Language and imagery: effects of language modality

Gabriella Vigliocco1,*, David P. Vinson1, Tyron Woolfe2, Matthew W. G. Dye3 and Bencie Woll2

1Department of Psychology, University College London, Gower Street, London WC1H 0AP, UK
2Department of Language and Communication Science, City University, Northampton Square, London EC1V 0HB, UK
3Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627, USA

Across spoken languages, properties of wordforms (e.g. the sounds in the word hammer) do not generally evoke mental images associated to meanings. However, across signed languages, many signforms readily evoke mental images (e.g. the sign HAMMER resembles the motion involved in hammering). Here we assess the relationship between language and imagery, comparing the performance of English speakers and British sign language (BSL) signers in meaning similarity judgement tasks. In experiment 1, we found that BSL signers used these imagistic properties in making meaning similarity judgements, in contrast with English speakers. In experiment 2, we found that English speakers behaved more like BSL signers when asked to develop mental images for the words before performing the same task. These findings show that language differences can bias users to attend more to those aspects of the world encoded in their language than to those that are not; and that language modality (spoken versus signed) can affect the degree to which imagery is involved in language.

Keywords: imagery; language; semantics; sign languages

1. INTRODUCTION

Language is traditionally considered to be propositional in nature and not iconic, abstracted away from any visual, spatial and motoric connotations of the corresponding experiences in the real world. As one most notable example of the abstract nature of linguistic representations, the relationships between wordforms (e.g. the sounds making up the word hammer in English) and referents are largely arbitrary (e.g. the pronunciation of hammer does not bear any resemblance to the form of a hammer, the sound it makes, etc.; de Saussure 1959). The arbitrariness in this relation is highlighted by the variation across languages in wordforms referring to the same objects (e.g. hammer in English, martello in Italian and kakas in Hungarian). This is not to say that all wordforms are arbitrary, onomatopoetic words (such as boom, ping in English) being one such class, but these represent a very small repertoire of exceptional words.

The auditory–vocal modality, however, can be considered an impoverished medium for creating iconic forms because most referents and actions have no associated auditory imagery. In sharp contrast, signed languages, created in space using the signer’s body and perceived visually, exploit the potential for a less arbitrary (i.e. more visually motivated) expression of a wide range of basic conceptual features (e.g. shape, movement, location, human actions) to a far greater extent (Taub 2001). For example, the sign HAMMER is produced by moving a fist up-and-down, resembling the typical use of a hammer, and the sign SCISSORS is produced by moving the index and medium fingers outward and inward, as if they were a pair of scissors. This correspondence suggests that different degrees of arbitrariness may arise as a consequence of modality, although it is important to note that despite the fact that visually motivated signs resemble the corresponding concrete objects/actions, they are clearly not equivalent to pantomimes (Corina et al. 1992; Emmorey 2002; Emmorey et al. 2004). First, they are specifically conventionalized and constrained by the structural features of a given signed language (Emmorey 2002). Second, dissociations have been described in aphasic signers between retained ability to pantomime, for instance, how to use a hairbrush, and impaired ability to retrieve the sign HAIRBRUSH (which is iconic of brushing the hair; Corina et al. 1992; Marshall et al. 2004; see also Emmorey et al. 2004).

We carried out two experiments designed to investigate whether there are cross-linguistic differences between a signed (BSL) and a spoken (British English) language that can be plausibly linked to the visual motivation of signforms (but not wordforms), and whether these differences can be related to differing recruitment of imagistic information in the two languages. We employ the triadic similarity judgement task, a task which has been successfully used in previous studies investigating semantic organization and its impairments (Fisher 1994; Romney et al. 1997; Grossman 1998; Garrard et al. 2004). These experiments focus upon three semantic classes: ‘tools’ (e.g. scissors, broom); ‘tool-actions’ (e.g. drilling, drawing) and ‘body-actions’ (e.g. slapping, punching). Signs from all three classes are visually motivated in BSL, but crucially, the signs for tool-actions and tools both evoke images of using tools, a property differentiating them from signs referring to body-actions which, instead, evoke images of the body movements involved in the
corresponding actions (see figure 1). No such difference is present for the English words, as none of the wordforms used in these semantic classes are iconically linked to their referents. In Experiment 1 we compared similarity judgements of BSL signers and English speakers for translation-equivalent signs/words from these semantic classes, in order to test whether language-specific effects stemming from visual motivation of BSL signforms could be observed.

