Perceptual Similarity in Visual Metaphor Processing

Lisanne van Weelden, Reinier Cozijn, Alfons Maes, Joost Schilperoord

Communication and Cognition, School of Humanities, Tilburg University
Warandelaan 2, 5037 AB Tilburg, The Netherlands
L.vanWeelden@uvt.nl, R.Cozijn@uvt.nl, Maes@uvt.nl, J.Schilperoord@uvt.nl

Abstract

In visual metaphor processing, one object, the *target*, is compared to and understood in terms of another object, the *source*. Several studies suggest that perceptual similarity between two objects enhances a conceptual link between the two. However, little is known about *how* perceptual features contribute to the establishment of this link. In the present experiment we investigated the processing of the four possible combinations of conceptually and perceptually similar picture pairs using a same-different task. In order to determine whether particular processes are bound to a particular time range, we manipulated the delay between the two successively presented pictures. We expected perceptual processing effects at a short delay and conceptual processing effects at a longer delay. We did not find evidence for this expectation. However, the results did show that (i) it took participants longer to give a 'different' response if two objects shared perceptual features than when they did not; (ii) this presence of perceptual similarity also resulted in more response errors; and (iii) if objects shared only perceptual features, participants in the long delay condition produced more erroneous responses than the participants in the short delay condition did. These results are discussed in light of metaphor processing models.

Introduction

Theories of metaphor processing claim to answer the question how people make sense of ordinary metaphors, such as 'a cigarette is (like) a bullet' (Bowdle & Gentner 2005; Lakoff & Johnson 1980; Steen 2007). The main problem to be solved seems to be how to specify the cognitive process that leads people to understand the target concept 'cigarette' in terms of the source concept 'bullet'. Although opinions differ as to how this should be done, existing theories converge on the idea that metaphor processing initially involves the cognitive process of *comparing* target and source, hence of juxtaposing two conceptual domains in order to find out what justifies them being aligned. Understanding a cigarette in terms of a bullet is therefore a process of mapping source features to the target, for example, the deadly potential of bullets to cigarettes, or to include cigarettes into the category of 'deadly things'.

Consider the advertisement from the NHA (New Zealand Health Association) about the dangers of smoking shown in Figure 1. In the picture, a large number of lined up bullets is depicted. Within the row of bullets one bullet is replaced by a cigarette. Conceptually, this image expresses the very same metaphor of the verbal example given earlier, [CIGARETTE IS BULLET]. So, to interpret this image the viewer should compare the concepts bullet and cigarette. Notice, however, that in this image the metaphor is not expressed 'just' by showing the objects, but also by making them *look similar*. The conceptual link is furthermore suggested by artful visual manipulations, such as perspective, spatial orientation, and by projecting the objects on a hypothetical horizontal axis (Schilperoord, Maes, & Ferdinandusse 2009). Although such types of visual figuration have been described extensively (Schilperoord & Maes 2009; Schilperoord et al. 2009; Teng & Sun 2002), little is known about how such perceptual features may influence the cognitive process of mapping in processing visual metaphors.

Therefore, the present study aims at exploring the role of perceptual similarity between objects in visual metaphor processing. According to Tversky's (1977) contrast model,

Figure 1: NHA advertisement

similarity is a feature-matching process. That is, objects are represented by features (perceptual or otherwise), and the number of commonalities between the features defines the amount of similarity. As a result, the similarity of an object pair increases with the number of common features and decreases due to its distinctive features.

