Cognition 114 (2010) 96–104

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

The linguistic and embodied nature of conceptual processing

Max M. Louwerse [*], Patrick Jeuniaux

University of Memphis, Memphis, TN 38152, United States

ARTICLE INFO

Article history:
Received 14 April 2009
Revised 2 September 2009
Accepted 3 September 2009

Keywords:
Conceptual processing
Embodied cognition
Iconicity
Linguistic factors
Embodiment factors
Semantic judgment
Iconicity judgment
Symbolic
Embodied
Amodal
Modal
Perceptual simulation

ABSTRACT

Recent theories of cognition have argued that embodied experience is important for conceptual processing. Embodiment can be contrasted with linguistic factors such as the typical order in which words appear in language. Here, we report four experiments that investigated the conditions under which embodiment and linguistic factors determine performance. Participants made speeded judgments about whether pairs of words or pictures were semantically related or had an iconic relationship. The embodiment factor was operationalized as the degree to which stimulus pairs were presented in the spatial configurations in which they usually occur (i.e., an iconic configuration, e.g., *attic* presented above *basement*). The linguistic factor was operationalized as the frequency of the stimulus pairs in language. The embodiment factor predicted error rates and response time better for pictures, whereas the linguistic factor predicted error rates and response time better for words. These findings were modified by task, with the embodiment factor being strongest in iconicity judgments for pictures and the linguistic factor being strongest in semantic judgments for words. Both factors predicted error rates and response time for both semantic and iconicity judgments. These findings support the view that conceptual processing is both linguistic and embodied, with a bias for the embodiment or the linguistic factor depending on the nature of the task and the stimuli.

Published by Elsevier B.V.

1. Introduction

Recent theories of cognition have emphasized that embodied experiences are activated in cognitive processing (Barsalou, 1999; Glenberg, 1997; Zwaan, 2004). According to these theories, meaning construction heavily relies on perceptually simulating the information that is presented to the comprehender: when we see the word *rose* we also 'see' its colors, 'feel' its thorns, and 'smell' its fragrance (Pecher & Zwaan, 2005; Semin & Smith, 2008). Over the last decade ample evidence has supported this view. For instance, several studies have shown that when words or their presentations are analogous to their meaning (e.g., the word *ceiling* presented above the word *floor*), processing is easier than when there is not such an iconic relationship between words and their meaning. These iconicity findings have been obtained with words for animals (Šetić & Domijan, 2007), object words (Estes, Verges, & Barsalou, 2008) and motion verbs (Meteyard, Baharami, & Vigliocco, 2007; Richardson & Spivey, 2000). Zwaan and Yaxley (2003) demonstrated an iconicity effect with participants seeing the word *attic* presented above the word *basement* (iconic configuration) or the word *basement* presented above *attic* (reverse-iconic configuration). When participants were asked whether a word pair was semantically related, they responded faster to the iconic pairs than to the reverse-iconic pairs.

Studies like these not only show evidence for embodiment in conceptual processing, but might also give the impression that there is not much more to conceptual processing than the activation of embodied representations. Various theories of embodied cognition have started to

---

[*] Corresponding author. Address: Department of Psychology/Institute for Intelligent Systems, University of Memphis, Psychology Building, Memphis, TN 38152, United States. Tel.: +1 901 678 5017; fax: +1 901 678 2579.

E-mail address: mlouwerse@memphis.edu (M.M. Louwerse).

0010-0277/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.cognition.2009.09.002

acknowledge that conceptual processing tasks involve both perceptual simulations and linguistic structure (Barsalou, Santos, Simmons, & Wilson, 2008; Louwerse & Jeuniaux, 2008; Zwaan, 2008). For instance, there is evidence that language encodes embodied relations (Louwerse, 2008, in press). Consequently, language users might rely on language, on embodied relations, or on both.

However, it remains unclear to what extent conceptual processing is primarily embodied or linguistic. Some have argued that conceptual processing is predominantly embodied (Barsalou et al., 2008), whereas others have argued that it is predominantly linguistic in nature (Louwerse & Jeuniaux, 2008). In fact, even though there is a wealth of recent evidence showing that language tasks evoke perceptual simulations, surprisingly little is known to what extent, and under what conditions, both embodied and linguistic factors are used in conceptual processing.

Louwerse (2008) replicated the iconicity study by Zwaan and Yaxley (2003), showing that word pairs matching embodied experience facilitate processing. In addition, Louwerse (2008) tested whether a linguistic explanation, like frequency of word order, should be ruled out. When the frequency of word pairs in an iconic (*attic–basement*) or a reverse-iconic order (*basement–attic*) was computed, this frequency explained response times better than iconicity did. Finally, when order frequency and iconicity made opposite predictions, order frequency explained response times better than iconicity.

In a nutshell, Zwaan and Yaxley (2003) found evidence that embodiment factors affected semantic judgments, and Louwerse (2008) found evidence that linguistic factors did also. Yet finding a linguistic explanation (i.e., frequency of word pairs) for the results in a task in which participants make semantic judgments about verbal stimuli (e.g., Louwerse, 2008) might not be as remarkable as finding an embodied explanation (i.e., iconicity) for these verbal stimuli (e.g., Zwaan & Yaxley, 2003). In order to conclude that both embodiment and linguistic factors play a role in conceptual processing, two types of studies are needed. On the one hand, evidence for embodiment factors should be obtained for a predominantly linguistic task with verbal stimuli, and on the other hand, evidence for linguistic factors should be obtained for a predominantly embodied task with non-verbal stimuli such as pictures.

