METAPHOR AND SYMBOLIC ACTIVITY, 2(4), 239–250

Metaphorical Uses of Language in the Expression of Emotions

Lynn Fainsilber and Andrew Ortony
Center for the Study of Reading
University of Illinois

Two hypotheses concerning the function of metaphor production were examined in participants' descriptions of instances of emotional states and events. Participants provided verbal descriptions of emotional states they had experienced and of actions in which they had engaged as they experienced those states. Results showed that descriptions of emotional states contained more metaphorical language than did descriptions of behaviors. This is interpreted as supporting the hypothesis that an important function of metaphorical language is to permit the expression of that which is difficult to express using literal language alone. The results also indicated that intense emotions led to greater metaphor use than mild emotions for descriptions of feeling states but not for descriptions of actions associated with emotions. This result was interpreted as consistent with the view that metaphors provide a particularly vivid form of description.

The last few years have seen a sharp increase in the attention devoted to the comprehension of metaphorical uses of language (e.g., Honeck & Hoffman, 1980; Ortony, 1979), and these inquiries have raised a number of interesting questions regarding the cognitive processes involved in understanding metaphors. However relatively little systematic effort has been devoted to examining metaphor production. This lack of attention to metaphor production is somewhat surprising, particularly given that it is now generally accepted by psychologists and linguists that metaphors, and their close cousins, analogies, are important tools of cognition and communication, providing us with unfamiliar ways of conceptualizing unfamiliar things (Lakoff & Johnson, 1980; Ortony, 1979; Vosniadou & Ortony, in preparation). It has yet to be demonstrated, rather than being merely assumed, that nonliteral

Requests for reprints should be sent to Andrew Ortony, Center for the Study of Reading, University of Illinois, 51 Gerty Drive, Champaign, IL 61820.

uses of language are sometimes *necessary* for accomplishing such goals, as opposed to being simply convenient or elegant ways of doing so. In part, this may reflect the cognitive psychologist's traditional focus on comprehension mechanisms in language, but it may also be a reflection of a problem specific to the examination of metaphor production: how to create a situation in which metaphors will be produced with sufficient frequency to permit their systematic study.

In this article we present a sort of empirical existence proof that there are some things whose descriptions appear to invoke much more use of metaphorical language than others. This, although not establishing the necessity of metaphors, is certainly a first step. Our approach to the problem of finding a discourse domain rich in the spontaneous production of metaphorical uses of language was to consider first what some of the communicative functions of metaphors might be.

In theory, there are at least three communicative functions that metaphor may serve (Ortony, 1975). First, metaphors may allow one to express that which would be difficult or impossible to express if one were restricted to literal uses of language. Consider the metaphor "*The thought slipped my mind like a squirrel behind a tree.*" The characteristics of squirrels slipping behind trees that one would want to predicate of thoughts (e.g., swiftness, suddenness, ungraspableness) are difficult to express using literal language alone. These characteristics seem to be applicable to thoughts only metaphorically—so that attempts to translate the metaphorical sentence into literal language result in the speaker still having to resort to metaphorical explication (e.g., "*The thought went away*" and "*The thought evaded me*"). Thus, the *inexpressibility hypothesis* proposes that metaphors may enable the communication of that which cannot be readily expressed using literal language. Evidence for this inexpressibility claim would constitute encouraging support for the necessity-for-metaphors view.

A second possible function of metaphors is captured by the *compactness hypothesis,* which asserts that metaphors may constitute a particularly compact means of communication. Although conscious experience is continuous in form, the linguistic system we use to talk about this experience is comprised of discrete elements (lexical items). Language partitions the continuity of experience into discrete units comprised of words and phrases having a relatively narrow referential range. So, for example, "The apple is red" communicates only a single piece of information about some particular apple. It may be that metaphors help to convey "chunks" of information rather than discrete units. When we assert that "*Love is like a red, red rose,*" we communicate a wealth of information regarding love (i.e., that it is sweet, delicate, beautiful, etc.), and we do so using relatively few words. Thus, unlike more literal forms of language, metaphor may enable us to convey a great deal of information succinctly.