If the goal is to explore the way that perceptual similarity enhances the processing of metaphors, the literature on visual object processing suggests that we should look especially at the very early stages of processing. That is, research on visual processing has shown that perceptual features are fundamental in processes prior to (Langley, Cleary, Kostic, & Woods 2008; Snodgrass, Hirshman, & Fan 1996) and throughout object identification (Loftus & Ginn 1984). The study of Langley et al. (2008) tried to examine the claim of Snodgrass et al. (1996) that even an unidentifiable picture can only be recognized by its perceptual characteristics. In this so-called pre-identification phase, a picture is seen and processed, however not yet identified nor understood. Langley et al.'s (2008) experiments showed that, for pictures (black and white line drawings) that were identified, participants were able to discriminate between old (seen in a previous study session) and new pictures. However, the same result was found for pictures which were impossible to identify due to masking by straight and curved black lines. The latter finding suggests the existence of recognition without identification. So, the participants saw the pictures and thereby collected enough information to recognize them in a subsequent task. The question whether this type of information is based on perceptual or conceptual features was addressed in another experiment. During the study session of this experiment, picture names were used instead of the pictures themselves. When picture names were used, participants were no longer able to discriminate between old and new unidentified pictures. Since the picture names corresponded to the conceptual features of pictures, it was concluded that an unidentified picture cannot be recognized by its conceptual characteristics. These results were taken as support for the claim of Snodgrass et al. (1996) that recognition of unidentifiable pictures is based on perceptual characteristics. Flores d'Arcais and Schreuder (1987), furthermore, demonstrated that perceptual features are more compelling than conceptual features in priming experiments. They conducted an object naming task in order to explore semantic activation. The pictures (black and white line drawings) that had to be named were primed by another picture, which shared either perceptual, conceptual, or both features with the targets. The authors found the largest facilitation effect when prime and target shared both perceptual and conceptual features. However, when the two pictures shared only perceptual features, the priming effect was stronger than when the pictures were only conceptually similar.

Yet another study that investigated picture perception is that of Loftus and Ginn (1984). They suggest that picture perception involves two different processes: perceptual and conceptual processes. They propose that perceptual processes result in the identification of the picture and that conceptual processes operate on the output of the perceptual processes, but do not require the presence of the visual input itself. Loftus and Ginn distinguished conceptual from perceptual masking in a picture-memory paradigm in which one target picture of a naturalistic scene (for example, street scenes and landscapes) was presented. The subject's task was to recall as many details as possible from the picture. The effect of different masks (photograph mask or noise mask, bright or dim) on picture-memory performance was studied for two different delay conditions. Masks were presented either immediately following the offset of the picture or were delayed by 300 ms following the offset of the picture. The results showed that the different delay conditions led to different psychological processes. For the immediate mask, the luminance of the mask affected picture-memory, while for the delayed mask, type of mask affected picture-memory. Loftus and Ginn concluded that a mask immediately following the offset of a 50 ms picture interrupts perceptual processing and a mask from 300 ms interrupts conceptual processing.

All in all, these studies suggest that perceptual processes occur at a very early stage and conceptual processes at a later stage of picture processing. The present experiment investigates this issue in relation to the processing of picture pairs that are conceptually and perceptually similar, conceptually similar, perceptually similar, conceptually and perceptually dissimilar, or identical. The difference in picture pairs will be a within subjects factor. By employing the simple task of picture matching with either a short or a long delay, the experiment hopes to tease apart the distinct roles of perceptual and conceptual similarities in time during the matching process, which is assumed to be part of processing visual metaphors. The short or long delay between the sequentially presented pictures will be a between subjects factor. In order to capture the very early stages of processing in this experiment, the first picture of the match will be presented for 15 ms, a very short duration. In doing so, the picture is processed on a unconscious level which, according to Marcel (1983), is typified by automatic processes, instead of processes including selective intentionality.

Our experimental task is designed to measure the perceptual processes that take place when, instead of pictures in isolation, picture *pairs* are being processed. By presenting the two pictures successively, rather than simultaneously, we hope to be able to distinguish perceptual and conceptual processes. The experimental materials consist of simple black and white line drawings depicting utensils. Using simple drawings avoids any effects of high processing loads due to image characteristics like background, color, and texture of the objects. The five different picture pairs represent a match picture and the target picture (see Figures 2a-e). One of the match pictures is conceptually and perceptually similar to the target (pen - pencil, 2a), one is only conceptually similar (goose feather - pencil, 2b), one is only perceptually similar (needle - pencil, 2c), one is not similar at all (balloon - pencil, 2d), and one is completely identical to the target (pencil - pencil, 2e). So, in each set, there are two pairs depicting objects that belong to the same conceptual domain (pen - pencil, goose feather - pencil), and two pairs that depict objects originating from different conceptual domains (needle - pencil, balloon - pencil). Perceptual similarity as a factor is manipulated independently from this distinction. Hence, we find in each type of pairs one pair of objects that are perceptually similar. The effects of perceptual features during metaphor processing are expected to occur when participants process the pair of objects that look similar, while at the same time originate from distinct conceptual domains.