The purpose of the current study was to investigate the conditions under which embodiment and linguistic factors dominated conceptual processes. Four experiments were conducted, two of them with word pairs (Experiments 1a and 1b) and two with picture pairs (Experiments 2a and 2b). Participants were involved in a semantic judgment task, in which they were asked to determine whether two words had a similar meaning (Experiments 1a and 2a), or an iconicity judgment task, in which they were asked to determine whether two words had a relation in terms of position on the vertical axis that matched the one found in the physical world for their referents (Experiments 1b and 2b).

The embodiment factor was operationalized as the degree to which stimulus pairs were presented in the spatial configurations in which they usually occur (e.g., *attic* presented above *basement*) (Louwerse, 2008; Zwaan & Yaxley, 2003). The linguistic factor was operationalized as the degree to which stimuli were presented in the order in which they typically occur in language (Benor & Levy, 2006; Louwerse, 2008).

We predicted the linguistic factor to be more prevalent in the processing of verbal stimuli (Louwerse, 2008), and the embodiment factor in the processing of pictorial stimuli (Paivio, 1986). Furthermore, we predicted the linguistic factor to be more predominant in the semantic judgment task than in the iconicity judgment task, and the embodiment factor to be more predominant in the iconicity judgment task than in the semantic judgment task. Crucially, however, we expected both the embodiment and linguistic factors to be relevant in all four experiments, given that conceptual processing is both linguistic and embodied (Barsalou et al., 2008; Louwerse & Jeuniaux, 2008; Zwaan, 2008).[1]

2. Experiment 1a

2.1. Method

2.1.1. Participants

Thirty undergraduate students from the University of Memphis participated for Psychology course credit.

2.1.2. Materials

Sixty-four word pairs were used with each pair presented in the iconic (*attic–basement*), or reverse-iconic (*basement–attic*) configuration (see Appendix). These word pairs were taken from Zwaan and Yaxley's (2003) pool of 71 experimental pairs they used in three experiments. We excluded those word pairs which contained a duplicated word. Each participant saw 32 items in their iconic relations and the remaining 32 items in their reverse-iconic relation. In addition, 128 filler items were used, having no iconic relation. Half of the fillers had a high semantic relation (cos = .55) and half of them a low semantic relation (cos = .21), as determined by latent semantic analysis (LSA), a statistical, corpus-based, technique for estimating semantic similarities on a scale of -1 to 1 (Landauer, McNamara, Dennis, & Kintsch, 2007). All items were counterbalanced such that all participants saw all word pairs, but no participant saw the same word pair in both orders (i.e., both the iconic and the reverse-iconic order for the experimental items).

2.1.3. Procedure

Participants received instructions to judge the semantic similarity of word pairs presented to them on a computer screen (Fig. 1, top row). Upon presentation of a word pair, participants indicated as soon as possible whether the pair was similar in meaning or not by pressing a "yes" or "no"

---

[1] The literature has used "amodal", "symbolic" and "linguistic" as antonyms for "modal", "embodied" and "perceptual". Because we focus here on the frequency of word order, and because word order plays an important role on linguistic structure (Benor & Levy, 2006; Louwerse, 2008), we use "linguistic" instead of what we have called elsewhere (Louwerse & Jeuniaux, 2008) "symbolic".