Finally, Ortony (1975) proposed a *vividness hypothesis*, which suggests that metaphors may help capture the vividness of phenomenal experience. If metaphors convey chunks of information rather than discrete units, they can paint a richer and more detailed picture of our subjective experience than might be expressed by literal language. These details seem to embellish the communication, providing the listener with subtle nuances that may be part of the speaker's subjective experience. Thus, to say that "*Love is like a red, red rose*" is likely to conjure up perceptual and sensory images in the listener, and in so doing is likely to better reflect the vividness of that which is to be communicated.

To summarize, three potential functions of metaphor are suggested: (a) the expression of ideas that may be difficult or impossible to express using literal language, (b) the expression of ideas compactly, and (c) the expression of ideas vividly. We focused on the inexpressibility and vividness hypotheses. To do so, a discourse domain had to be selected for which a *prima facie* case could be made for supposing that literal language would often be inadequate, and in which one might find varying degrees of vividness of the to-be-described material. There are doubtless many domains that satisfy these conditions. The one that we selected was that of internal states, in particular, emotional states. The literature on the linguistic expression of emotions suggests a relatively high incidence of figurative language use. For example, Davitz (1969) asked people to describe emotional states and informally noted that many of the descriptions obtained were metaphorical. Similarly, Davitz and Mattis (1964) reported that sensitivity to metaphorical expressions seemed to relate to the vocal expression of emotion. Observations such as these provide pragmatic reasons for believing that the context of (linguistic) emotional expression may be a profitable one within which to study metaphor production. Emotional states also seemed well suited for our purposes because they tend to have an elusive, transient quality that is difficult to describe using literal language, although, of course, they can usually be *labeled* using literal language. Thus, while it might be easy for a person to label an emotional state as, for example, fear, it is difficult to provide a literal description of the *quality* of some particular experience of fear.

There seem to be two possible ways in which people might try to communicate the quality of an emotional state. Taking an indirect route, a speaker might use literal language to describe the events that triggered the emotional state and hope that the hearer correctly infers how he or she felt, presumably through some sort of imaginal or identification process. For example, a person might describe the details of being mugged, hoping that a listener would recognize the emotional experience as the type one would have if one were attacked by a mugger. In such a case, the literal description would not describe the *quality* of the subjective state itself but would merely identify its eliciting conditions (Ortony, Clore, & Collins, in presss). Alternatively, a speaker

might use a metaphor in an attempt to describe the quality of an emotional state. For example, one might say that one felt as though one's insides were a "butter churn." Here, the metaphorical description does represent an attempt to characterize the quality of the subjective state.

Although we think that emotions constitute a good domain for studying metaphor production, it does not follow that metaphorical modes of description will be equally prevalent for different facets of emotions. Emotion theorists frequently attribute differential significance to the subjective experience of emotion (de Rivera, 1977) or to their associated actions or action tendencies (Frijda, 1987). It may be, for example, that the subjective experience of an emotion can benefit more from a metaphorical description than the associated action or action tendency. Consider the subjective experience of some specific case of anger. The *quality* of such a subjective state cannot be publicly observed. In contrast, the actions to which an anger experience might give rise, for example, pounding one's fist on the table, *are* publicly observable. Thus, one might expect people to employ more metaphorical descriptions when trying to characterize the subjective experiential quality of emotional states than when trying to characterize the overt behaviors associated with such states. So the quality of emotional experience seems to be a good candidate for testing the inexpressibility hypothesis whereas the actions associated with such experiences might be expected to be more amenable to literal description.

Another aspect of emotions that might be expected to influence the use of metaphorical language is their intensity. It is possible that relatively mild emotional states are sufficiently unremarkable that speakers are more willing to settle for simply labeling them, whereas intense emotional states might sometimes generate a more pressing need for detailed description. If this is right, then one might expect verbal descriptions of intense emotional experiences to contain more metaphorical language than descriptions of less intense experiences. This possibility might be viewed as one of the consequences of the vividness hypothesis.