More specifically, we expect a difference in processing the picture pairs as follows. There will be a difference between the perceptually similar picture pair (needle - pencil) and the conceptually and perceptually similar picture pair (pen - pencil). With regard to this comparison we particularly expect a difference for the long delay group. At a short delay, when only perceptual features are processed, the matching task will not be expected to show a difference between the two pairs, since both pairs are perceptually similar. Their perceptual similarity will make it equally hard for the participants to judge them as being different. However, at a long delay, when conceptual processing is assumed to take place, the pair that is conceptually similar as well will be processed on its conceptual features. This comparison will lead to an erroneous match since the pictures are not the same. This erroneous match is expected to slow down the 'different' response in the same-different matching task. This additional conceptual similarity results in greater demands on the working memory (Kroger, Holyoak, & Hummel 2004) and, therefore, it should take longer to process the conceptual and perceptual similar object pairs (pen - pencil) than the perceptual similar object pairs (needle - pencil). A second difference is expected to show up in the comparison between the perceptually similar picture pairs (needle - pencil) and the conceptually similar picture pairs (goose feather - pencil). These pairs are expected to be differentially judged at both delays. The perceptual similarity will cause a slower response at an early delay, and the conceptual similarity a slower response at the late delay.

Method

The purpose of the present experiment is to gain insight in the processing of picture pairs that belong to the same conceptual domain or originate from different conceptual domains and, additionally, do or do not show perceptual similarity. An experimental paradigm with a same-different task was developed. In this task the participants had to match the two pictures and decide whether they were the same or different. The reaction times, as well as the actual 'yes' (same) and 'no' (different) responses, were expected to provide information about the way that the different picture pairs are processed.

Participants

Sixty participants (forty-three Tilburg University undergraduates and seventeen Tilburg University employees, 38 females and 22 males) received 5 euros for their participation in the experiment. The mean age was 25 years, ranging from 18 to 37. All participants were naive with respect to the purpose of the experiment and had normal or corrected-to-normal vision.

Apparatus

Stimuli were displayed in black on a white background, on a 19-inch color screen. Stimulus presentation and recordings of the behavioural data, that is reaction times and accuracy, were controlled by E-prime (www.pstnet.com; Psychology Software Tools, Inc.). E-prime also controlled the resolution of the screen, which was set to 640x480 pixels.

Materials

The pictures were black and white line-drawings depicting utensils. The line-drawings were very simple illustrations and each drawing contained an area of 200x200 pixels. For the experiment, 20 sets of five pictures pairs were created, that is, a Conceptual-Perceptual picture pair, a Conceptual picture pair, a Perceptual picture pair, a Different picture pair, and a Same picture pair (again, see Figure 2a-e). Considering the same-different task, only the Same picture pair should lead to a correct 'yes' (same) response (20 in total), whereas all other picture pairs should result in a correct 'no' response (80 in total). In order to counterbalance the 'yes' and 'no' responses in the experiment, 60 fillers were included that required a 'yes' response.

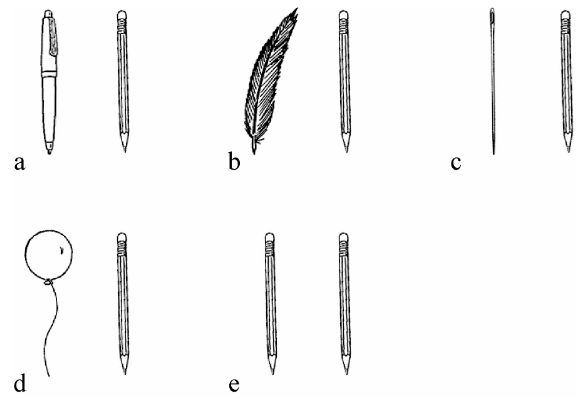

Figure 2a-e: Picture pairs indicating the different types of similarity between the match picture and the target picture, (a) Conceptual/Perceptual, (b) Conceptual, (c) Perceptual, (d) Different, and (e) Same.