Tools, tool-actions and body-actions are particularly interesting for the following reasons. Tool-actions and body-actions are both actions, thus differing from tools, which are objects. Previous work in English shows that the conceptual distinction between actions and objects (Gentner 1982) is a fundamental principle of meaning organization (Vigliocco et al. 2004). Moreover, the conceptual distinction between actions and objects coincides with the linguistic distinction between verbs (tool-actions and body-actions) and nouns (tools). Thus, English speakers and BSL signers alike may judge tool-actions and body-actions as the most similar among the semantic classes investigated.

However, tool-handling has been argued to be a fundamental dimension underlying the neural representation of concepts (e.g. Rizzolatti & Arbib 1998; Chao & Martin 2000) regardless of language. Considering tools, tool-actions and body-actions, we might expect tools and tool-actions to be judged as most similar on this basis (by both BSL signers and English speakers). Crucially, however, tool handling is evoked by the signform in BSL, whereas it is not evoked by the English wordform. Therefore, the visually motivated properties of signs referring to tools and tool-actions (e.g. the typical action associated with hammering, for the sign HAMMER; sweeping, for the sign BROOM; etc.) may be more salient in signers' mental representations for tools and tool-actions than for English speakers, and may therefore lead to cross-linguistic differences.

To anticipate the results of experiment 1, we found a cross-linguistic difference: whereas English speakers judged tool-actions as more similar to body-actions, thus distinguishing actions from objects, signers judged tool-actions as more similar to tools than to body-actions, thus distinguishing signs related to tool-handling from other actions. In experiment 2, we used the same task, and wordforms from the same semantic classes of tool-actions, body-actions and tools, to assess the extent to which imagery might mediate the difference between BSL and English. This experiment investigated whether English speakers would perform in a manner more similar to the signers, when the tool-handling component of the meanings of tool-actions and tools is made salient in the task. In order to increase the salience of tool-handling, we asked one group of English speakers to develop a mental image for the action/object described by the word before judging similarity. We compared their performance with that of another group of English speakers who were not asked to create a mental image, but were given the same instructions as the (English-speaking) participants in experiment 1.

2. MATERIAL AND METHODS

(a) Experiment 1. Triadic similarity judgements in BSL and English

(i) Materials and procedure

Twenty four translation-equivalent signs/words were chosen, eight referred to tools, eight referred to tool-actions and eight to body-actions. We avoided form similarity between signs/words by using signs/words that were not visually/acoustically highly similar to each other (especially avoiding ‘minimal pairs’ including a tool and its associated action, e.g. hammer/hammering or sweeping/broom). To make it clear whether a sign/word referred to an object or action (i.e. in English, the bare word hammer is ambiguous between a tool and a tool-action), we presented the tool-actions and body-actions in the progressive verbal form (e.g. hammering) in both languages. In BSL the progressive form increases the form similarity among the signs only to a limited extent; in English, it increases form similarity to a greater extent (i.e. all action-words end in ‘-ing’).

A native BSL model, naive to the goals of the experiment, was asked to produce the signs without English mouth patterns. In half of the tool and tool-action signs, the handshape represented the tool itself (e.g. SCISSORS), in the other half the hand shape represented handling the tool (e.g. SEWING). English words were recorded by a native speaker of British English. They were presented acoustically (rather than visually), in order to maintain the same temporal properties of presentation across the two languages, and to ensure that participants could not make their judgements on the basis of perception alone. However, because acoustic presentation does not maintain the same modality between languages, we also replicated experiment 1 in the visual modality, conducting a follow-up experiment with an additional group of 24 English speakers to ensure that any differences were not simply due to the contrast between auditory and visual modalities.