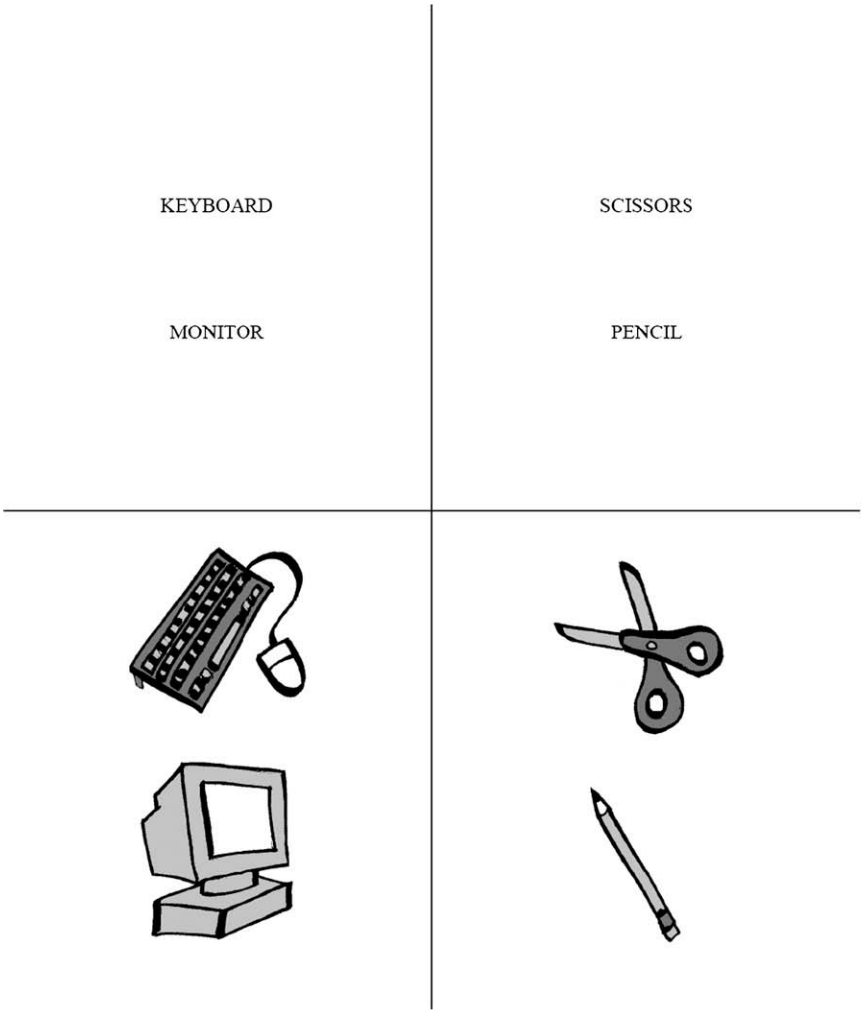

Fig. 1. Examples of stimuli used in Experiments 1a and 1b (top row) and Experiments 2a and 2b (bottom row). An example of an experimental item (reverse-iconic presentation) is presented in the left column, an example of a filler item is presented in the right column.

key. The fillers and experimental items were randomly ordered for each participant.

2.1.4. Factors

2.1.4.1. Order frequency. The linguistic factor was operationalized as the log frequency of *a–b* (e.g., *attic–basement*) and *b–a* (e.g., *basement–attic*) order of word pairs. The order frequency of all 64 word pairs within 3–5 word grams was obtained using the large *Web 1T 5-gram* corpus (Brants & Franz, 2006).

2.1.4.2. Iconicity ratings. The embodiment factor was operationalized as an iconicity rating. We asked 24 participants at the University of Memphis to estimate the likelihood that concepts appeared above one another in the real world. Ratings were made for 64 word pairs on a scale of 1–6, with 1 being extremely unlikely and 6 being extremely likely. Each participant saw all word pairs, but whether a participant saw a word pair in an iconic or a reverse-iconic order was counterbalanced between two groups, such that all participants saw iconic and reverse-iconic word pairs, but no participant saw a word pair both in an iconic and a reverse-iconic order. High interrater reliability was found in both groups (Group A: average $r = .76$, $p < .001$, $n = 64$; Group B: average $r = .74$, $p < .001$, $n = 64$), with a negative correlation between the two groups (average $r = -.72$, $p < .001$, $n = 64$).

2.2. Results and discussion

Zwaan and Yaxley (2003) included only those participants in their analysis who had an error rates less than 15%, and did not conduct analyses with error rate as the dependent variable. The error rate ($M = 17.53$, $SD = 3.8$) in the current experiment was not too far off the range of error rate of Zwaan and Yaxley's (2003) experiments. With no participant outliers, all participants were included in the analysis. The errors can be explained by the relatively

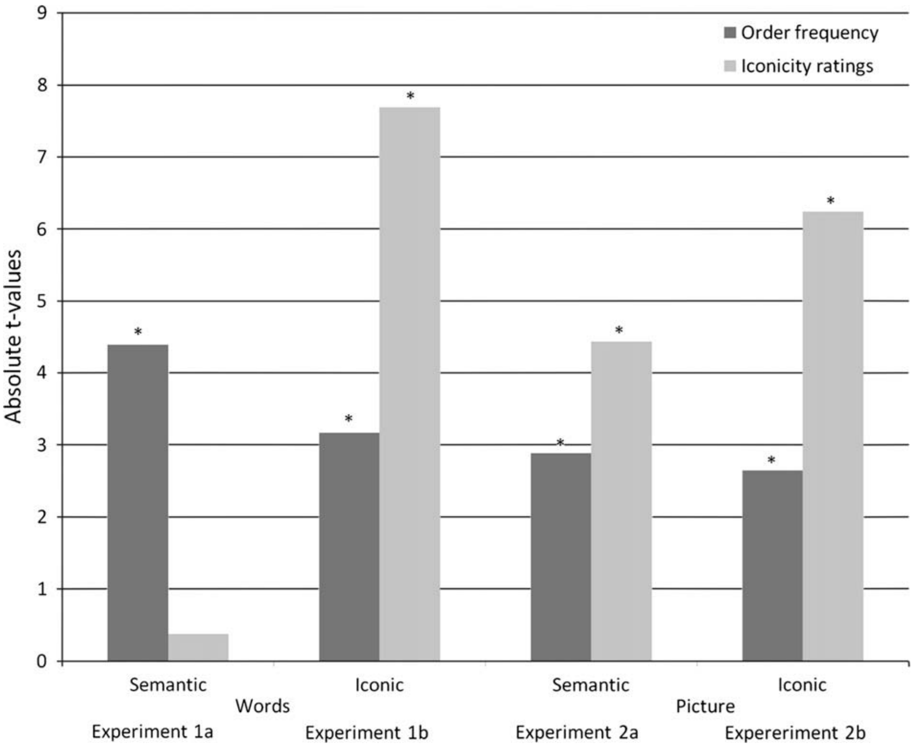

Fig. 2. Strength of logistic regressions to the error rates in absolute *t*-values for each of the four experiments for linguistic (order frequency) and embodiment (iconicity ratings) factors. Asterisks mark significant strengths ($p < .05$) of relationship with error rates. For Experiment 2b *t*-values are given for items with a high semantic association, as discussed in the text.