Our investigation examined the production of metaphors in descriptions of emotional states and events, with the predictions being derived from our analysis of the possible communicative functions of metaphors. First, from the inexpressibility hypothesis, we predicted that people would be more likely to use metaphors and metaphorical comparisons in descriptions of how they *felt* when they were experiencing an emotion than in descriptions of what they *did* when they experienced it. Second, the vividness hypothesis suggests that more metaphorical language might be used in descriptions of intense as compared to mild emotional states. The two hypotheses combined could be construed as predicting an interaction of description type (feelings vs. actions) and intensity, with the intensity factor having a greater effect on feeling descriptions than on action descriptions. Descriptions of feeling states,

which may already make use of metaphorical language, may be especially likely to include metaphors when the states are intense. On the other hand, it could be argued that although intense emotions are more vivid than less intense ones, the associated actions do not necessarily enjoy a corresponding increase in vividness. This is an admittedly tenuous argument, so the prediction of an interaction between description type and intensity is made with less confidence than the predictions of main effects for these variables. Finally, in the experiment to test these hypotheses, the valence of the emotions was manipulated to determine whether this factor has any systematic effect on metaphor use.

METHOD

Design

A 2 (Description Type: Feelings vs. Actions) × 2 (Emotional Intensity: Mild vs. Intense) × 2 (Valence: Positive vs. Negative) within-subjects factorial design was employed. Factors were completely crossed and repeated across subjects so that each participant experienced all combinations of all conditions. Presentation order of each variable was randomized in a sequential fashion. First, emotion words were randomly ordered. For each emotion, order of presentation of the two intensity levels (mild vs. intense) was then randomized. Finally, the description type variable was randomized for each intensity level of each emotion.

Participants and Materials

Participants were 25 introductory psychology students at the University of Illinois who received course credit for their participation. Approximately half the participants were male, and half female. All participants were tested individually and the testing session was audiotaped and later transcribed.

The emotions used included four positive ones (happiness, pride, gratitude, and relief) and four negative ones (sadness, fear, resentment, and shame). These specific emotions were selected to include both general emotions (e.g., happiness, sadness) and more specific ones (e.g., gratitude, shame). Some effort was made to use a rather heterogeneous set of emotions, although it should be noted that the particular hypotheses to be tested do not depend in any important way on exactly which emotions are used.

Procedure

Participant and experimenter were seated opposite each other in a comfortable room. The experimental procedure consisted of a structured interview

lasting from 1 to 1½ hr. Participants gave oral responses to the experimenter's questions, and both questions and responses were tape recorded.

They were told that the experiment was intended to investigate how people experience different events, and that they would be asked to bring to mind situations in which they had experienced different emotions. For each emotion, participants were asked to recall two situations in which they had experienced it, one in which they had experienced the emotion mildly and one in which they had experienced it to an intense degree. In an effort to ensure that the situations chosen were of relatively greater and lesser intensities, we instructed them to silently compare those two situations to make sure that one event really had generated a much stronger emotional reaction than the other. Participants were then asked to provide the experimenter with labels, descriptive of the events, that they could later use as a memory aid to again bring the events to mind.

Once labels were obtained for both the mild and intense forms of all emotions, participants were asked to recall the situations one at a time. The experimenter prompted each individual with the label that he or she had previously selected as a reminder of each specific event. Then, for each situation, they were encouraged to bring the incident to mind as vividly as possible—to remember where they were, with whom, and what they were doing. They were then asked to describe either the feelings they had experienced during the event, or the actions they took in response to their feelings. Both feeling and action questions were asked of each intensity level of each emotion, so that each individual responded to 32 questions. The order in which these questions were asked was randomly determined, with the restriction that half the responses for each individual begin with the feeling question and the other half begin with the action question.

For the feeling question, participants were asked to describe the subjective quality of their emotion, or *how they felt inside* when they felt that emotion. For the action question, they were asked to describe *what they did* in response to their feelings. They were instructed to describe any actions they had performed that might have indicated to others that they were experiencing that emotion. To help circumvent any reluctance about reporting the details of the emotion-inducing events themselves, they were explicitly told *not* to describe the events, but to focus on their feelings and actions instead.

The description task was conducted only after completion of the labeling task. This was done to avoid the possibility that knowledge of the task requirements might bias participants in their choice of emotional events.