BSL signers were presented with triplets of signs (e.g. SCISSORS, SEWING, PUSHING), as video clips embedded in a MICROSOFT POWERPOINT presentation on a computer screen. English speakers were presented with the same triplets of acoustically presented words. In both languages, signs/words were presented sequentially, followed by a decision task in which participants were asked to indicate which two signs/words out of the three were the most similar in meaning. Examples and instructions made clear that the decision had to be made on the basis of meaning rather than form similarity among the signs/words; and that there were no ‘right’ or ‘wrong’ responses.

(ii) Design and analysis

All possible combinations of three signs/words from the 24 items (total of 2024 triads) were presented to participants, each of whom rated a subset of the triads. Separate analyses of
variance were conducted treating subjects and items (referring to tool-actions only) as the random factor. The factors included in the experimental design were: language (BSL versus English, manipulated between subjects and within (translation-equivalent) items), and pairing-type (item pairs including one tool-action and one tool, and item pairs including one tool-action and one body-action). The dependent variable was the proportion of word pairs of a given type selected as ‘similar’ by each participant (subjects analysis), or the number of times each word/sign referring a tool-action was selected along with words/signs of each type (items analysis).

(iii) Participants
Ten native BSL signers (deaf adults with deaf parents; median age = 21, range 16–50) and 10 native English speakers individually matched to the BSL signers on age (median age = 20.5, range = 18–48) and sex (3 men, 7 women) were presented with triplets of signs/words and asked to choose the two most similar in meaning. In both languages, each of the 10 participants completed 202 or 203 triads.

(b) Experiment 2. Effects of imagery upon triadic similarity judgements in English

(i) Materials and procedure
Twenty-four words were used, eight referred to tools, eight referred to tool-actions and eight to body-actions. These words were not the same as those used in experiment 1 as we chose only words that were not ambiguous between objects and actions, or in case of object–action homophones, for which the frequency of usage (Baayen et al. 1995) was dominant for use as a noun (for objects) or as a verb (for actions). This allowed us to present verbs in their bare forms (without ‘-ing’), avoiding the form overlap for the actions present in experiment 1.

Words in each triad were printed in 10 point (capitalized) Times New Roman font, separated by dashes; two columns of triads were presented per page. Participants were randomly assigned to one of the two groups. One group received the same instructions as in experiment 1 (i.e. to decide which two words of a triad were most similar in meaning). The other group received ‘imagery’ instructions (‘This study concerns the impact of mental imagery in making decisions on word meanings. You will see sets of three words. For each word you should develop a mental image that includes your typical experience with the thing or action described by the word. For example, given the word ‘cat’ you might imagine seeing a cat and stroking it. Once you have developed a mental image for the three words, we would like you to decide which two of the three are more similar in meaning…’).

(ii) Design and analysis
As in experiment 1, the dependent variable was proportion of word pairs of a given type selected as similar. Analyses were conducted using ANOVA by subjects and items. The experimental design included: instructions (imagery versus no-imagery; manipulated within items, and because participants under different instruction conditions were presented with matched lists, also treated as within-subjects), and pairing-type (tool-actions & tools versus tool-actions & body-actions).

(iii) Participants
Twenty-four native English speakers (median age = 19, range = 17–24) were asked to form mental images for each word before expressing their judgements, and 24 additional native speakers (median age = 19, range = 18–26) received the same instructions as in experiment 1. Each participant completed 337 or 338 triads. All participants gave their informed consent before participating.

3. RESULTS AND DISCUSSION

(a) Experiment 1
Proportion of selection of pairings of different types, reported in figure 2, differed in BSL and English: BSL signers judged tool-actions as more similar to tools than to body-actions; in sharp contrast, English speakers judged tool-actions as more similar to body-actions than to tools.