In order to make sure all pictures depicted what we wanted them to depict, they were pre-tested with an object naming task[1]. Furthermore, all picture pairs were randomly divided into two blocks, each with the same amount of picture pairs in each condition.

Procedure

Participants were told that they were going to take part in a reaction time experiment involving the comparison of pictures and that their task was to judge as fast as possible whether the two presented pictures were the same or different. Each trial consisted of the following sequence (see Figure 3). First, the participants had to press a response key to start a stimulus trial. When the key was pressed, a fixation cross appeared for 2 seconds[2]. Subsequently, a 50 ms visual noise mask appeared, which consisted of random black and white pixels within a 300x300 square. Immediately after this mask, the match picture was presented for a duration of 15 ms. After the match picture, a 50 ms visual noise interruption mask was presented. This kind of interruption masking arises when processing of a first pattern (the match picture) is interrupted by a second pattern (the mask) that appears in the same spatial location before the first pattern has been fully processed (Enns & Di Lollo 2000). The task of this mask was to interrupt both perceptual and conceptual processes and to make sure that no iconic information of the preceding match picture was present once the target was presented. The mask was followed by a delay – a blank screen – of 250 ms or 950 ms, which led to a *stimulus onset asynchrony* (SOA) of 300 ms or 1000 ms. This difference in SOA was based on Loftus and Ginn's finding (1984) that perceptual processes last till 300 ms after the offset of the picture and that from 300 ms conceptual processes are operating. After the delay, the target was presented and participants had to judge as fast as possible whether the target picture was the same as or different from the first picture. They gave their judgment by pressing a key on a button panel. The 'yes' response was always located at the dominant hand side of the participant. Immediately after their judgment, feedback indicated whether the answer was correct, incorrect, or too late, that is, later than 2 seconds. Right after the feedback, the next trial started with a cross on the screen. The procedure was practiced in a session of 10 trials. The difference in delay was a between participants factor and participants were randomly assigned to one of the two conditions. The experiment lasted approximately 25 minutes.

[1] Forty-nine Tilburg University undergraduates participated in this pre-test. None of these participants took part in the picture matching experiment.

[2] All presentation durations were synchronized with the refresh rate of the screen (13.6 ms).

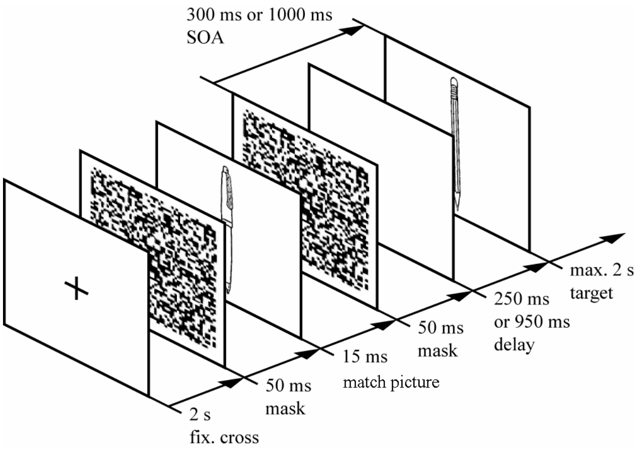

Figure 3: Stimulus trial; order and duration of the fixation cross, mask(s), match picture, delay and target picture.

Results

The analyses focused on reaction times (RTs) and on accuracy (correct and incorrect responses). First, all RTs were screened for outliers. RTs below 120 ms and RTs above 1075 ms were eliminated. This resulted in a loss of 5.1% of the data.