ambiguous nature of the task, with no clear cut off for semantically related semantically unrelated stimuli. For instance, according to LSA the word pair *rib* and *spinach* has a low semantic relation (cos = .07), and *clarinet* and *violin* a high semantic relation (cos = .74). However, in the semantic field of food, the former two words are very much related, whereas in the semantic field of string instruments, the latter two words are not related (see Louwerse, Cai, Hu, Ventura, & Jeuniaux, 2006).

Analyses of the errors revealed no evidence for a speed-accuracy trade-off. In fact, correct answers were associated to a faster RT than incorrect answers, $t(1699.29) = -6.33$, $p < .001$.

A mixed-effect regression model analysis was conducted on error rates and RTs with order frequencies and iconicity ratings as fixed factors and participants and items as random factors (Baayen, Davidson, & Bates, 2008; Locker, Hoffman, & Bovaird, 2007; Richter, 2006). The problem with standard ANOVAs and regression models is that analyses are conducted by-subjects or by-items, whereas they should be analyzed together in one model. Mixed effects models allow for such an analysis (see Brysbaert (2007) for an excellent primer on mixed effect models). The model was fitted using the restricted maximum likelihood estimation (REML) for the continuous variable (RT). *F*-Test denominator degrees of freedom for both error rates (logistic regression) and RT (regression) were estimated using the Kenward–Roger's degrees of freedom adjustment to reduce the chances of Type I error (Littell, Stroup, & Freund, 2002).

A mixed effects logistic regression showed order frequency significantly predicted error rates, $F(1, 113.5) = 19.25$, $p \leqslant .001$, $R^2 = .15$, with higher frequencies yielding less errors.[2] No such effect was found for iconicity ratings, $F(1, 109.8) = .14$, $p = .71$, $R^2 < .001$ (see the first two bars in Fig. 2). No interactions were found.

For mixed-effect regression models on RT, values below 200 ms and values above 2.5 standard deviations from an individual participant's mean were removed from the analysis (2.6% of the data).

Order frequency significantly predicted RT, $F(1, 105.02) = 18.87$, $p < .001$, $R^2 = .15$, with higher frequencies yielding faster RTs. A similar pattern was observed for iconicity ratings, but this factor did not yield a significant relation with RT, $F(1, 104.18) = .03$, $p = .859$, $R^2 < .01$ (see the first two bars in Fig. 3). The direction of the factors was, however, the same with higher order frequencies yielding lower RTs and higher iconicity ratings tending to yield lower RTs. No interactions were found.

---

[2] Note that the strength of a model association is represented as a weighted ratio of the *F* statistic. R^2 and *F* used in ordinary regression analysis are closely related, since $F = \frac{\frac{R^2}{k}}{\frac{1-R^2}{N-k-1}}$ where *k* is the number of model parameters and *N* is the number of cases, such that *F* has $(k, N-k-1)df$. See also Pedhazur (1997, p. 105).