In the ANOVA, the main effect of pairing-type was not significant ($F_{1,18}=2.187, p=0.183$); there was no general difference between similarity ratings for the two types of pairings. The main effect of language was significant ($F_{1,18}=29.376; p<0.001$; $F_{1,7}=26.380, p=0.001$) indicating that BSL signers overall tended to rate sign-pairs of both types as more similar than did English speakers. Crucially, there was a strong interaction between pairing-type (tool-actions/tools versus tool-actions/body-actions) and language (BSL versus English): ($F_{1,18}=71.873; p<0.001$; $F_{1,7}=129.941, p<0.001$). While BSL signers judged tool-actions as more similar to tools than to body-actions, English speakers judged tool-actions as more similar to body-actions than to tools.

For example, SAWING and SPANNER were chosen as similar by BSL users on 77% of the triads in which those two items occurred, while English speakers chose them only 5% of the time; in contrast, SAWING and SCRATCHING were chosen as similar by BSL users on only 5% of triads, but English speakers chose them 50% of the time.

This is a novel and surprising finding. Whereas the pattern of results for English speakers reflects the conceptual distinction between objects and actions, as well as the linguistic distinction between nouns and verbs, the results from the signers show the impact of the visually motivated properties of the signforms in rendering tool-handling a salient property of meaning shared between signs referring to tool-actions and signs referring to tools.
These findings demonstrate that the arbitrary relation between form and meaning cannot be taken as a foundation assumption for studies of word meaning that aim at establishing universal properties of languages. Before exploring further the implications of these results concerning the relationship between language and imagery, let us consider and dismiss some alternative explanations for the cross-linguistic difference we observed.

First, the fact that English speakers considered tool-actions as more similar to body-actions than to tools cannot be accounted for in terms of the form overlap between the words referring to actions (which all ended in ‘-ing’) because, as we will see in experiment 2, we replicated the results for English speakers when bare verbs (without the ‘-ing’) were used. Second, the cross-linguistic difference cannot be accounted for in terms of a difference between presentation in the visual and acoustic modalities: the follow-up English experiment in the visual modality also replicated these findings. Finally, because BSL signers and British English speakers were presented with the same instructions, and because they live in the same cultural environment, the cross-linguistic difference we observed between the two groups cannot be accounted for in terms of other culture-based factors (e.g. different interpretation of the task by the two groups). Thus we take our results to genuinely reflect the manner in which properties of meaning similarity among the same referents (tool-actions, body-actions and tools) can be modulated by language-specific properties.

This cross-linguistic difference, we argue, arises because signs referring to tool-actions and to tools are both iconic of tool-handling in BSL. If this is the case, increasing the salience of tool-handling should cause English speakers’ performance to become more similar to the performance of signers. This was tested in experiment 2.

(b) Experiment 2

Results of this experiment are shown in figure 3. Under imagery instructions English speakers judged tool-actions to be more similar to tools, compared to the speakers to whom no imagery instructions were given. In the ANOVA, the main effect of pairing-type (tool-actions/tools versus tool-actions/body-actions) was significant \( F(1,17) = 15.583, p = 0.001; F(3,17) = 19.966, p = 0.004 \), indicating that overall, tool-actions were judged to be more similar to body-actions than tools. The main effect of instruction was significant only by items \( F(1,17) = 2.244, p = 0.152; F(3,17) = 7.353, p = 0.030 \), indicating a tendency for participants under imagery instructions to select word pairs of both types more often. Crucially, there was a significant interaction between instructions (imagery versus no-imagery) and pairing-type (tool-actions/tools versus tool-actions/body-actions); \( F(1,17) = 9.625, p = 0.006; F(3,17) = 54.823, p < 0.001 \), such that the effect of pairing-type was confined to the no-imagery condition; with imagery instructions, there was no significant difference in similarity ratings between the two pairing-types \( t(17) = 1.081, p = 0.297; t(21,7) = 1.594, p = 0.155 \).

These results support our claim that the cross-linguistic difference we observed in experiment 1 is mediated by the imagistic properties of signs that render tool-handling imagery shared between tool-actions and tools salient (or at least provide a transparent mapping between meaning and sign form) and therefore, overwriting any preference to distinguish between objects and actions, and between nouns and verbs.

