COGNITION

Cognition 85 (2002) B61–B69

www.elsevier.com/locate/cognit

Brief article

Semantic distance effects on object and action naming

Gabriella Vigliocco[a,*], David P. Vinson[a], Markus F. Damian[b], Willem Levelt[c]

[a]*Department of Psychology, University College London, Gower Street, London WC1, UK*
[b]*University of Bristol, Bristol, UK*
[c]*Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*

Received 6 March 2002; received in revised form 17 April 2002; accepted 23 May 2002

Abstract

Graded interference effects were tested in a naming task, in parallel for objects and actions. Participants named either object or action pictures presented in the context of other pictures (blocks) that were either semantically very similar, or somewhat semantically similar or semantically dissimilar. We found that naming latencies for both object and action words were modulated by the semantic similarity between the exemplars in each block, providing evidence in both domains of graded semantic effects. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Semantic memory; Objects and actions; Nouns and verbs; Lexical semantics

1. Introduction

Miller and Fellbaum (1991) wrote: "When psychologists think about the organization of lexical memory it is nearly always the organization of nouns that they have in mind" (p. 214). Even more specifically, we may add, often it is nouns referring to objects that we have in mind.

Although the object-noun domain is certainly relevant to studies of lexical memory, it only represents part of adults' lexical knowledge; theories and tools developed to investigate semantic organization must generalize beyond words for things. In this article we take up this challenge, addressing the question: can we capture semantic relatedness effects in object-noun and action-verb domains using parallel principles and tools?

Meaning similarity among words affects many tasks involving speech production. For

* Corresponding author. Tel.: +44-20-7679-5419; fax: +44-20-7436-4276.
E-mail address: g.vigliocco@ucl.ac.uk (G. Vigliocco).

PII: S0010-0277(02)00107-5

example, speakers are slower in naming a picture when a meaning-related distracter word is presented, relative to an unrelated word (e.g. Glaser & Düngelhoff, 1984). Similar interference effects arise when speakers name pictures in the context of naming other pictures from the same semantic field, relative to naming pictures in the context of other pictures from different semantic fields (semantic context effects; Damian, Vigliocco, & Levelt, 2001; Kroll & Stewart, 1994).

It is generally agreed upon that these effects reflect properties of the conceptually-driven lexical retrieval process (Levelt, Roelofs, & Meyer, 1999). During retrieval, a target lexico-semantic representation is activated, along with other meaning-related representations. Picture-word interference and semantic context effects reflect competition between these different representations (Damian et al., 2001). The reliability of semantic interference effects in production tasks, and the observation that these effects are relatively impervious to lexical dimensions such as frequency, length and phonological overlap (Levelt et al., 1999) render them well-suited to investigate the semantic representation of words in memory. Here, we capitalize on semantic context effects to investigate the fine-grained semantic organization of words referring to objects and actions.

Many previous studies investigating words referring to objects (e.g. Schriefers, Meyer, & Levelt, 1990) and one study concerning actions (Roelofs, 1993) have established differences in naming latencies between semantically *related* and *unrelated* conditions. However, interference effects in other content domains (numbers and colors) have been shown to be graded: modulated by the degree of semantic similarity between the to-be-named target and the distracter. For example, Klopfer (1996) reported that, in a Stroop-like task, words representing colors perceptually similar to the color to be named produced greater interference than words representing perceptually dissimilar colors. Distance effects also occur in the number domain (Brysbaert, 1995; Moyer & Landauer, 1967; Pavesi & Umiltà, 1998). However, colors and numbers may be special content domains because it is easy to describe their primary conceptual dimensions: hue and saturation for colors, quantity for numbers.

What about more complex domains: objects and actions? Because identifying the primary conceptual dimensions is far more difficult in these domains, can graded effects be observed? To date little empirical work has addressed this question (beyond studies using the "release from proactive interference" paradigm[1]; Wickens, 1970; Wickens, Dalezman, & Eggemeier, 1963), although a number of different theories of meaning representation predict semantic distance effects, especially those assuming distributed semantic representations and featural overlap between them (e.g. Martin & Chao, 2001; McRae, de Sa, & Seidenberg, 1997; Plaut, 1995).

