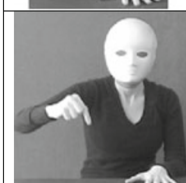
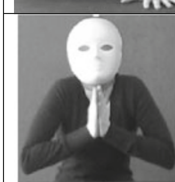
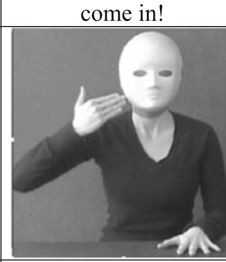
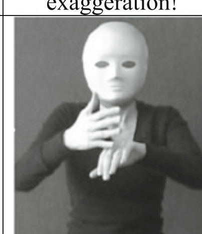
Symbolic gesture prime	Congruent target	UnRelated target (Topic Related)	UnRelated target (Topic UnRelated)
	גרוע - fail	טבעי-natural	שלכת-autumn
	כסף - money	בית-house	עשן-smoke
	עכשיו! – now!	בשמחה! with pleasure	שריפה burning
	נו נו נו! – telling off	הללויה! hallelujah!	סופת חול sand storm
	מתחנן - אני begging	הוא מחליט he decides	עצים ברוח windy trees
	אני חזק - strong	אתה רעב? are you hungry?	פרי נושר fallen fruit

Table 3 continued

Landscape prime	Congruent target	UnRelated target (Topic Related)	UnRelated target (Topic UnRelated)
	רוח - wind	עמק – valley	!אתה-you!
	חוף ים - beach	פרח בר wild flower	ביי ביי! see you!
	ים סוער – storm sea	גן ירוק green garden	אני שבע satiety
	ברקים - bolt	שלולית puddle	הצדעה salute
	זריחה - sunrise	חורשה forest	ייאוש despair
	פריחה - blossom	מערות caves	מצויין splendid

Meaningless gesture prime			
Verbal target	לא נורא! never mind	כנס פנימה! come in!	אתה מגזים exaggeration!
Meaningless gesture			
Verbal target	אתה מבין? you see?	לבד alone	דואג worried

Appendix 2

Stimuli characteristics

See Tables 4, 5, 6, 7 and 8.

Table 4 Prime agreement (conventionality or meaningless agreement score), semantic agreement, lexical agreement, target length and frequency, according to the experimental condition, averaged across the different experimental lists

Prime type	Target	Prime agreement	Semantic agreement	Lexical agreement	Target length	Target frequency
Meaningless gesture	Words	95.20[a] (0.04)			6.73 (2.49)	5.77 (1.21)
	Pseudo-words	95.20[a] (0.04)			6.56 (1.76)	
Symbolic gesture	Congruent	93.99[b] (0.07)	83.63 (0.18)	57.79 (0.25)	6.77 (2.24)	5.84 (1.29)
	Unrelated	94.27[b] (0.07)			6.70 (2.37)	5.77 (1.09)
	Pseudo-words	91.55[b] (0.08)			6.52 (2.09)	
Landscape	Congruent		82.23 (0.18)	70.48 (0.23)	5.56 (2.11)	5.66 (0.87)
	Unrelated				5.56 (2.17)	5.68 (0.85)
	Pseudo-words				5.78 (2.42)	

[a] The 40 MGs used were randomized during the experimental run across the trials (words and pseudo-words)

(1) The SG–congruent and unrelated congruent conditions were matched for conventionality agreement; (2) The congruent–SG and LS conditions were matched for semantic agreement, but not for lexical agreement ($p < 0.01$). Judges tended to name the landscapes in a more consistent manner than the symbolic gestures; (3) On average gesture–targets were lengthier than LS–targets ($p < 0.01$); however, they were matched for number of letters. This difference is due to the use of punctuation marks and higher frequency of two-words in the gesture–targets

[b] With respect to (2) and (3), the experimental results confirm that the lexical agreement and length factors did not advantage the LS–targets over the gesture–targets

(4) All word conditions were matched for target frequency; (5) The different experimental lists were comparable to one another (statistics used: LSD comparisons)

Table 5 Gesture–targets distribution into linguistic categories in each word condition

Prime type	Target	Noun (%)	Adjective (%)	Verb (%)	Adverb (%)
Symbolic gesture	Congruent	24	18	40	21
	Unrelated	25	18	39	21
Meaningless gesture	Unrelated	29	17	38	21