RESULTS

Metaphors were identified in the transcripts of interview sessions. Two scoring methods were adopted. Transcripts were coded by judges who were

trained to recognize metaphors in practiced items, and who only began coding the experimental data once they had achieved a .70 level of interrater reliability. Interrater reliability also remained at this level during the coding of the experimental data.[1]

Metaphors were also identified in the transcripts by the authors. This procedure was used for two reasons. First, we wanted to check validity, because the interrater reliability level, though acceptable, was less than ideal. Second, we wanted to classify the metaphors produced with respect to their frozenness (i.e., the degree to which they had become conventional English phrases). The authors coded the data blind to experimental condition. Disagreements occurred in less than 5% of the cases.

Verbal output was scored in terms of idea units (Johnson, 1970). Idea units were adopted as the basis for the dependent measure because the items of interest—metaphors—are generally better conceptualized as single ideas than as individual words.

Five different indices of metaphor production were used. Because the results of analyses using these different measures were virtually identical, only those based on the most conservative index of metaphor production are reported, except where otherwise noted. This index was computed by calculating the proportion of all *distinct* idea units that were metaphorical in nature.[2] In other words, this measure is based on the ratio of metaphor types to the total number of different idea-unit types (i.e., the number of idea-unit types) appearing in a protocol. A variety of considerations led us to prefer operationalizing metaphor production in this manner. We were concerned that possible systematic differences in the amount of verbal output produced during descriptions of the different emotion-inducing events might contaminate the measure of metaphor production, such that high verbal output might lead to a high production of metaphor, and low verbal output might be associated with low metaphor use. If so, metaphor production would be a consequence of verbal output per se and this effect might conceal any differential use of metaphor during descriptions of feelings and actions. By looking at the ratio of metaphor types to all idea-unit types (i.e., the sum of both metaphorical and nonmetaphorical idea units), the potential confounding of metaphor production and amount of linguistic output was partially avoided. A second concern was that participants' tendency to repeat words and phrases during an oral account might distort the results if measures in terms of tokens (as opposed to types) were used.

---

[1]This figure compares favorably with studies of metaphor production that report interrater reliability between two coders. Gore (1977) found pairwise agreement ranging from .40 to .64, with a mean of .51. Pollio and Barlow (1975) reported values ranging from .74 to .84.

[2]The other four indices were (a) proportion of the total number of idea units that were metaphorical (i.e., ratio of metaphor tokens to idea-unit tokens), (b) proportion of the total number of idea units that were metaphor types (i.e., ratio of metaphor types to idea-unit tokens), (c) number of metaphor types, and (d) number of metaphor tokens.

Due to an error on the part of the interviewer, approximately 15% of the data were missing. Chi-square tests indicated that missing data were randomly distributed across all cells of the design, suggesting that the missing values were not confounded with any factors in the design. To handle the problem of missing data, average values of metaphor production and verbal output were calculated across the four emotions within each valence type. Analyses were conducted using these mean values as raw data. Separate analyses were performed on the data coded by trained judges and on the data coded by the authors.

Coding by Trained Judges

An analysis of variance was performed on arcsin tranformations (Winer, 1962) of the proportion of metaphor types. Variables were description type (feeling states vs. actions), intensity (mild vs. intense), and valence (positive vs. negative), with subjects crossed with all variables. The main effect of description type was significant, $F(1, 24) = 25.40, p < .01$, with a higher percentage of metaphor types used in descriptions of feeling states ($M = 13\%$) than in descriptions of actions ($M = 5\%$). Using this measure, no other significant effects were found. The choice of index of metaphor production did make a difference in the analyses of intensity. When looking at the number of metaphor types produced without taking total verbal output into account, there was a main effect for intensity, $F(1, 24) = 7.81, p < .01$, as well as a Valence $\times$ Intensity interaction, $F(1, 24) = 3.95, p < .05$. More metaphors were used to describe intense emotions than to describe mild ones, with this difference being particularly pronounced for descriptions of positive emotions (mean number of metaphors: for intense positive emotions, $M = 16$; for mild positive emotions, $M = 13.6$). People's use of metaphor during negative emotional states was not as influenced by the intensity of the emotion being experienced. Finally, although there was no significant interaction between description type and intensity, $F(1, 24) = 2.75, p = .11$, the pattern of the data was in the expected direction when raw frequency of metaphors was the dependent measure (mean number of metaphors: for intense feelings, $M = 19.2$; for mild feelings, $M = 16.8$; for actions associated with intense emotions, $M = 12.8$; for actions associated with mild emotions, $M = 12.0$).