Furthermore, semantic representations of objects and actions are different. Within the object domain, category membership has powerful effects, most striking in patients who are selectively impaired or spared in one category of knowledge, such as animals (Caramazza & Shelton, 1998), body-parts (Shelton, Fouch, & Caramazza, 1998) and fruits and vegetables (Hart, Berndt, & Caramazza, 1985). These findings have led some researchers to postulate that domains playing a fundamental role for our survival (e.g. animals, plants

[1] This paradigm has been criticized as a tool to investigate semantic memory because it involves a number of strategic retrieval processes.

and body-parts) are represented categorically in semantic memory within dedicated neural substrates (Caramazza & Shelton, 1998). In this view, semantic distance effects may not be observed between evolutionarily motivated categories. These should act as isolable clusters because they are independent from other domains of knowledge. In contrast, graded effects may be observed between categories which are not evolutionarily motivated. This contrasts with proposals according to which featural overlap (regardless of category membership) determines semantic similarity within and between categories (e.g. Martin & Chao, 2001). Furthermore, it remains an empirical question whether "categorical" and "featural overlap" effects for objects may be disentangled in behavioral tasks.

In the action domain, category boundaries are not as well defined.[2] For example, consider "speaking". Is this verb categorized as "communication", like "teaching", or perhaps as "body-noise", like "snoring"? To account for such differences between objects and actions, Huttenlocher and Lui (1979) proposed that category membership and well-defined hierarchical organization are important organizational principles in the object domain, while the action domain is organized in a matrix-like manner, with exemplars in different fields sharing general properties (e.g. intentionality) crossing semantic fields, and lacking clear hierarchical organization. It is an empirical question whether such differences have consequences for the likelihood of observing graded effects in both domains.

In order to assess graded semantic similarity effects, we operationalized semantic similarity in a way that allows us to select materials in both the object and action domains. We used empirically-based measures of semantic distance for 456 English words (referring to objects and to actions), obtained in the following manner (for detailed descriptions of the methods, see Vinson & Vigliocco, 2002). First, feature norms were obtained by asking speakers of English to generate features that define and describe each word. Second, we used self-organizing maps (SOMs) (Kohonen, 1997) to reduce the dimensionality of the featural space and obtain a semantic space in which each word is represented as the unit best responding to an input vector in the resulting map. Semantic distance is operationalized as the Euclidean distance between two best responding units in the space. This model, hence, serves as an empirical tool to select materials on the basis of semantic distance, without making a priori or arbitrary decisions about the degree of semantic similarity between words.

2. Method

2.1. Participants

Ninety-four native English speakers from the UCL community participated in the experiment in exchange for payments of £3. All had normal or corrected-to-normal vision.

[2] It can be argued that differences in field boundaries in object and action domains are both a matter of degree; objects can also be cross-classified (e.g. "cat" is an animal and a pet) and boundaries are not always clear-cut.

2.2. Materials

Action and object pictures were selected separately based upon semantic distance. Selected items were picturable objects or actions from separable semantic fields. Items were relatively similar within a semantic field, and separable (in terms of semantic distance) between fields. Finally, between-field distances were such that two fields were semantically near each other, while the third was relatively far from the other two (thus allowing us to investigate graded effects of semantic distance between blocks).

Object pictures (24 in total, eight per field) were taken from the semantic fields of vehicles (average within-field distance = 2.71 units), clothing (5.35), and body-parts (7.35). Clothing and body-parts were "near" fields (average distance between exemplars from the two fields = 13.51 units) while the other two pairings were "far": vehicles and clothing (18.30) and vehicles and body-parts (18.53). Most object pictures were taken from Snodgrass and Vanderwart (1980) with a few specifically prepared for our purposes.

For action pictures, 24 exemplars were found within fields roughly identifiable as "body actions" (typically manner of motion; average within-field distance = 7.66), "tool actions" (11.44), and "actions involving the mouth" (12.21). Body and tool actions were "near" fields (16.74), the other two pairings "far": body-mouth (21.60) and tool-mouth (20.45). Action pictures were taken from Druks and Masterson (2000) or prepared for our purposes. Action distances were slightly greater than object distances, a consequence of the necessity that action pictures be distinguishable from each other. Items are listed in Appendix A.

