

On possible factors in the aesthetic appreciation of metaphors*

PÉTER CSATÁR, GERGELY PETHŐ, AND ENIKŐ TÓTH

Abstract

Gibbs (2002) investigated by experimental methods the question whether the recognition of metaphors in a literary text affects the aesthetic appreciation of recognized expressions. He found that recognition of metaphors improves aesthetic judgments. He explains this by reference to the conceptual theory of metaphor: Identifying a linguistic metaphor provides access to the underlying conceptual metaphor and opens up new ways of interpretation, thereby making the expression more appealing. In this paper, we note some conceptual arguments against Gibbs's conclusion and then present an experiment which aims to verify whether the improvement of judgments observed by Gibbs could have been caused by independent factors, such as social expectations. Surprisingly, although we did not find evidence for such effects, we were not able to reproduce Gibbs's results either. However, the data reveal that aesthetic judgments concerning metaphors are measurable to some extent.

0. Introduction

The fundamental purpose of this paper is to examine various factors that might be able to influence aesthetic judgments about metaphorical expressions. The starting point for our discussion is a paper written by Ray Gibbs (2002). That paper puts forward the hypothesis (which is later apparently confirmed by experimental methods) that judgments about the beauty of metaphors are affected, more exactly improved, if the informant has not only understood the expression in question, but has also recognized it as a metaphor.

Our goal has been on the one hand to check the validity of Gibbs's results, and on the other hand to examine the question whether at least some linguistic expressions possess an inherent aesthetic value. Although this latter idea is in a way presupposed by Gibbs's experiment, it is not trivial whether it is a justified

assumption, nor whether it can be considered a proper object of rational scientific investigation at all.

We first summarize Gibbs's experiment and our reservations, which have led us to try to reproduce his results. Next we discuss an experiment that we have carried out, and evaluate Gibbs's results in the light of ours. We conclude that none of the different factors that were hypothesized to be able to influence the aesthetic appreciation of metaphors do in fact have this effect. We could not verify the results of Gibbs (2002), and therefore his conclusions regarding the cognitive effects of poetic analysis turn out not to be supported by further empirical observations. However, we also conclude that the reason for this is not that the questions asked by Gibbs are wrong per se: we have found that aesthetic judgments concerning linguistic expressions are in fact open to rational investigation.

1. Summary of Gibbs's experiment (2002)

The aim of Gibbs's experiment (2002) was to find out whether the perceived aesthetic value of metaphors decreases or increases when they are consciously identified in a literary text. Although he does not state this explicitly, it can be inferred from his introduction that he examined this issue because he assumed that Lakoff and Johnson's (1980) theory of conceptual metaphor (of which Gibbs is a well-known representative as well) can be thereby supported by empirical means. At the same time, he assumed that the pragmatic theory of metaphor, which is regarded as a rival to the conceptual theory (cf. Gibbs 1994), could be falsified empirically in this way.¹ We will not examine here the problems that these two competing models try to solve in detail, nor do we evaluate the relevant conclusions of Gibbs's experiment. We will rather concentrate directly on the topic of the aesthetic appreciation of metaphors.

Gibbs intended to find out what happens when poems and other literary texts are analysed, or more specifically, when it is examined what poetic techniques are employed by the author in a certain text. **Many would regard the analysis of literary works as equal to overexplaining them. The magic of the text is assumed to be lost thereby,** i.e. the intensity of the aesthetic experience is supposed to decrease. Gibbs challenges this widely-held view. He finds that aesthetic judgments concerning metaphors actually improve when readers not only understand the text, but also realize which expressions in the texts can be considered metaphors.