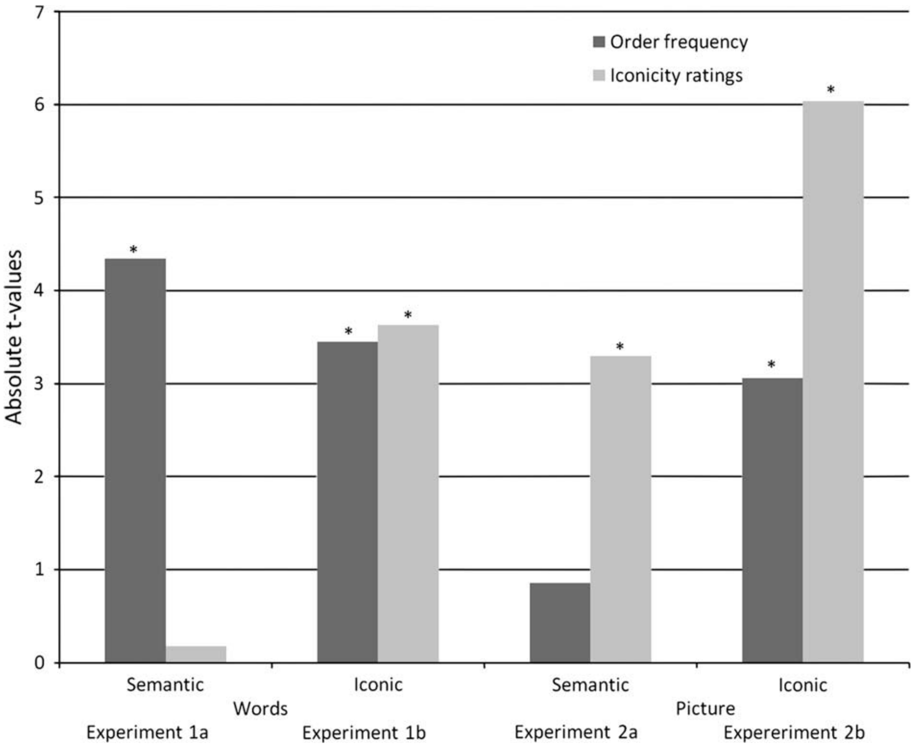

Fig. 3. Strength of regressions to the RTs in absolute *t*-values for each of the four experiments for linguistic (order frequency) and embodiment (iconicity ratings) factors. Asterisks mark significant strengths ($p < .05$) of relationship with RT.

In one sense these results are in line with Louwerse (2008) in that in a semantic judgment task with verbal stimuli the order frequency explained response times better than the iconicity rating ($z = 4.34$, $p < .001$). However, in another sense the results are surprising because iconicity did not significantly explain response times although such an effect was found in Zwaan and Yaxley (2003). We do not have an explanation for this discrepancy. Perhaps most important to note is that as far as one could speak about a relation between iconicity ratings and RT, this relation was in the predicted direction, with higher iconicity ratings yielding lower RTs.

But perhaps the absence of an iconicity effect is not surprising, given the fact that the task concerns evaluating the semantic similarity of linguistic stimuli. Accordingly, making the task more embodied is expected to increase the likelihood of activating embodied representations. This was investigated in Experiment 1b.

3. Experiment 1b

In Experiment 1b participants were asked to make iconicity judgments instead of semantic similarity judgments.

3.1. Method

3.1.1. Participants

Thirty undergraduate students from the University of Memphis participated for Psychology course credit.

3.1.2. Materials

The same 64 experimental items and 128 fillers were used as those in Experiment 1a.

3.1.3. Procedure

The procedure was the same as in Experiment 1a, except that participants were asked whether the vertical configuration of the words matched their referents' relative position as it would usually be found in the world (i.e., had an iconic order) (Fig. 1, top row). As before, participants answered "yes" or "no".

3.2. Results and discussion

The error rate was approximately 25% ($M = 25.77$, $SD = 4.3$). Error performance can again be explained by the task. For instance, word pairs such as *priest* and *flag* are not iconic, but an iconic relation could be imagined. No clear evidence was found for a speed-accuracy trade-off, $t(1757.85) = 1.95$, $p = .06$.

Both order frequency and iconicity ratings significantly predicted error rates, $F(1, 91.4) = 10.02$, $p < .001$, $R^2 = .1$ and $F(1, 108.4) = 59.26$, $p < .001$, $R^2 = .35$ (see Fig. 2), with higher frequencies and ratings yielding fewer errors. No interactions were found.

As before, for analyses on RT, values below 200 ms and values above 2.5 standard deviations from an individual participant's mean were removed from the analysis (2.1% of the data).

As in Experiment 1a, all participants were included in the analysis, with no outliers for accuracy. Different than Experiment 1a both order frequency and iconicity ratings significantly explained RT, $F(1, 94.27) = 11.88$, $p = .001$, $R^2 = .11$, and $F(1, 93.38) = 13.18$, $p < .001$, $R^2 = .12$, respectively (see Fig. 3). As in Experiment 1a, higher frequency and higher ratings both yielded lower RTs. No interactions were found.

The results of the semantic judgment experiment (1a) and the iconicity judgment experiment (1b) suggest that the strength of the relationship between order frequency and RT is larger for the semantic judgment task than the iconicity judgment task. The iconicity rating only presented a significant relationship with RT in the iconicity judgment task.

The order frequency seems to explain conceptual processing of verbal stimuli the best. The fact that the iconicity rating had no significant impact on RT in the semantic judgment task of Experiment 1a is perhaps not surprising given that the verbal stimuli encourage linguistic processing. It is therefore noteworthy that the iconicity rating does significantly explain RT of verbal stimuli, however, presumably only if the task encourages it, like in Experiment 1b. However, we need to be cautious comparing strength of relationship across experiments, because task difficulty cannot be assumed to be constant.

The question next is whether the word order frequency also has an impact on the evaluation of non-linguistic stimuli, while we expect that the impact of the iconicity rating to emerge with non-linguistic stimuli. This was investigated in a second series of experiments whereby we used pictures instead of words.

4. Experiment 2a

Experiment 2a was identical to Experiment 1a, except that pictures were used instead of words.

4.1. Method

4.1.1. Participants

Thirty undergraduate students participated for course credit.

4.1.2. Materials

For all 384 words from Experiments 1a and 1b, corresponding pictures were obtained from the Internet. Because it turned out to be difficult to find appropriate pictures for all words, a subset of the experimental and filler items was formed, resulting in 56 experimental and only 56 filler items. Pictures size was standardized to one common size.