Visual similarity ratings were collected for objects and actions, following the procedure used in Damian et al. (2001). The scale ranged from 1 (not similar at all) to 5 (very similar) (see Table 1 for ratings). All conditions differed significantly, with the exception of near-far for actions. However, visual similarity was very low overall regardless of condition.

Parallel experimental lists were prepared for object and action items as follows. In blocks of the "Same field" condition, all eight exemplars in a given field were included. Two distinct blocks of "same field" items were included for each semantic field, with half (randomly selected) treated as fillers (e.g. for clothing, one block would include [belt, sock, hat, waistcoat] as target items, and the others as fillers; the second block was reversed). For blocks of the "Near field" condition, four exemplars from both "nearby" fields (body-parts + clothing for objects, body and tool actions for actions) were included; finally, for blocks of the "Far field" condition, four exemplars were selected from two distant fields (body-parts + vehicles, or clothing + vehicles for objects; body + mouth or tool + mouth actions). For both objects and actions, 12 different specific blocks were

Table 1
Average visual similarity ratings between objects and actions as a function of semantic condition ($N = 4$)

	Semantic condition		
	Same	Near	Far
Objects	1.93	1.60	1.27
Actions	1.71	1.60	1.50

created. Within a block, each of the eight pictures was repeated four times, in pseudorandom sequence (each item was sampled once before any item was repeated in a block, and no item appeared twice in succession). Each specific block was presented twice in a given experiment; blocks were presented in pseudorandom order for each participant (each block was sampled once before any block was repeated in the experiment, and no specific block type appeared twice in succession). Hence an experimental list consisted of 24 blocks of 32 trials each (objects and actions were in separate lists).

Stimuli were presented using E-Prime experimental software (Schneider, Eschmann, & Zuccolotto, 2002) on IBM-PC compatible computers; response latencies were collected using a PST Serial Response Box (Psychology Software Tools) and tape-recorded for error analysis.

2.3. Procedure

Participants were assigned to a Word Type condition (object naming, $n = 40$; or action naming, $n = 54$). They were told the experiment focused upon spoken reaction times: pictures would be presented on the computer screen, and their task was to say the picture name aloud as quickly and accurately as possible. Objects were named using bare nouns, actions using stem + "ING". The experiment began with an untimed phase in which each picture was presented for the participant to name. Experimenters provided the correct label in the (few) instances in which participants failed to recognize or name a picture. Then, participants were presented with a practice block in which each of the pictures was presented once, in random order. Immediately after, the experimental blocks were presented.

Each block began with a button press from the participant. A fixation cross appeared on screen for 300 ms, followed by a blank screen of 450 ms. The target picture appeared in the center of the screen and remained until the voice key detected a response, or for 2500 ms if no response was detected. Responses were followed by a 200 ms blank screen, followed by the fixation cross for the next trial. Each session was tape-recorded and scored for accuracy.

3. Results

3.1. Errors

Each participant's responses were scored for errors, including failure to detect initial word onset, false detections, and erroneous or dysfluent utterances (6.9% of all trials). Analysis of variance (ANOVA) showed no significant effect of semantic context (same, near, and far) on the number of errors, either for object or action naming (all $F < 1$). Error trials were then removed, as were response latencies faster than 250 ms or slower than 1500 ms (1.6% of all trials).

3.2. Response latencies

Fig. 1 reports the response latencies for objects and actions in the different semantic context conditions.

First, a 2 (word type: objects vs. actions; between subjects and items) × 3 (semantic context: same, near and far field; within subjects and items) ANOVA was carried out. The main effect of word type was significant ($F1(1, 92) = 11.61$, $P < 0.001$; $F2(1, 30) = 14.54$, $P < 0.001$) reflecting longer naming latencies for actions than objects. The main effect of semantic context was also significant ($F1(2, 184) = 19.01$, $P < 0.001$; $F2(2, 60) = 16.14$, $P < 0.001$), reflecting the effects of semantic distance. However, crucially, the interaction of the two factors was not significant ($F1(2, 184) = 1.34$, $P = 0.26$; $F2(2, 60) < 1$), indicating the same pattern of semantic context effects for both.