Prime type	Target	Interjection (%)	Pronoun (%)	Preposition (%)	Question (%)
Symbolic gesture	Congruent	12	17	7	1
	Unrelated	11	18	7	0
Meaningless gesture	Unrelated	13	17	4	0

The categories do not complete 100 %, as there are items which belong to two categories. All LS–targets were nouns. The different experimental lists were comparable to one another

Table 6 Categorization of gesture–targets into concrete/abstract

	SG–congruent (%)	SG–unrelated (%)	MG–words (%)
Concrete	40.5	39.3	41.7
Abstract	59.5	60.7	58.3

All LS–targets were nouns. The different experimental lists were comparable to one another

Table 7 Prime agreement (conventionality or meaningless agreement score), target length and frequency, according to the experimental condition, averaged across the different experimental lists

Prime type	Target	Prime agreement	Target length	Target frequency
Symbolic gesture	Congruent	92.91 (0.08)	5.08 (1.66)	5.83 (1.10)
	Incongruent	93.09 (0.08)	4.88 (1.61)	5.70 (1.12)
	Unrelated	93.70 (0.07)	5.03 (1.68)	5.52 (1.24)
	Pseudo-words	88.24 (0.11)	4.96 (1.56)	
Meaningless gesture	Words	85.72 (0.07)	5.36 (1.66)	5.70 (1.15)
	Pseudo-words	84.20 (0.07)	5.28 (1.62)	

The SG–congruent, incongruent and unrelated conditions were matched for conventionality agreement. All conditions were matched for target length. All word conditions were matched for target frequency. The different experimental lists were comparable to one another (statistics used: LSD comparisons)

Table 8 Targets distribution into linguistic categories in each word condition

Prime type	Target	Noun (%)	Adjective (%)	Verb (%)	Adverb (%)	Else (%)
Symbolic gesture	Congruent	42.7	16.0	17.3	20.0	13.3
	Incongruent	42.7	21.3	18.7	25.3	2.7
	Unrelated	42.7	20.0	18.7	20.0	9.3
Meaningless gesture	Unrelated	42.7	16.0	18.7	26.7	10.7

The categories do not complete 100 %, as there are items which belong to two categories. The different experimental lists were comparable to one another

References

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: a matter of some moment? *Journal of Experimental Psychology. Learning, Memory, and Cognition, 27*(2), 514–544.

Bates, E., & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology, 40*(3), 293–310.

Bernardis, P., & Caramelli, N. (2009). Meaning in words, gestures, and mental images. In *Proceedings of the 31st annual conference of the Cognitive Science Society* (pp. 1693–1697).

Bernardis, P., & Gentilucci, M. (2006). Speech and gesture share the same communication system. *Neuropsychologia, 44*(2), 178–190.

Bernardis, P., Salillas, E., & Caramelli, N. (2008). Behavioural and neurophysiological evidence of semantic interaction between iconic gestures and words. *Cognitive Neuropsychology, 25*(7–8), 1114–1128.

Blair, I. V., Urland, G. R., & Ma, J. E. (2002). Using Internet search engines to estimate word frequency. *Behavior Research Methods, Instruments, & Computers, 34*(2), 286–290.

Brookes, H. (2005). What gestures do: some communicative functions of quotable gestures in conversations among Black urban South Africans. *Journal of pragmatics, 37*(12), 2044–2085.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407–428.

Corballis, M. C. (1998). Cerebral asymmetry: Motoring on. *Trends in Cognitive Sciences, 2*(4), 152–158.

Duscherer, K., & Holender, D. (2005). The role of decision biases in semantic priming effects. *Swiss Journal of Psychology, 64*(4), 249–258.

Frick-Horbury, D., & Guttentag, R. E. (1998). The effects of restricting hand gesture production on lexical retrieval and free recall. *The American Journal of Psychology, 111*(1), 43–62.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*(2), 593–609.

Gentilucci, M., Bernardis, P., Crisi, G., & Volta, R. D. (2006). Repetitive transcranial magnetic stimulation of Broca's area affects verbal responses to gesture observation. *Journal of Cognitive Neuroscience, 18*(7), 1059–1074.