At the end of Experiments 2a and 2b, both using pictorial stimuli, we asked participants to rate how well a picture represented a word on a scale of 1–6. Interrater reliability ($N = 60$) was excellent (Cronbach's $\alpha = .96$), with participants agreeing that pictures represented words accurately ($M = 5.35$, $SD = 1.26$).

4.1.3. Procedure

Participants were instructed to make semantic judgments on picture pairs (Fig. 1, bottom row).

4.2. Results and discussion

Error rates were higher than the ones found in Experiment 1a ($M = 38.2$, $SD = 5$). It is noteworthy that after a median split on the LSA cosine values, results showed picture pairs with a low semantic relation, error rates were higher (47.5%) than for word pairs with a high semantic relation (27%). Apparently participants assessed a semantic relation even when none was supposed to be found. A potential explanation is that participants could have used visual features from the pictures themselves (e.g., color, shape) than semantic information to make their decisions. No speed-accuracy trade-off was found, $t(1511.30) = 1.09$, $p = .27$. As in the previous experiments, no participant outliers were observed and all participants were included in the analysis.

Neither order frequency nor iconicity ratings predicted error rates (Fs < 1). This can likely be attributed to the increased error rates. That is, if the above explanation for the high error rates is correct, order frequency and/or iconicity ratings might predict error rates for items with a high LSA cosine value. This was indeed the case. Both order frequency and iconicity ratings predicted error rates, $F(1, 35.43) = 8.3$, $p = .01$, $R^2 = .19$ and $F(1, 691) = 19.65$, $p < .001$, $R^2 = .03$ (see Fig. 2), whereby higher frequencies and ratings yielded lower error rates. No interactions were found.

For the RT analyses, values below 200 ms and values above 2.5 standard deviations from an individual participant's mean were removed from the analysis (2.6% of the data).

Iconicity ratings significantly predicted RT, $F(1, 1159.41) = 11.34$, $p < .001$, $R^2 = .01$, but order frequency did not, $F(1, 123.46) = .83$, $p = .363$, $R^2 = .01$ (see Fig. 3). Higher iconicity ratings yielded lower RT, and higher order frequency tended to yield lower RT. The iconicity rating explained RT in the semantic judgment task for pictorial stimuli better than the order frequency ($z = 3.24$, $p < .001$). No interactions were found.

The findings of Experiment 2a are the opposites of those of Experiment 1a. In Experiment 1a the order frequency explained RT for verbal stimuli, but the iconicity rating did not. In Experiment 2a, the iconicity rating explained RT for pictorial stimuli but the order frequency did not. One could interpret these findings by stating that information about semantic similarity is more linguistic in nature when it is accessed through words, but is more embodied when it is accessed through pictures. Whether the impact of iconicity rating on RT also increases for iconicity judgments on pictures was assessed in Experiment 2b.

5. Experiment 2b

Experiment 2b was identical to the iconicity judgment of Experiment 1b except that instead of word pairs Experiment 2b used picture pairs.

Table 1
Average response times (and *SD*) for Experiments 1a, 1b, 2a, and 2b.

	Word pair	Picture pair	
Semantic judgment	1493.27 (584.76)	1551.14 (742.47)	1530.70 (691.38)
Iconicity judgment	2085.79 (1015.35)	2518.97 (1224.04)	2360.55 (1170.76)

5.1. Method

5.1.1. Participants

Thirty undergraduate students participated for Psychology course credit.

5.1.2. Materials

The same pictures were used as in Experiment 2a.

5.1.3. Procedure

The procedure was the same as in Experiments 1b and 2a, except that participants were instructed to judge whether or not the pairs of pictures represented an iconic order (Fig. 1, bottom row).

5.2. Results and discussion

Error rates were similar to those of Experiment 1b ($M = 26.01$, $SD = 4.3$), and as before, no speed-accuracy trade-off was found, $t(1623.99) = -0.77$, $p = .44$.

Because there were no participant outliers, all participants were included in the analysis. The same outlier criteria were used as before, affecting 2.4% of the data.

Both order frequency and iconicity ratings predicted error rates, $F(1, 73.94) = 6.97$, $p = .01$, $R^2 = .09$, and $F(1, 1232) = 38.98$, $p < .001$, $R^2 = .03$ (see Fig. 2), and as before, higher frequencies and ratings yielded fewer errors. No interactions were found. Both iconicity ratings and order frequency significantly predicted RT, $F(1, 1151.07) = 36.44$, $p < .001$, $R^2 = .03$, and $F(1, 139.61) = 9.36$, $p = .002$, $R^2 = .06$, respectively (see Fig. 3). When iconicity ratings or order frequencies were higher, RT was faster. The strength of relationship with RT was higher for the iconicity rating than the order frequency ($z = 5.2$, $p < .01$). No interactions were found.