Separate (because semantic fields were at different distances for object and action items) linear trend analyses were then performed on the simple main effects of semantic context to assess the role of semantic distance. Linear contrasts were calculated on the basis of the semantic distances between items in Same Field, Near Field, and Far Field blocks, as intervals were not equidistant. The resulting contrast coefficients were [−6.4, 0.8, 5.6] for objects, and [−6.2, 1.0, 5.2] for actions. The linear trend was significant for objects ($F1(1, 39) = 20.87$, $P < 0.001$; $F2(1, 15) = 18.30$, $P < 0.001$); it was also significant for actions ($F1(1, 53) = 24.44$, $P < 0.001$; $F2(1, 15) = 21.73$, $P < 0.001$). Corresponding quadratic trends were not significant (all $F < 1$).

4. Discussion

We obtained parallel, graded patterns of semantic context interference for separate sets of object nouns and action verbs, selected not on the basis of a priori category membership, but based upon semantic distances obtained from a model of lexical-semantic representation derived from speaker-generated features (Vinson & Vigliocco, 2002).

It is important to note here that semantic distance is correlated with visual similarity (especially in tasks using pictures; see Vitkovitch, Humphreys, & Lloyd-Jones, 1993) and visual similarity tended to differ across conditions in our study. We believe, however, that visual similarity per se cannot account for our results. First, rated visual similarity was overall low. Second, Damian et al. (2001) reported semantic context effects for objects when visual similarity between pictures was minimized. Finally, Vigliocco and Vinson (2002) reported that semantic distance significantly predicted the likelihood of committing

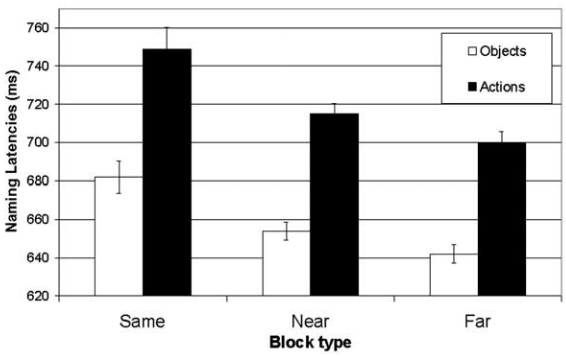

Fig. 1. Naming latencies as a function of word and block type.

semantic errors (e.g. saying "hand" for "foot") even when visual similarity (established in a rating task) between pictures was taken into account in a regression model. Most importantly, they also showed that semantic similarity, but not visual similarity, predicted performance for actions. In our study, visual similarity ratings did not significantly differ for near and far conditions for actions, nonetheless graded semantic effects were observed.

Our results are important from a number of different perspectives. First, from a methodological standpoint, they demonstrate the possibility of going beyond assessing semantic effects between semantically related and unrelated items, by using empirically-driven measures of semantic similarity in domains in which the primary conceptual dimensions are not easily described. In these regards, our results provide evidence concerning the value of using speaker-generated features to derive measures of semantic similarity converging with the work of McRae and Boisvert (1998) who showed how some of the conflicting priming results reported in the literature could be resolved when semantic similarity was re-defined on the basis of empirical measures. They further provide converging evidence with the work by McRae et al. (1997) who showed that meaning similarity (in terms of individual and correlated features) significantly predicted priming (in a semantic decision task) and similarity ratings in the object domain.

Second we found graded semantic effects between object categories, suggesting that categorical effects are not all-or-nothing. Namely, it is not the case that category boundaries determine whether interference effects are present (within categories) or absent (between categories). This fact is especially interesting because one of our critical fields was body-parts, a category that can be considered to be a good candidate for being evolutionarily motivated (Shelton et al., 1998). If this field acted as an isolable category, we should have observed comparable interference effects for body-parts in the context of clothing (semantically near, in our measures) or vehicles (semantically far). Instead we found that reaction times for clothing + body-part blocks were significantly slower than for vehicles + body-part blocks. Hence our findings argue against any special status for category boundaries independent of semantic distance among exemplars. Our data, instead, are compatible with theories that assume a featural organization (e.g. Martin & Chao, 2001).