Gibbs's experiment was designed in the following way. **A total of 60 university students participated in the experiment, divided into three groups.** The first task was to read a poem. Half of the participants in each group were asked to read the poem "Daffodils" by Wordsworth, and the other half a text by Sharon Olds, a contemporary American poet. Next the groups received different tasks. Members of the first group (A) were asked by the supervisor of the experiment

to read the same poem for a second time and underline the metaphors in it. It was not specified to participants just how the term metaphor should be understood, so they had to rely on their own “naive metaphor intuitions”. The task of the second group (B) was to reread the poem and underline all verbs, whereas members of the third group (C) had to reread the poem without any additional instructions. After this, the participants in all three groups received another sheet of paper with the poem that they had read previously. However, on each of these sheets, eight expressions in the text had been underlined. The third task for all groups was to rate how they liked these expressions on a scale of 1 to 7 (1: did not like it; 7: liked it very much). All of the underlined expressions were metaphors (according to Gibbs’s judgment), but this was not disclosed to the participants.

Gibbs was mainly interested in learning how judgments of aesthetic value were affected by the fact that a participant in group (A) successfully recognized (i.e. underlined) individual expressions as metaphors in the first task. For according to Gibbs’s hypothesis, the identification (recognition) of a metaphor should influence its aesthetic appreciation. This assumption could be regarded as confirmed if the aesthetic judgments of identified metaphors in group (A) differed significantly from the values assigned to the same metaphors by those participants who had not identified them.

Gibbs summarizes the results of his experiment in Table 1 (Gibbs 2002: 107).

Table 1. *Mean ratings of informants evaluating recognized vs. unrecognized metaphors in Gibbs (2002)*

Task		Poem	
		Wordsworth	Olds
looked for metaphors (A)	recognized metaphor	5.11	4.82
	unrecognized metaphor	3.99	3.67
looked for verbs (B)	recognized metaphor	4.61	4.26
	unrecognized metaphor	4.42	4.38
read poem twice (C)	recognized metaphor	4.65	4.39
	unrecognized metaphor	4.5	4.13

The data in Table 1 represent the arithmetical averages of the ratings given to recognized vs. unrecognized metaphors. It is striking that in the case of group (A) the average of judgments is obviously greater for expressions that have been identified as metaphors (i.e. underlined) by the participant than for those which have not been recognized. Furthermore, the recognized ones were judged to be more beautiful by group A than by the other two groups, B and C. These differences were statistically significant at $p < .01$.

A note is due regarding the rows labelled as “recognized” and “unrecognized” metaphors for groups (B) and (C) in Table 1. Although members of the groups in question did not have to recognize metaphors, Gibbs randomly assigned to every participant in group (A) a participant both from groups (B) and (C). If a specific metaphor was recognized by the participant in group (A), then the judgments made by these other two persons were also evaluated as if they had recognized the metaphors themselves. In this way it could be controlled that the reason why the values for recognized metaphors in group (A) were higher than the ones which have not been recognized was not simply that more beautiful metaphors can be recognized more easily than less beautiful ones.

Gibbs observed that there were no significant differences between the results of the three groups overall. Hence the task itself does not influence the aesthetic experience. This is sufficient in itself to refute the claim that the analysis of poetic works automatically diminishes the appreciation of poetry. For in the case of members of (A) in comparison with the control groups (B) and (C), not only did judgments about recognized metaphors differ, in this case positively, but also ones about unrecognized metaphors, although negatively.

In summary, Gibbs was able to confirm his hypothesis that the recognition of metaphors has an effect on aesthetic value judgments. Specifically, he observed that following their recognition metaphors receive higher approval ratings. Gibbs briefly explains the improvement of aesthetic value judgments in the framework of the cognitive theory of metaphor by claiming essentially that through the process of metaphor recognition, the so-called conceptual metaphors that the verbal metaphors are based upon become accessible. These conceptual metaphors activate new meanings and associations, and as a result, the perceived beauty of the respective linguistic expressions is supposedly enhanced.²

2. Problems in connection with Gibbs’s experiment

We believe that the argumentation of Gibbs (2002) is debatable for several reasons. The problems that arise can be sorted into at least three classes which we discuss in order below: 1) problems in connection with Gibbs’s critique of the pragmatic theory of metaphor; 2) difficulties in connection with the purported aesthetic effect of the poetic analysis of literary works; 3) further possible factors that might affect judgments of aesthetic value but were not considered by Gibbs.