When the RTs for the four experiments were compared, RTs for the pictorial stimuli were higher than those for the verbal stimuli, $F(1, 389.29) = 124.85$, $p < .001$, $R^2 = .24$, and RTs for the iconicity task are higher than those for the semantic judgment task, $F(1, 3751.18) = 470.19$, $p < .001$, $R^2 = .11$. Moreover, an interaction between stimuli and task was found, $F(1, 3651.64) = 5.93$, $p = .02$, $R^2 < .01$ (Table 1). This might suggest that the overall error rates and RT for the four experiments changes as a function of the nature of the stimulus and the nature of the task, with the linguistic factor being strongest on the semantic judgment task and verbal stimuli, and the embodiment factor being strongest on the iconicity judgment task and pictorial stimuli (Figs. 2 and 3). However, because of differences in task difficulty and number of items, these findings should be treated with caution.[3]

6. General discussion

A wealth of empirical evidence has shown that cognition is embodied. The current paper investigated the conditions under which embodiment and linguistic factors affect conceptual processing. The linguistic factor, operationalized by frequency of word order, explained RT better for word pairs than picture pairs. The embodiment factor, operationalized by iconicity ratings, explained RT better for picture pairs than word pairs. These findings were modified by task (semantic judgment or iconicity judgment), with the embodiment factor being strongest in iconicity judgments for pictures and the linguistic factor being strongest in semantic judgments for words. Identical findings were obtained for error rates. These RT and error rate results suggest that neither linguistic nor embodiment factors reign supreme in any absolute sense, but depend on the nature of the task and stimuli.

How could the results be interpreted? One could attempt to answer this question in terms of depth of processing, whereby depth corresponds to the extent to which meaningfulness is extracted from the stimulus (Lockhart & Craik, 1990). At the task level, iconicity judgments could be considered as requiring a deeper level of processing than semantic judgments, if one assumes that iconicity judgments imply semantic judgments but not vice versa. Moreover, at the stimulus level, some studies have shown processing of words is easier than processing of pictures (Fraisse, 1968; Paivio, 1978). Although it is debatable whether words require shallower processing than pictures (cf. Banks & Flora, 1977), the patterns observed in the total RTs by task, discussed earlier, seem to warrant such a conclusion, with longer RTs for pictures on one hand, and longer RTs for iconicity judgments on the other hand. Based on these distinctions in terms of depth, and the results visible in Fig. 2, it can then be observed that the linguistic factor dominates in shallow cognitive tasks, whereas the embodiment factor dominates in deeper cognitive tasks (Louwerse & Jeuniaux, 2008).

A second answer as to how our findings can be interpreted, and one related to the previous one, comes from Paivio's (1986) Dual Coding Theory. Paivio identified three levels of meaning, a representational, referential, and associative level. At the level of representational meaning, verbal and non-verbal stimuli activate the corresponding representational comprehension processes. At the second level of meaning, referential meaning, interconnections are formed between the verbal and non-verbal representational processes. At the third level, associative meaning involves intraverbal associations (associative connections

---

[3] Even though task difficulty was not controlled, number of items did not affect the results. When an analysis was conducted on only those items that were identical in Experiments 1a, 1b, 2a, and 2b (see Appendix), similar results were obtained as those reported elsewhere in this paper (Experiment 1a: order frequency: $t = -3.48^*$, iconicity ratings: $t = -.47$; Experiment 1b: order frequency: $t = -3.18^*$, iconicity ratings: $t = -3.42^*$; Experiment 2a: order frequency: $t = -.854$, iconicity ratings: $t = -3.29^*$; Experiment 2b: order frequency: $t = -3.06^*$, iconicity ratings: $t = -6.04^*$ ($^*p < .01$).

between words) and interimaginal representations (associative connections between pictures). The Dual Coding Theory predicts processing of representations that match the corresponding stimuli being easiest, and processing across representations being hardest. Consequently, according to the Dual Coding Theory linguistic and embodiment factors each dominate when they match the nature of the task and stimuli.

As a final remark pertaining to the interpretation of results, we acknowledge that caution is advised in considering a direct cause–effect relationship between the independent and dependent variables, because the linguistic and embodiment factors might correlate with other unmeasured variables that caused the differences in RT instead. However, we are unable to speculate about the nature of other variables that have a systematic relationship with word order than the ones used here.

We started out this paper with the emphasis recent cognitive studies have placed on the fundamental role of perceptual simulation in cognitive processing, and the wealth of evidence showing that embodied representations are activated in cognitive processing tasks. This is very important research, but with the available evidence in favor of embodied cognition, the question that becomes increasingly relevant is whether other types of representations are formed, and to what extent, and under what conditions they are formed. The current study has shown that the embodiment factor affects RT in iconicity tasks with pictorial stimuli, as well as in semantic judgment tasks with linguistic stimuli. We have also shown that a linguistic factor plays a role in semantic judgment tasks with linguistic stimuli, as well as in iconicity tasks with pictorial stimuli. In that sense we can conclude conceptual processing is both linguistic and embodied. However, differences in strengths of relationship were found across the four experiments, showing that embodiment and linguistic factors affect conceptual processing differently depending on task and stimulus. We recommend that future research on conceptual processing takes into account the nature of the task, stimuli, and different mental representations.

Acknowledgements

We would like to thank George Relyea and Gilbert Parra for their useful feedback on the statistical methodology. The usual exculpations apply.