Finally and most importantly, we showed that we can capture important aspects of the semantic organization of objects and actions using the same empirical tool. The semantic distances we used here were obtained in the same manner for the object and action words, crucially suggesting that the two domains follow the same underlying organizational principles. Importantly this claim does not require that the semantic space for objects and actions is similarly organized. As discussed in Vigliocco and Vinson (2002) and Vinson and Vigliocco (2002), the semantic space resulting from the speaker-generated features differs for objects and actions. In particular, it is "lumpy" for object concepts with clusters of tightly packed concepts belonging to a category and with well-defined boundaries between categories. For actions, instead, the space is smooth with less dense neighborhoods and no well-defined boundaries between different clusters (as was suggested by Huttenlocher & Lui, 1979). However, these differences emerge not as a consequence of using different tools but in terms of differences in the properties of the speaker-generated features (such as: different content of the listed features; different numbers of features listed in the two domains; and differences in the degree of correlation among features).

Despite these differences, we have demonstrated that empirically derived measures of semantic similarity generated using the same tools in both domains successfully predict the magnitude of interference effects in a naming task.

Acknowledgements

The work reported here was supported by a Human Frontier Science Program Grant (RG148/2000) to Gabriella Vigliocco.

Appendix A. Items used in the experiment

A.1. Objects

Body-parts: arm; finger; foot; hand; leg; neck; shoulder; thumb.
Clothing: belt; glove; hat; shirt; shoe; sock; trousers*; waistcoat*.
Vehicles: airplane; bicycle; bus; car; helicopter; lorry*; motorcycle; train.
*Semantic features were obtained for US English translation equivalent.

A.2. Actions

Body actions: hop; kick; march; run; sit; slide; stop; walk.
Mouth actions: drink; eat; frown; smile; sneeze; spit; taste; yawn.
Tool actions: cut; dig; draw; drill; paint; rake; saw; shovel.

References

Brysbaert, M. (1995). Arabic number reading: on the nature of the numeric scale and the origin of phonological recoding. *Journal of Experimental Psychology: General*, *124*, 434–452.

Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: the animate-inanimate distinction. *Journal of Cognitive Neuroscience*, *10*, 1–34.

Damian, M. F., Vigliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*, B77–B86.

Druks, J., & Masterson, J. (2000). *An object and action naming battery*. London: Psychology Press.

Glaser, W. R., & Düngelhoff, F. -J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 640–654.

Hart, J., Berndt, R. S., & Caramazza, A. (1985). Category-specific naming following cerebral infarction. *Nature*, *316*, 439–440.

Huttenlocher, J., & Lui, F. (1979). The semantic organization of some simple nouns and verbs. *Journal of Verbal Learning and Verbal Behavior*, *18*, 141–179.

Klopfer, D. S. (1996). Stroop interference and color-word similarity. *Psychological Science*, *7*, 150–157.

Kohonen, T. (1997). *Self-organizing-maps*, New York: Springer.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Current Opinion in Neurobiology*, *11*, 194–201.

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*, 558–572.

McRae, K., de Sa, V., & Seidenberg, M. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*, 99–130.

Miller, G. A., & Fellbaum, C. (1991). Semantic networks of English. *Cognition*, *41*, 197–229.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*, 1519–1520.

Pavesi, A., & Umiltà, C. (1998). Symbolic distance between numerosity and identity modulates Stroop interference. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1535–1545.

Plaut, D. C. (1995). Double dissociation without modularity: evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, *17*, 291–321.

Roelofs, A. (1993). Testing a non-decompositional theory of word retrieval in speaking: retrieval of verbs. *Cognition*, *47*, 59–87.

Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh, PA: Psychology Software Tools.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: picture-word interference studies. *Journal of Memory and Language*, *29*, 86–102.

Shelton, J. R., Fouch, E., & Caramazza, A. (1998). The selective sparing of body-part knowledge: a case study. *Neurocase*, *4*, 339–351.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *6*, 174–215.

Vigliocco, G., & Vinson, D. P. (2002). The meaning of object and action words, submitted for publication.

Vinson, D. P., & Vigliocco, G. (2002). A semantic analysis of grammatical class impairments: semantic representations of object nouns, action nouns, and action verbs. *Journal of Neurolinguistics*, *15*, 317–351.

Vitkovitch, M., Humphreys, G. W., & Lloyd-Jones, T. (1993). On naming a giraffe a zebra: picture naming errors across different categories. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 243–259.

Wickens, D. D. (1970). Encoding categories of words: an empirical approach to meaning. *Psychological Review*, *77*, 1–15.

Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *2*, 440–445.