There is a significant distortion involved in Gibbs’s presentation of the pragmatic conception of metaphor processing (our first class of difficulties). For he suggests that proponents of the pragmatic approach would identify the phase of the processing of a metaphor when the interpreter encounters a semantic anomaly (which is the notion of metaphor recognition that would indeed be considered relevant in this approach) with the conscious identification of the expression as a metaphor, which involves the explicit analysis of the type of se-

mantic anomaly. However, this latter process is not considered a relevant part of the recognition process according to the pragmatic approach; in fact they do not concern themselves with that at all. According to them, the necessary element is the non-conscious procedural knowledge that is a part of every speaker's semantic-pragmatic competence and that enables speakers to resolve this anomaly. By contrast, the search for metaphors and their classification is crucially a conscious process, and is therefore (contrary to Gibbs's claim) not assumed to be reducible to the process of recognizing an anomaly. To take an analogy from syntax: interpreting a sentence that contains a subject is a very different process from explicitly identifying the subject expression in the sentence. Hence it may be reasonably assumed that when persons in an experiment explicitly identify metaphors, they are not introspectively analysing their interpretation process, but are rather comparing the expressions to be classified to the metaphors they have learned at school as prototypes. This would in no way contradict the pragmatic model of metaphor.

Therefore, the process of metaphor recognition that Gibbs attributes to the pragmatic theory does not in fact follow from it. The speakers' introspective analysis of their interpretation process, with the aim of deciding whether they had encountered a certain kind of anomaly during this process, is not a part of the pragmatic model of metaphor processing.³ Because Gibbs is mistaken in his reconstruction of the pragmatic model of processing, his general conclusions concerning the theory of metaphors consequently miss the mark.

Independently of this, even if Gibbs had been right that the identification of an anomaly and the explicit underlining of a metaphor basically amount to the same thing, the results of his experiment would not refute this processing model. For let us consider the following:⁴ Assuming that the underlining of an expression reflects correctly whether the participants had recognized a metaphor during its processing, they could have understood only those metaphors which they had identified, since (according to Gibbs) identification is a prerequisite of the interpretation of a metaphor according to the pragmatic model. It follows that the participants had understood metaphors that they underlined, whereas they had not understood those that they did not underline. This being so, it is quite plausible to assume that people tend to appreciate expressions that make sense (to them) more than those that do not. And this in fact correlates perfectly with what Gibbs found: The recognized (i.e. successfully understood) metaphors were rated higher than both the unrecognized ones and the control groups' metaphors, and exactly the reverse was true for the unrecognized ones (i.e. those that had not been understood). The fact that the control groups' rates were halfway between these two poles simply means that there we do not know which metaphors were recognized (successfully understood) by the participants, and so the positive effect that accompanies understanding and the negative effect that accompanies lack of understanding balanced each other out. It

is worth noting that in fact Gibbs's mistakenly reconstructed and rejected pragmatic model fares much better with regard to his experimental results than his own explanation, since it not only answers the question why the ratings of the recognized metaphors were higher, but also why those of the unrecognized ones were lower (which Gibbs was not able to explain).

Turning to the second class of difficulties, we believe that the way Gibbs assesses the effects of poetic analysis on the basis of his data is not completely convincing. He claims that the improvement of the judgments of recognized metaphors he has observed provides an affirmative answer to his (primary) question, whether the analysis of poetry affects its appreciation, and, if so, in what way. However, this is incorrect. In fact, his data show that a metaphoric expression is only perceived as more beautiful if the respondent has recognized it earlier as such. In the opposite case it is actually perceived as less beautiful. These two effects counterbalance each other: as we mentioned above, the overall judgments given by group (A) did not differ significantly from the ones given by the other two groups. This means that looking for metaphors does not have any effect on the appreciation of a work of poetry, unless one recognizes all metaphors (unlike the subjects in the experiment). It is also strange that Gibbs equates the analysis of poetry with the identification of poetic techniques, which requires a rather simplistic notion of analysis. If at all, the experiment can only answer the question about the effects of analysis if this notion is understood in this restricted sense.