Goodwyn, S. W., Acredolo, L. P., & Brown, C. A. (2000). Impact of symbolic gesturing on early language development. *Journal of Nonverbal Behavior, 24*(2), 81–103.

Gunter, T. C., & Bach, P. (2004). Communicating hands: ERPs elicited by meaningful symbolic hand postures. *Neuroscience Letters, 372*, 52–56.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*(2), 301–307.

Holle, H., & Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *Journal of Cognitive Neuroscience, 19*(7), 1175–1192.

Kang, M. S., Blake, R., & Woodman, G. F. (2011). Semantic analysis does not occur in the absence of awareness induced by interocular suppression. *Journal of Neuroscience, 31*(38), 13535–13545.

Kelly, S. D., Creigh, P., & Bartolotti, J. (2010a). Integrating speech and iconic gestures in a stroop-like task: Evidence for automatic processing. *Journal of Cognitive Neuroscience, 22*(4), 683–694.

Kelly, S. D., Özyürek, A., & Maris, E. (2010b). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science, 21*(2), 260–267.

Kendon, A. (1994). Do gestures communicate? A review. *Research on language and social interaction, 27*(3), 175–200.

Kendon, A. (2004). Gesture: Visible action as utterance. Cambridge University Press, Cambridge.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge: Cambridge University Press.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48*(1), 16–32.

Krauss, M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261–283). New-York: Cambridge.

Krauss, R., & Hadar, U. (1999). The role of speech-related arm/hand gesture in word retrieval. In L. S. Messing & R. Campbell (Eds.), *Gesture, speech, and sign* (pp. 93–116). Oxford: Oxford University Press.

Lakoff, G. (1987). Women, fire, and dangerous things. Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). Philosophy in the flesh: The embodied mind and its challenge to Western thought. New York: Basic books.

Lovseth, K., & Atchley, R. A. (2010). Examining lateralized semantic access using pictures. *Brain and Cognition, 72*(2), 202–209.

McNamara, T. P. (2005). Semantic priming: Perspectives from memory and word recognition. New York: Psychology Press.

McNeill, D. (1992). *Hand and mind. What the hands reveal about thought*. Chicago: University of Chicago Press.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Learning, Memory, 18*(3), 615–622.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology, 20*, 750–756.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In *Basic processes in reading: Visual word recognition, 11.*

Neely, J. H., & Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective/retrospective processing theory. In G. H. Bower (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 24, pp. 207–248). New York: Academic Press.

Obermeier, C., Holle, H., & Gunter, T. C. (2011). What iconic gesture fragments reveal about gesture-speech integration: when synchrony is lost, memory can help. *Journal of Cognitive Neuroscience, 23*(7), 1648–1663.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Oliveri, M., Finocchiaro, C., Shapiro, K., Gangitano, M., Caramazza, A., & Pascual-Leone, A. (2004). All talk and no action: a transcranial magnetic stimulation study of motor cortex activation during action word production. *Journal of Cognitive Neuroscience, 16*(3), 374–381.

Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research, 66*(3), 180–194.

Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience, 17*(6), 884–892.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7*(4), 226–231.

Rimé, B., & Schiaratura, L. (1991). Gesture and speech. In R. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 239–281). Cambridge: Cambridge University Press.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences, 21*(5), 188–194.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience, 2*(9), 661–670.

Schegloff, E. A. (1984). On some gestures' relation to talk. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 266–296). Cambridge: Cambridge University Press.

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*(2), 273–281.

Wu, Y. C., & Coulson, S. (2007). Iconic gestures prime related concepts: An ERP study. *Psychonomic Bulletin & Review, 14*(1), 57–63.

Wu, Y. J., & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *Journal of Neuroscience, 30*, 7646–7651.

Xu, J., Gannon, P. J., Emmorey, K., Smith, J. F., & Braun, A. R. (2009). Symbolic gestures and spoken language are processed by a common neural system. *Proceedings of the National Academy of Sciences, 106*(49), 20664–20669.

Yap, D. F., So, W. C., Melvin Yap, J. M., Tan, Y. Q., & Teoh, R. L. S. (2011). Iconic gestures prime words. *Cognitive Science, 35*(1), 171–183.

Zwaan, R. A. (2004). The immersed experiencer: Towards an embodied theory of language comprehension. In B. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 44, pp 35–62). San Diego: Academic Press.