4. GENERAL DISCUSSION

Our two experiments show that meaning similarity among translation-equivalent signs/words can differ by virtue of differences in language properties. In BSL, signs referring to tool-actions were judged as semantically more similar to signs referring to tools than to signs referring to body-actions. In contrast, English speakers judged words referring to tool-actions as more similar to the body-actions than to the tools. Importantly, however, the English speakers’ preference for judging tool-actions as more similar to body-actions than to tools was eliminated (although not reversed, as was the case for BSL signers for whom action imagery may be more established) when participants were asked to make their meaning similarity judgements after developing a mental image related to action or use of an object.

There are two broad implications of these findings. First, we showed that properties of the language used by a community can bias users of that language to attend to certain properties of the world rather than others (Slobin 1996). In principle, given our translation-equivalent signs/words, both signers and speakers could have chosen to cluster signs/words based on the conceptual distinctions (a) between actions and objects, or (b) between words related to tool-handling and words related to the body. But this is not what we found. Crucially, the linguistic difference we found has implications regarding which conceptual distinctions are more salient: the distinction between words related to using tools and other concepts for BSL signers and the distinction between actions and objects (coinciding with the distinction between verbs and nouns) for the English speakers. Importantly, the salience of these distinctions is not only modulated by differences in language properties, but also by the characteristics of the task (as we have shown in experiment 2), thus underscoring the flexibility of our mental representations (Barsalou et al. 2003).

Figure 3. Proportion of the different pairing-types (tool-actions/tools and tool-actions/body-actions) selected in the two instruction conditions (imagery and semantic) for experiment 2. Dark columns depict selection proportions for tool-action/tool word pairs; light columns depict selection proportions for tool-action/body-action word pairs (error bars depict standard errors by subjects).
Second, considering experiments 1 and 2 together, our results illuminate the relation between language and imagery, linking visually motivated properties of signs to imagery, or at least those aspects of meaning reflected in the form of the signs. Thus they question the traditional view in cognitive psychology, developed primarily on the basis of research on spoken languages, according to which language is abstracted away from our sensorial and motoric experience of the world, utilising abstract propositional representations stripped from any imagistic connotations (e.g. Miller & Johnson-Laird 1976).

Other findings in the literature on signed languages also suggest a close link between language and imagery. For instance, Emmorey et al. (1993) showed that American Sign Language (ASL) signers form visual images better than do English speakers. Moreover, activations in parietal areas (particularly the superior parietal lobule), traditionally associated with mental imagery (e.g. Kosslyn et al. 2001) have been consistently reported in brain imaging studies of signed languages (and cannot be ascribed simply to the greater degree of visuo-manual coordination engaged in signed rather than spoken languages). These activations have been found in studies that have investigated spatial language for which the actual topographic layout is preserved in signing space (Emmorey et al. 2002; MacSweeney et al. 2002) but also studies that have investigated lexical signs (Emmorey et al. 2002, 2004). Crucially, activations of these areas were not reported in spoken languages under the same conditions (Damasio et al. 2001; Emmorey et al. 2002; MacSweeney et al. 2002).

Spoken languages, bound to the auditory modality, have only a limited repertoire of words that can evoke images of the corresponding entity/event in the world (onomatopoeic words, e.g. boom, ping). Because the vast majority of studies have investigated spoken languages, it has been argued that language (as a cognitive system) is separate and independent from imagery. Signed languages, perceived in the visual modality, exploit the rich visuo-spatial potential for a large number of visually motivated signs, and as we have shown, these imagistic properties are salient aspects of the mental representations of the corresponding entities/events in the world. Thus, our study shows that the modality in which languages are expressed modulates the degree of cross-talk between language and imagery, traditionally considered to be separate cognitive modules.

This work was supported by a James S. McDonnell Foundation Pilot Award (#21002048) and Economic and Social Research Council grant (RES000230038) to Gabriella Vigliocco. We would like to thank Tina Kelberman and Katie Roach for modelling the BSL signs, and Art Glenberg for comments.

REFERENCES