The third class of difficulties relates to a fundamental aspect of psycholinguistic experiments: the importance of identifying and differentiating the factors that lead to the results of the experiment. If this is not secure, the explanation of the results can easily be led astray. In this respect, Gibbs's conclusions are problematic because the only factor he took into consideration was the recognition of the metaphor, and he tried to explain the emerging patterns on the basis of this. However, the observed differences might also have been caused by other factors that were not controlled by Gibbs, and the possible involvement of which is not even mentioned in his paper.

One possible factor that might have led to the greater values in the case of recognized metaphors is that members of group (A) had already focussed their attention in the second task on these expressions when they identified them as metaphors, and therefore they might have viewed it as a positive feedback, as a confirmation of their earlier work when they met these expressions again in the third task. Incidentally, along these lines it could also be explained why the expressions that these participants had not recognized previously as metaphors received worse values than the same expressions in the control groups: the expressions in question can be regarded as a kind of negative feedback.

Another factor that is worth considering is related to special social norms connected to the notion of metaphor. We have learned as early as primary

school that metaphors are beautiful things, characteristic of refined use of language. If one does not assign a high enough aesthetic value to a metaphor, one would reveal that one does not have a keen sense of aesthetics, is not well-educated or does not care about the cultural traditions that one has been socialized in. We can therefore assume that participants in the experiment are under a significant unconscious social pressure when they have to evaluate an expression that they assume to be a metaphor, and they will accordingly judge these as more beautiful.

Since Gibbs did not control these two variables in his experiment, it is impossible to tell what role they actually played in the observed differences, or what effects other cognitive and social factors may have had. One of the aims of our experiment was exactly to investigate this question.

In the remainder of this paper, we will concentrate on this third group of problems. We have differentiated the following three factors in connection with Gibbs's experiment that may have influenced his results:

1. "The hypothesis of metaphorical meaning": The recognition of metaphorical meaning structures (and only this) leads to improved judgments of metaphors. This was Gibbs's conclusion.

2. "The hypothesis of social pressure": Judgments of expressions identified as metaphors improve because of social expectations connected to the notion of metaphor.

3. "The hypothesis of increased attention and positive feedback": The appreciation of those and only those expressions improves to which the participant in the experiment had to devote attention in a separate task.

In order to examine which factors indeed play a crucial role in determining the aesthetic appreciation of metaphors we carried out an experiment, which will be discussed in the following sections in detail.⁵

3. An experiment concerning possible factors of the aesthetic appreciation of metaphor

3.1. *Persons, materials, and methods*

The participants were Hungarian university students, 160 persons in total. They were assigned to eight groups, each consisting of about 20 students. Each group was characterized by different combinations of three binary factors: 1. *Kind of text read*: One half of the students (80 persons) read a well-known poem, the other half a relatively unknown prose text. 2. *Kind of task to be carried out*: Half of the students were asked to look for metaphors in the text, the other half to look for other items. 3. *Kind of expressions to be judged*: Only one half of the participants had to judge only metaphorical expressions. We had included some non-metaphorical (although stylistically marked) expressions among the ones to be rated by the other half. The emerging groups are represented in Table 2.

Table 2. Classification of groups by task and materials

		Poem	Prose
looked for metaphors	evaluated not only metaphors	1A	3A
	evaluated only metaphors	1B	3B
did not look for metaphors	evaluated not only metaphors	2A	4A
	evaluated only metaphors	2B	4B

With the help of this experimental design we were able to investigate the following:

1. In those groups that were looking for metaphors (1A, 3A, 1B, 3B), we could check whether the evaluation of metaphors that were recognized by the participants differed on the one hand compared to the metaphors not found by other participants in the same group, and on the other hand to the control group of this group, members of which were not looking for metaphors, but used the same answer sheets (i.e. had to rate the same expressions).

2. As we have mentioned, there were groups whose members had to judge not only metaphors. One half of these groups were asked to look for metaphors in the first task. Thereby, we wanted to suggest to them that the expressions they were judging in the evaluation task were in fact metaphors. In this way we checked the possible effects of social pressure.