Appendix. Experimental items used in Experiment 1a–b and 2a–b

Airplane–runway*	Curtain–stage	Hiker–trail*
Antenna–radio*	Eyes–whiskers	Hood–engine*
Antler–deer	Faucet–drain*	Icing–doughnut
Attic–basement*	Fender–tire*	Jam–toast
Belt–shoe*	Flame–candle*	Jockey–horse*
Billboard–highway*	Flower–stem*	Kite–string*
Boat–lake*	Foam–beer	Knee–ankle*

Appendix (*continued*)

Boot–heel	Fountain–pool*	Lamp–table*
Bouquet–vase*	Froth–coffee	Lid–cup*
Branch–root*	Glass–coaster*	Lighthouse–beach*
Bridge–river*	Grill–charcoal*	Mailbox–post*
Car–road*	Handle–bucket*	Mane–hoof*
Castle–moat	Hat–scarf*	Mantle–fireplace*
Ceiling–floor*	Head–foot*	Mast–deck
Cork–bottle*	Headlight–bumper*	Monitor–keyboard*
Mustache–beard*	Roof–porch	Sprinkler–lawn*
Nose–mouth*	Runner–track*	Steeple–church*
Pan–stove*	Saddle–stirrup*	Sweater–pants*
Pedestrian–sidewalk*	Seat–pedal*	Tractor–field*
Penthouse–lobby*	Sheet–mattress*	Train–railroad*
Pitcher–mound*	Sky–ground*	
Plant–pot*	Smoke–chimney*	

Note: Items marked with an asterisk were used as picture pairs in Experiments 2a and 2b. Three picture pairs did not appear as word pairs, but were selected because they were part of Zwaan and Yaxley's (2003) stimulus pairs and received the highest ratings in the picture rating task. These were *rocket–launchpad*, *cup–saucer*, and *stoplight–street*.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Banks, W. P., & Flora, J. (1977). Semantic and perceptual processes in symbolic comparisons. *Journal of Experimental Psychology, 3*, 278–290.

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

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. de Vega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 245–283). Oxford: Oxford University Press.

Benor, S. B., & Levy, R. (2006). The chicken or the egg? A probabilistic analysis of English binomials. *Language, 82*, 233–278.

Brants, T., & Franz, A. (2006). *Web 1T 5-gram version 1*. Philadelphia: Linguistic Data Consortium.

Brysbaert, M. (2007). *"The language-as-fixed-effect fallacy": Some simple SPSS solutions to a complex problem (version 2.0)*. Royal Holloway, University of London. Technical report.

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: Object words orient attention to the object's typical location. *Psychological Science, 19*, 93–97.

Fraisse, P. (1968). Motor and verbal reaction times to words and drawings. *Psychonomic Science, 12*, 235–236.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20*, 1–55.

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). (2007). *Handbook of latent semantic analysis*. Mahwah, NJ: Erlbaum.

Littell, R. C., Stroup, W. W., & Freund, R. J. (2002). *SAS for linear models* (4th ed.). Cary, NC: SAS Publishing.

Locker, L., Jr., Hoffman, L., & Bovaird, J. A. (2007). The use of multilevel modeling in the analysis of psycholinguistic data. *Behavior Research Methods, 39*, 723–730.

Lockhart, R. S., & Craik, F. I. M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology, 44*, 87–112.

Louwerse, M. M. (2008). Embodied representations are encoded in language. *Psychonomic Bulletin and Review, 15*, 838–844.

Louwerse, M. M. (in press). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science (TopiCS)*.

Louwerse, M. M., Cai, Z., Hu, X., Ventura, M., & Jeuniaux, P. (2006). Cognitively inspired natural-language based knowledge representations: Further explorations of latent semantic analysis. *International Journal of Artificial Intelligence Tools, 15*, 1021–1039.

Louwerse, M. M., & Jeuniaux, P. (2008). Language comprehension is both embodied and symbolic. In M. de Vega, A. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 309–326). Oxford: Oxford University Press.

Meteyard, L., Baharami, B., & Vigliocco, G. (2007). The perceptual grounding of motion verbs. *Psychological Science, 11*, 1007–1013.

Paivio, A. (1978). Imagery, language, and semantic memory. *International Journal of Psycholinguistics, 5*, 31–47.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.

Pecher, D., & Zwaan, R. A. (Eds.). (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. New York: Cambridge University Press.

Pedhazur, E. (1997). *Multiple regression in behavioral research*. New York: Holt, Rinehart and Winston.

Richardson, D., & Spivey, M. (2000). Representation, space, and Hollywood squares: Looking at things that aren't there anymore. *Cognition, 76*, 269–295.

Richter, T. (2006). What is wrong with ANOVA and multiple regression? Analyzing sentence reading times with hierarchical linear models. *Discourse Processes, 41*, 221–250.

Semin, G. R., & Smith, E. R. (Eds.). (2008). *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches*. New York: Cambridge University Press.

Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes, 22*, 297–312.

Zwaan, R. A. (2004). The immersed experience: Toward an embodied theory of language comprehension. In B. H. Ross (Ed.). *The psychology of language and motivation* (Vol. 44). New York: Academic Press.

Zwaan, R. A. (2008). Experiential traces and mental simulations in language comprehension. In M. DeVega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 165–180). Oxford: Oxford University Press.

Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic-relatedness judgments. *Psychonomic Bulletin and Review, 10*, 954–958.