RT analysis

The RT analysis was performed on the correct responses (85.7 %). As depicted in Figure 4, there was a difference in RT between the different picture pairs. That is, for the picture pairs with perceptual similarity (left and middle), the RTs were slower than for the three other pairs. A Repeated Measures analysis, with Type of picture pair (within participants) and Delay (between participants) as factors confirmed this pattern. There was a main effect of Type: $F1(4, 232) = 99.25$, $MSE = 1379.70$, $p < 0.001$ and $F2(4, 95) = 20.21$, $MSE = 2309.14$, $p < 0.001$, indicating slower responses for the Conceptual/Perceptual and Perceptual picture pairs than the Conceptual, Different and Same[3] picture pairs. However, the expected difference between the Perceptual picture pair and the Conceptual/Perceptual picture pair was not found. The analysis showed no effect of Type ($F1 < 1$). The analysis did reveal a difference in the reaction times on the Perceptual picture pairs and the Conceptual picture pairs. It took longer to say 'no' to a match of a Perceptual picture pair than to that of a Conceptual picture pair: $F1(1, 58) = 129.65$, $MSE = 1197.80$, $p < 0.001$, /d/ = 71.95; 95% CI [59.30, 84.60]. Finally, the RTs of the Same picture pairs differed significantly from all four other picture pairs: $F1(1, 58) = 17.96$, $MSE = 1514.79$, $p < 0.001$, /d/ = -30.11; 95% CI [-15.89, -44.34]. This was not very surprising, since the Same

[3] Please note that the correct response for the Same picture pair was 'yes' (same).

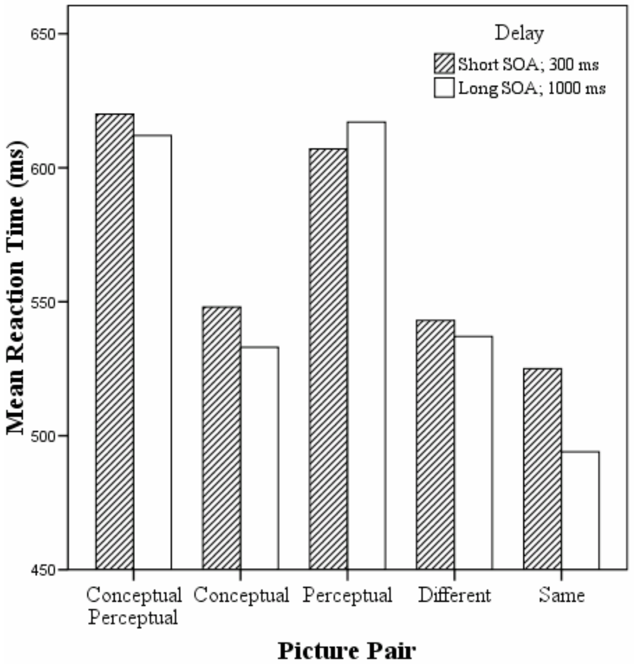

Figure 4: Mean RTs of the correct responses for the five different picture pairs as a function of Delay (SOA).

picture pairs required a ‘yes’ answer and any other pair a 'no' answer, and affirmative answers are easier to produce than negative answers.

As depicted in Figure 4, the RT patterns of both the short SOA and the long SOA group were similar. The analyses showed no effects of Delay ($F1 < 1$), nor any interactions between Type and Delay.

Accuracy analysis

The overall mean error rate was 14.3%. For the short SOA group an error rate of 13.3% was found and for the long SOA group a rate of 15.3% errors. However, as in the RT analyses, no main effect of Delay was found. The analysis revealed a trend though: $F1(1, 58) = 3.13$, $MSE = 19.00$, $p = 0.08$, /d/ = -1.99; 95% CI [-4.24, 0.26] There was a trend of more errors in the long SOA group than in the short SOA group.

Looking at Figure 5, it can be seen that the error rates for the Conceptual and the Different picture pairs were lower than the error rates of the Conceptual/Perceptual, Perceptual, and Same picture pairs. So, the error rates were lower when the picture pairs shared no perceptual similarity. The analyses confirmed this difference by yielding a significant main effect of Type: $F1(4, 232) = 108.38$, $MSE = 73.94$, $p < 0.001$ and $F2(4, 95) = 11.38$, MSE = 0.02, $p < 0.001$. Similar to the RT analysis, there were no differences between the error rates of the Perceptual picture pairs and the Conceptual/Perceptual picture pairs ($F1 < 1$). In line with our expectations, there was a difference in error rates between the Perceptual and the Conceptual picture pairs:

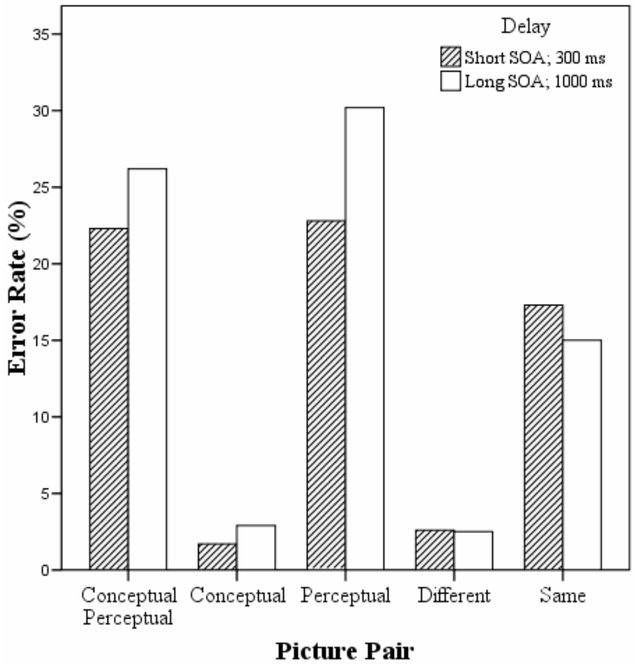

Figure 5: Error rates in percentages for the five different picture pairs as a function of Delay (SOA).

$F1(1, 58) = 258.70$, $MSE = 67.96$, $p < 0.001$, /d/ = 24.21; 95% CI [21.20, 27.22]. There also was a significant interaction between Type and Delay: $F1(4, 232) = 2.89$, $MSE = 73.94$, $p < 0.05$. A closer look at the error rates of both groups revealed a difference between the short SOA and the long SOA group for the Perceptual picture pairs. A *t*-test confirmed this difference between the groups: $t(58) = -2.405$, $p < 0.05$, /d/ = -7.44; 95% CI [-13.63, -1.25]. The long SOA group made more errors than the short SOA group (30.2% versus 22.8%). Furthermore, similar to the RT analysis, the error rates of the Same picture pairs differed significantly from those of the other four types, but since the Same pairs required 'yes' responses and the other pairs 'no' responses, comparisons are hard to make.

Discussion and Conclusion

The goal of the present experiment was to gain insight into the distinct roles of perceptual and conceptual similarities in the process of matching picture pairs, a process which is assumed to be part of processing visual metaphors. The effects of perceptual similarity during picture pair processing were expected when participants processed a picture pair that looked similar, but was conceptually different, that is, originated from distinct conceptual domains. Therefore, we expected differences in processing time between the Perceptual similar picture pairs, like needle - pencil, and the Conceptual/Perceptual similar picture pairs, like pen - pencil, particularly for the long SOA group, and between the Perceptual similar picture pairs, like needle - pencil, and the Conceptual similar picture pairs, like goose feather - pencil. For the first comparison we found no differences, regardless of delay. This might indicate that the conceptual similarity in these cases had been overruled by the perceptual similarity, even for the long SOA group. This explanation is supported by the fact that there were significant differences between the Perceptual similar picture pairs and the Conceptual similar picture pairs, with longer RTs for the Perceptual similar picture pairs. In other words, it took the participants of both groups longer to give a 'different' response when two pictures shared perceptual features than when they shared conceptual features, or no features at all. An explanation would be that, regardless of delay, the participants in the experiment mainly focused on the object-constitutive factor 'shape'. The lack of an effect of delay might be explained by the fact that the participants were specifically instructed to pay attention to perceptual features, and this behaviour was reinforced by the feedback during the experiment. So, any influence of the delay may have been annulled by this perceptual focus. A second explanation would be that an SOA of 1000 ms is too short to detect conceptual processes. Although this explanation is possible, it is not probable, since it is in conflict with studies like Loftus and Ginn (1984). Though taking in consideration that the findings of Loftus and Ginn are based on scene perception, these explanations fail to account for the significant difference in error rates between the short SOA and the long SOA group for the Perceptual similar picture pairs: Participants of the long SOA group gave more erroneous 'same' responses for picture pairs that shared perceptual features than the short SOA group. As proposed in this paper, more processing time leads to additional conceptual processing, besides perceptual processing. So, the increase of the delay may have led to conceptual processes, resulting in the establishment of a conceptual relationship between two pictures. As a result, participants may have given more (erroneous) 'same' responses, which were based on a conceptual link between the two pictures.