Each participant was asked to complete a questionnaire that consisted of two parts. In the first task, one half of the groups had to read carefully a relatively long excerpt from a very popular poem, *Családi kör* ('Family circle'), written by one of the most well-known Hungarian poets, János Arany. The other half had to read a passage from a less known short story entitled *A pusztában kiáltó* ('Screaming in the wastes') written by a similarly popular poet, Gyula Juhász. The texts were selected bearing in mind that they should be literal depictions of scenes or everyday events containing relatively few genuinely metaphorical expressions, because the task of recognizing specific metaphors would have seemed pointless with a densely metaphorical text.

The second task was to underline metaphors in the case of one half of the groups, whereas the other half of the groups was asked to underline infinitives in the poem and names of occupations in the prose. It was only the metaphors that we were in fact interested in. The infinitives and names of occupations served only as a control.⁶

Then the participants were asked to turn over the page and the third task followed. On the other side of the page the previously read poem and prose appeared again, respectively. However, five expressions were highlighted in each text (by boldface type and underlining). We asked the participants to judge these five expressions on a scale of 1 to 9, 1: do not like it at all, 9: like it very much.

None of the groups were told that the expressions to be evaluated were metaphors in order to be able to repeat Gibbs's experiment as faithfully as possible.

The following pairs of groups were given the same expressions in the poem or prose to rate: 1A and 2A, 1B and 2B, 3A and 4A, 3B and 4B. In other words, the questionnaire they worked with was completely identical. The only difference between the members of each of these pairs of groups was that they had to look for different expressions (metaphors and not metaphors, respectively) in the first task.

We do not include a more detailed summary of the materials here, partly because of lack of space and partly because we do not think that raw translations of the texts and the specific expressions would render the following discussion significantly easier to follow. We refer those who can read Hungarian to Pethő et al. (2005), which includes all of the materials and a much more detailed presentation of the results.

3.2. *Reproducing Gibbs (2002)'s results*

Let us now see the results for each group and whether we were able to replicate the results of Gibbs's experiment. To check this we have to look at those four groups that were only evaluating metaphors in the third task. We also repeat Gibbs's relevant data for the sake of comparison in Table 4.

Table 3. *Group ratings averages from test replication*

	Average	Average of recognized expressions (percentage of recognition)	Average of unrecognized expressions	Average of the control group
Poem (1B)	6.70	6.66 (64%)	6.76	6.71
Prose (3B)	6.22	6.27 (53%)	6.16	6.32

(1B): 23 persons, (2B) (control): 18 persons

(3B): 21 persons, (4B) (control): 19 persons

Table 4. *Summary of results from Gibbs (2002) test*

Gibbs's data	Average of recognized expressions	Average of unrecognized expressions	Average of the control group
Wordsworth	5.11	3.99	4.61
Olds	4.82	3.67	4.26

Group (1B) was asked to look for metaphors, whereas the control group (2B) was looking for infinitives in the poem. Group (3B) was also looking for metaphors, its control group (4B) was looking for names of occupations. There were no significant overall differences between the averages of the groups and their controls (the second and fifth column in Table 3, respectively). This confirms

Gibbs's findings: the differences between his two groups and their controls were not statistically significant either. However, we were not able to reproduce the significant differences that he observed between the values given by participants who recognized a particular metaphor and those who did not recognize it. In his experiment, the recognized metaphors received significantly higher values, whereas in ours there was no significant difference at all in any direction, compare the grey columns in Tables 3 and 4.

Thus, we were not able to repeat Gibbs's results in a task that was essentially identical to his own, with groups that were put together in essentially the same way as his. Not only was there no difference in the overall averages of the groups, there was none with respect to the recognized and unrecognized values of the individual metaphors either. (The analysis was based on the Mann-Whitney test in all cases.)

We see no obvious explanation for very clear and statistically significant differences between the relevant groups appearing in Gibbs's experiment and missing in the case of our texts. We were unable to reproduce either the improvement of the values the metaphors received when they were recognized or their deterioration when they were not recognized.