Coding by the Authors

Results for data coded by the authors replicated and extended those found by trained judges. As before, a main effect for description type was found, $F(1, 24) = 47.83, p < .01$, for the proportion of metaphors used; main effects for description type, $F(1, 24) = 76.67, p < .01$, and intensity, $F(1, 24)$

$= 14.51, p < .01$, were found for metaphor production when verbal output was not taken into account. Some new patterns also emerged for the proportional measure of metaphor production. A main effect for intensity, $F(1, 24) = 6.13, p < .01$, was found parallel to the analysis reported above based on raw frequency of metaphors. In addition, the trend toward an Intensity × Description Type interaction found in the analysis based on the judges' coding appeared as a significant interaction in this analysis, $F(1, 24) = 4.23$, $p < .05$. More metaphors were used in descriptions of intense emotions than in descriptions of mild ones, although this was mainly during descriptions of feeling states (mean percentage of metaphor types: for intense feelings, $M = 19.6\%$; for mild feelings, $M = 14.7\%$). There was a much smaller difference in metaphor use in descriptions of actions associated with mild versus intense emotions (mean percentage of metaphors: for actions during intense emotions, $M = 4.5\%$; for actions during mild emotions, $M = 3.5\%$).

DISCUSSION

Our findings are consistent with the predictions of the inexpressibility hypothesis of metaphor production. The quality of an unobservable internal feeling state is often difficult to describe using literal language, and our results showing the predominance of metaphorical language during descriptions of feeling states as opposed to actions suggest that metaphors may help people capture that which is otherwise difficult or impossible to convey. Evidence in support of the vividness hypothesis was also found in the form of a main effect for intensity. However, this effect was somewhat sensitive to the particular measure and the particular coder that were used. When coding was done by the authors, the results revealed that descriptions of intense emotions did lead to greater use of metaphor than did descriptions of mild ones. This effect was also found using the less conservative measure of raw frequency of metaphor types in the trained judges' coding of the data. Finally, there was modest support for the tentatively predicted Description Type × Intensity interaction. The authors' coding revealed that the increase in metaphor production for intense emotions was almost entirely due to this difference in descriptions of feelings as opposed to actions. The same pattern of data, although not statistically significant, was evident in the trained judges' coding using raw frequencies.

To understand what the relationship between emotional intensity and metaphor production might be, we conducted an informal analysis of the frozenness of the metaphors. We reasoned that not only might intense feeling states yield a greater *number* of metaphors than mild ones, but that the *quality* of the metaphors produced during intense emotions might also be different. Unfortunately, on the whole the metaphors that participants produced were

disappointingly banal. They used eight times as many frozen, or dead, metaphors as novel ones. However, the ratio of novel to frozen metaphors was greater for intense emotions (12%) than for mild ones (8%), suggesting perhaps that when people are describing intense feeling states, they are more likely to generate striking and complex metaphors to explain how they feel. These results lend mild support to the prediction, related to the vividness hypothesis, that more intense (and presumably, therefore, more vivid) emotional states not only generated more metaphors, but also resulted in richer and more vivid metaphors. Examples of some metaphors classified as novel in describing positive emotions included "*It was like a very bright light was just shining outward*" and "*I felt like I was in the middle of a pressure chamber and pressures were coming every which way—then the pressure just went away*." Examples of novel metaphors in response to negative emotions included "*It was like someone had just dropped a bomb on me*," "*I was just kind of hoping to go off in a big pile of smoke*," and "*I felt like a reconstructive system*" (sic). Frozen metaphors included references to such expressions as "*warm feelings*," "*a load off one's back*," "*bottling things up*," "*burning sensations inside*," and others.