In conclusion, perceptual similarity seems to play the leading role in the early stages of picture pair processing. However, in order to explore to what extent perceptual features enhance conceptual processing, subsequent research will focus on an experimental task that more directly instructs participants to pay attention to conceptual features. That is, in a follow-up study the task will be to answer the more conceptual question 'Can you use the two objects for the same purpose?'. If perceptual features enhance conceptual processing, then perceptual similarity should lead to facilitation when the two pictures are conceptually similar. Therefore, we expect lower reaction times for the Conceptual/Perceptual picture pairs than for the Conceptual picture pairs. On the other hand, if the two pictures show no conceptual similarity, then perceptual similarity is expected to inhibit the conceptual processing. This would lead to more erroneous 'yes' responses for the Perceptual picture pairs in comparison to the Different picture pairs.

Acknowledgments

This research is part of the project "Visual metaphor: a Psycholinguistic Perspective", which is funded by an 'Open Competition' grant from the Netherlands Organization for Scientific Research (NWO).

References

Bowdle, B. F., and Gentner, D. 2005. The Career of Metaphor. *Psychological review* 112(1): 193-215.

Enns, J. T., and Di Lollo, V. 2000. What's new in visual masking? *Trends in Cognitive Sciences* 4(9): 345-352.

Flores d'Arcais, G. B., and Schreuder, R. 1987. Semantic activation during object naming. *Psychological Research* 49: 153-159.

Kroger, J. K., Holyoak, K. J., and Hummel, J. E. 2004. Varieties of sameness: the impact of relational complexity on perceptual comparisons. *Cognitive science* 28: 335-358.

Lakoff, G., and Johnson, M. 1980. *Metaphors we live by.* Chicago: University of Chicago Press.

Langley, M. M., Cleary, A. M., Kostic, B. N., and Woods, J. A. 2008. Picture recognition without picture identification: A method for assessing the role of perceptual information in familiarity-based picture recognition. *Acta psychologica* 127(1): 103-113.

Loftus, G. R., and Ginn, M. 1984. Perceptual and Conceptual Masking of Pictures. *Journal of Experimental Psychology: Learning, Memory and Cognition* 10: 435-441.

Marcel, A. J. 1983. Conscious and Unconscious Perception: An Approach to the Relations between Phenomenal Experience and Perceptual Processes. *Cognitive Psychology* 15: 238-300.

Schilperoord, J., and Maes, A. 2009. Visual metaphoric conceptualization in political cartoons. In Forceville, C. & Urios-Aparisi, E. eds., *Multimodal Metaphor*. 211-238. Berlin: Mouton de Gruyter.

Schilperoord, J., Maes, A., and Ferdinandusse, H. 2009. Perceptual and Conceptual Visual Rhetoric: The Case of Symmetric Object Alignment. *Metaphor and Symbol* 24: 155-173.

Snodgrass, J. G., Hirshman, E., and Fan, J. 1996. The Sensory Match Effect in Recognition Memory: Perceptual Fluency or Episodic Trace? *Memory & cognition: a journal of the Psychonomic Society* 24(3): 367-383.

Steen, G. J. 2007. *Finding metaphor in grammar and usage: a methodological analysis of theory and research.* Amsterdam: Benjamins.

Teng, N. Y., and Sun, S. 2002. Grouping, Simile, and Oxymoron in Pictures: A Design-Based Cognitive Approach. *Metaphor and Symbol* 17(4): 295-316.

Tversky, A. 1977. Features of Similarity. *Psychological Review* 84(4): 327-352.