3.3. *Testing for the effects of social expectations*

After having summarized the part of our experiment that was structured analogously to Gibbs's groups, let us now turn to its other part. We wanted to know what happens if we asked respondents to look explicitly for metaphors in the second task but made them judge expressions including non-metaphorical ones. Our hypothesis was that the social expectations connected with metaphors would influence in this case the judgments given about the expressions. We assumed that this would be confirmed if the group that was asked to look for metaphors gave higher ratings to non-metaphorical expressions hidden among real metaphors, in comparison with the control group, at least if the respondents believed us that they were judging metaphors.

The questionnaires contained the same passages as in the case of the groups described above. However, in this case 3 metaphors and 2 non-metaphorical expressions were highlighted in both texts. Let us see now how the non-metaphorical expressions were evaluated in the groups looking for metaphors (1A and 2A) and in their respective control groups (3A and 4A).

Table 5. *Evaluation of non-metaphorical expressions in test and control groups*

	Average (1A, 3A)	Average in control groups (2A, 4A)
non-metaphorical expressions	5.9	5.66

(1A): 20 persons, (2A) (control): 19 persons

(3A): 19 persons, (4A) (control): 19 persons

It can be observed that if we suggest that an expression is metaphorical, this leads to a slight improvement of its ratings over the control, although the difference is not significant. Thus the results of the experiment do not clearly confirm that social expectations have a positive effect on aesthetic value judgments concerning metaphors, although they do not contradict this hypothesis either.

3.4. *Conclusions*

We have not been able to reproduce Gibbs's results. However, we were not able to support our hypothesis that social expectations exert a significant influence on the reported aesthetic appreciation of metaphors, either. Although some data weakly point toward this conclusion, the expected effects could not be observed reliably.

4. Can the aesthetic appreciation of metaphors be measured?

Overall, the analysis of the experimental data showed no clear, consistently reproducible effects. Therefore, the question arises whether it makes sense at all to ask people in an experiment how beautiful they think certain expressions are. Whereas in normal communication we clearly do need to understand expressions, it is not at all trivial to assume that aesthetic evaluation is an integral part of our linguistic capabilities as well. But when Gibbs investigated in his experiment whether a certain factor (the recognition of metaphoricity) affects the aesthetic appreciation of metaphors, he had to assume implicitly that some aesthetic value is always assigned to expressions under the given circumstances.

Because of this, we analysed the judgments we had gathered from this point of view as well, i.e. we intended to verify whether the participants of the experiment were able to make sense of the appreciation task and solve it accordingly. It became clear from the analysis of the data that the respondents agreed to a fairly high degree as to which expressions they liked or disliked. The differences in ratings for pairs of expressions were in part rather small across the board (when the participants agreed that two expressions were similarly beautiful) and in part very large (when participants agreed that one was beautiful and the other one less so). These latter differences were consistently significant to a very high degree ($p < .001$ based on the Friedman test for a large number of pairs of expressions). Therefore we were able to conclude that the participants had not rated the expressions at random. Instead, there was extensive interpersonal agreement as to whether the aesthetic value of a particular expression should be 4 or 7, for example (the standard deviation for the ratings of each expression was consistently around 2). There was no difference in this respect between metaphors and non-metaphors, i.e. both were judged equally consistently by the participants.

Furthermore, we also checked how participants in the experiments had gone about rating the expressions, in particular (1) whether they evaluated each individually, in other words, the rating was absolute, or (2) whether they evaluated each compared to the other expressions, in other words, relatively. We had constructed the experiment so as to be able to examine this issue as well. On the questionnaires for groups (3–4A) and (3–4B), the participants had to judge in part identical expressions: the three metaphors that were selected for the A groups were among those in the B groups. Comparison of the ratings given for these identical metaphors across the groups showed that the differences between them were not statistically significant (according to the Friedman test). At the same time, the remaining two expressions (non-metaphors for A and two further metaphors for B) received differing ratings. So the respondents had probably rated absolutely rather than relatively.