The data raise many unanswered questions about the relationship between emotional intensity and metaphor production for individual emotions. For example, is metaphor production comparable for mild instances of different emotions (e.g., mild anger vs. mild fear)? Our procedures made it impossible to conduct such a comparison because there is no way to know whether the subjective intensity of, for example, an instance of mild anger that an individual brought to mind was equivalent to that for an instance of mild fear. It might be possible to investigate such issues by having people rate the intensity of each emotion, although we suspect that the problem of cross-emotion calibration would make this very difficult to do.

An interesting possibility suggested by the data is that the use of metaphorical expressions may vary with different emotions. An examination of means indicated that relief produced greatest metaphor use, followed by pride, shame, sadness, and fear. Happiness and resentment came next, with gratitude producing the fewest metaphors. However, before conclusions can be drawn from this observations, it would be necessary to develop procedures to explicitly examine the issue in order to obtain reliable effects and to rule out uninteresting explanations. For example, differential metaphor production for different emotions might merely reflect different frequencies of experience of different emotions. Other (unpublished) research that we have conducted shows that relief, sadness, shame, pride, and fear occur more frequently than do resentment and gratitude. Perhaps participants were better able to recall events that had induced the more frequently occurring emotions because they were more likely to have experienced them in the recent past. Better recall of events associated with the more frequently occurring emo-

tions may have led to higher metaphor use because recent events probably have more detailed memorial representations than distant events.

To summarize, our findings are consistent with the inexpressibility and vividness hypotheses of metaphor production, suggesting that metaphorical language may sometimes be a valuable tool that enables people to convey the *quality* of their internal feeling states to others. Metaphors may be used because the subjective qualities of our emotions cannot be described using literal descriptions. So, for example, when one of our participants reported that he felt like "*a storm was brewing inside*," he succeeded in conveying a particular quality of his subjective experience that is richer and more specific than could have been conveyed had he merely labeled the experience as "resentment." For the most part, the types of metaphors that people used to describe their emotions were figurative forms that have become conventionalized in the English language. When (relatively) novel metaphors were used, they seemed to be particularly evident in descriptions of intense feeling states. To the extent that this is true, one would have to reject the classical Aristotelian view of metaphor as merely linguistic decoration, in favor of a view that accords it an indispensable communicative function.

ACKNOWLEDGMENTS

This research was funded in part by a doctoral fellowship from the Social Sciences and Humanities Research Council of Canada to the first author, in part by National Science Foundation Grant BNS-83-18077, and in part by National Institute of Education Contract NIE-400-81-0030.

The authors thank Bruce Rapkin, Mark Foss, and William Chaplin for their statistical advice and Gerald Clore and Stephen Levine for their helpful comments on an earlier version of this article.

REFERENCES

Davitz, J. R. (1969). *The language of emotion.* New York: Academic.

Davitz, J. R., & Mattis, S. (1964). The communication of emotional meaning by metaphor. In J. R. Davitz (Ed.), *The communication of emotional meaning* (pp. 157–176). New York: McGraw-Hill.

deRivera, J. (1977). *A structural theory of the emotions.* New York: International Universities Press.

Frijda, N. (1987). *The emotions.* New York: Cambridge University Press.

Gore, N. S. (1977). Psychological functions of metaphor. *Dissertation Abstracts International, 38,* 2861B. (University Microfilms No. 77-26, 248)

Honeck, R.P., & Hoffman, R. R. (Eds.). (1980). *Cognition and figurative language.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Johnson, R. E. (1970). Recall of prose as a function of the structural importance of the linguistic unit. *Journal of Verbal Learning and Verbal Behavior, 29,* 12–20.

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

Ortony, A. (1975). Why metaphors are necessary and not just nice. *Educational Theory, 25,* 45–53.

Ortony, A. (Ed.). (1979). *Metaphor and thought.* New York: Cambridge University Press.

Ortony, A., Clore, G. L., & Collins, A. (in press). *The cognitive structure of emotions.* New York: Cambridge University Press.

Pollio, H. R., & Barlow, J. M. (1975). A behavioral analysis of figurative language in psychotherapy: One session in a single case study. *Language and Speech, 18,* 236–254.

Vosniadou, S., & Ortony, A. (Eds.). (in preparation). *Similarity and analogical reasoning.* New York: Cambridge University Press.

Winer, B. J. (1962). *Statistical principles in experimental design.* New York: McGraw-Hill.