We can conclude that it does indeed make sense to talk about an objective measure of aesthetic judgments in some sense for both metaphorical and non-metaphorical expressions. Consequently, it should also make sense to ask firstly what such aesthetic judgments are based upon, and secondly by what factors these judgments may be influenced, if at all. However, since the reported questionnaire studies failed to yield consistent data, it is still unclear what methods should be employed in order to get useful results.

University of Debrecen (Csátár and Tóth)
University of Edinburgh (Pethő)

Notes

- * The authors would like to thank the participants of the Semantic Intuitions research seminar at the University of Debrecen for their comments on previous versions of this paper. Our research was supported by grant F42664 of the Hungarian Scientific Research Fund. Gergely Pethő's work was supported by the Research Group for Theoretical Linguistics of the Hungarian Academy of Sciences at the University of Debrecen.
1. On the pragmatic theory of metaphor, cf. Loewenberg (1975) and Searle (1979), and on the conceptual theory, Lakoff and Johnson (1980) and Lakoff (1993). Note that Gibbs does not refer to these two approaches explicitly in these terms, but the two positions concerning the interpretation of metaphors that he writes about can be clearly identified as the pragmatic approach on the one hand and Gibbs's own (conceptual) position on the other. Cf. the following two characterizations: "listeners/readers find some utterance to be anomalous [...] and then apply special mental processes to infer the speaker's/writer's intention to express metaphoric meaning" vs. "metaphor understanding and appreciation need not require that people first identify, consciously or otherwise, that the linguistic statement they are hearing/reading is a 'metaphor'" (Gibbs 2002: 102).

2. In Gibbs (2002), a second experiment is also presented where the participants were asked to judge the strength of their own emotional reactions to the same metaphors as in this experiment on a similar scale of 1 to 7. We will not have anything to say about that part of the paper.
3. On a side note, it is highly questionable whether speakers are able to access their linguistic processing abilities retrospectively, and whether they are able to reconstruct specific elements of the (e.g. interpretation) processes that automatically take place in their minds, cf. Schütze (1996).
4. We have to thank Csilla Rákosi for having pointed out this argument to us.
5. The effect of the third factor is not considered in the experiment reported here. In a pilot study we have examined the effect of this factor as well, and were unable to confirm it, however, we cannot discuss the exact results here due to lack of space. We refer the reader to Pethő et al. (2005) for a more complete discussion.
6. The choice of items that members of the control groups were asked to look for in the text was mainly determined by the following factors: the number of their tokens had to be neither too large, nor too small in that text, and the task should not be too trivial. Since the two texts were very different from each other in every respect, it was not possible to find a common type of control items for both. However, the fact that the control task was not identical for all participants could not have influenced the results in any way, because both tasks were chosen so as not to interfere with the appreciation task: there was no overlap between the control items and the expressions to be judged in either group, and both the infinitives and the names of occupations were aesthetically neutral.

References

- Gibbs, R. W. (1994). *The Poetics of Mind: Figurative Thought, Language, and Understanding*. Cambridge: Cambridge University Press.
- (2002). Identifying and appreciating poetic metaphor. *Journal of Literary Semantics* 31: 101–112.
- Lakoff, G. (1993). The contemporary theory of metaphor. In *Metaphor and Thought*, 2nd ed., A. Ortony (ed.), 202–225. Cambridge: Cambridge University Press.
- Lakoff, G. and M. Johnson (1980). *Metaphors We Live By*. Chicago: Chicago University Press.
- Pethő, G., P. Csátár, and E. Tóth (2005). A metaforák esztétikai értékelésének tényezői [Factors of the aesthetic appreciation of metaphors]. In *A láthatatlan nyelv [The invisible language] II*, J. Gervain and Cs. Pléh. Budapest: Osiris.
- Loewenberg, I. (1975). Identifying metaphors. *Foundations of Language* 12: 315–338.
- Schütze, C. (1996). *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Chicago: University of Chicago Press.
- Searle, J. R. (1979). Metaphor. In *Expression and Meaning*, J. R. Searle, 76–116. Cambridge: Cambridge University Press.

Copyright of *Journal of Literary Semantics* is the property of Walter de Gruyter GmbH & Co. KG. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.