Literal Meaning and the Comprehension of Metaphors

Steven L. Lytinen and Robert R. Burridge and Jeffrey D. Kirtner

Artificial Intelligence Laboratory
The University of Michigan
Ann Arbor, MI 48109

Abstract

Based on psychological studies which show that metaphors and other nonliteral constructions are comprehended in the same amount of time as comparable literal constructions, some researchers have concluded that literal meaning is not computed during metaphor comprehension. In this paper, we suggest that the empirical evidence does not rule out the possibility that literal meaning is constructed. We present a computational model of metaphor comprehension which is consistent with the data, but in which literal meaning is computed. This model has been implemented as part of a unification-based natural language processing system, called LINK.

Introduction

How are metaphors and other nonliteral constructions understood? A point of contention with regard to this question has been whether or not it is necessary to compute the literal meaning of a metaphorical utterance en route to understanding its intended (nonliteral) meaning. A large body of research in psycholinguistics has revealed that, in appropriate contexts, metaphors and other nonliteral constructions are comprehended no slower than comparable literal constructions, and sometimes even faster. This evidence, along with other related studies, has led some researchers to conclude that computation of literal meaning is not a necessary step in the understanding of metaphors.

In this paper, we suggest that the existing psychological data does not provide the necessary evidence to determine whether or not literal meaning is computed during the comprehension of metaphors and other nonliteral expressions. Moreover, from a computational standpoint, there are reasons to prefer an approach in which literal meaning is always computed. This motivates our own model of metaphor processing, in which literal meaning is computed. We present our approach, and show that our model is consistent with the existing psychological data.

An important feature of our model is that the nonliteral interpretation of a potentially metaphorical expression is automatically computed from the expression's literal meaning, irregardless of the context in which it appears. We argue that the existing data cannot distinguish between this processing model and one in which literal meaning is not computed.

Our approach to processing metaphors has been implemented in a natural language processing (NLP) system called LINK. LINK is an integrated, unification-based system, in which syntactic, semantic, and domain knowledge are all implemented in unification constraint rules. In order to process metaphors, we have added a set of mapping rules, which construct alternative (metaphorical) interpretations from literal meanings constructed during parsing. Context then determines which of the literal or metaphorical interpretations of the expression is most appropriate.

This paper is organized as follows: first, we present an argument for why, from a computational standpoint, there is a reason for computing the literal meaning of a metaphor and for the use of mapping rules. Next, we examine previous computational models which have proposed the use of mapping rules, and the psychological evidence which seems to refute these theories. Finally, we present our model, and show why it is consistent with the psychological data.

The need for mapping rules

Can the meanings of metaphorical expressions be computed directly, or is it necessary to compute their literal meaning first? Consider the following example:

The stock market rose today.

What would it mean to understand this metaphor directly? Assuming the traditional compositional approach to semantic interpretation, it would mean that the intended meaning of the sentence must be constructed from the meanings of its pieces (either lexical items or phrases). However, if we combine the literal meanings of "stock market" and "rose," we obviously will not arrive at the correct interpretation of this sentence. We could define one or both of these terms as
ambiguous, with "literal" and "metaphorical" senses, but it is not clear that we can easily separate the meaning of the metaphor into components to be associated with individual lexical items or phrases, as the compositional approach requires. For example, "rose" could be defined ambiguously, as either referring to an increase in altitude or an increase in a numeric indicator, but this does not capture the generalization that almost any word which refers to a change in altitude can be used in the same way. Consider the following examples:

The stock market plummeted today.
My fever has gone up.
Computer science enrollments are leveling off.
The 10% jump in property taxes this year was outrageous.

If we defined "rose" as ambiguous, we would also have to define every other word which refers to a change in altitude as ambiguous, in order to account for the productivity of this metaphor, not a plausible solution.¹

Metaphors such as the above example, which are tied to general concepts rather than specific lexical items, are not uncommon. We will not attempt to argue this point here, as many others have presented convincing arguments as to the prevalence of highly productive metaphors and other nonliteral constructions (e.g., Lakoff and Johnson, 1980; Langacker, 1987). Given their prevalence, it is important that we properly account for them, with a mechanism which explains their productivity.

How can we write appropriately general interpretation rules for these kinds of metaphors? Since they are often syntactically and lexically flexible, it seems that the only way to do so is at the semantic level. Instances of a given metaphor share commonalities in "literal" meaning (e.g., all lexical items that can be used in the "rose" metaphor normally refer to a change in altitude) and also in "intended" meaning (e.g., when used metaphorically, words like "rose" refer to a change in value of a numerical indicator). Thus, it seems natural to write an interpretation rule that transforms the literal meaning of a construction to an alternate nonliteral one.

We will refer to this kind of rule as a mapping rule. In the case of the "rose" metaphor, the mapping is from a change in altitude to a change in value of a numerical indicator. The direction of the change in value is "the same" as the direction of the change in altitude: that is, increase in altitude means increase in the numerical indicator.

How can mapping rules be used in semantic interpretation? Comprehension of a nonliteral construction using a mapping rule would involve constructing the literal interpretation of the construction, identifying the appropriate mapping rule, and applying the rule to the literal interpretation, thereby yielding the intended (nonliteral) interpretation.

The Standard Pragmatic Model
Many previous theories of metaphor comprehension have proposed the use of mapping rules to account for metaphor comprehension (e.g., Indurtyka, 1987; Carbonell, 1982; Searle, 1978). In fact, this approach has been called the Standard Pragmatic Model (Gerrig, 1989). Gerrig outlined three general steps in the Standard Model:

1. Construct the literal meaning of the utterance
2. Identify a failed semantic or pragmatic constraint in the literal meaning
3. Apply a mapping rule to transform the literal meaning to the intended meaning.

This model makes two predictions about processing. First, since steps 2 and 3 must be performed after completion of step 1, it predicts that comprehension of metaphors is slower than comprehension of comparable literal utterances. Second, since step 3 above is only executed if "normal" semantic and pragmatic processing fails to interpret an utterance, it predicts that the metaphorical interpretation of an utterance is not computed in a context in which the literal interpretation is acceptable.

Numerous psychological studies contradict both of these predictions. First, many studies have shown that, in an appropriate context, metaphor comprehension is as fast as comprehension of comparable literal expressions. In (Ortony et al., 1978) subjects read one or more sentences which set up a context, then read the target sentence, and then indicated as quickly as possible that they understood the target by pressing a key. There were two variations on this basic design. In experiment 1, the context was either short or long. In the short context, subjects took longer to understand nonliteral than literal targets, but in long contexts the reading times were not significantly different. In experiment 2, the targets were phrases that could have either a literal or idiomatic interpretation, depending on context. It was found that targets with an idiomatic interpretation took no longer to process (and may be processed faster) than the literal interpretations of the same targets.

Gibbs (1979) reported similar contextual effects on the comprehension of indirect requests. Subjects read stories, one line at a time, which ended in either an indirect request (e.g., "Must you open the window?") or a direct request (e.g., "Please open the window"). In an appropriate context (i.e., a story in which the context set up the expectation for a request), indirect
requests took no longer to comprehend than the direct requests.

More recent experiments have on the whole confirmed that metaphors take no longer to process than literal statements in appropriate contexts. See Gerrig (1989) or Gibbs (1984) for a more extensive discussion of the evidence.

Gibbs has also examined the question of whether or not literal meanings of indirect requests are activated during comprehension. In (Gibbs, 1989), he found no evidence for activation of literal meaning if an indirect request was comprehended in an appropriate context. Immediately after reading an indirect request such as “Can you pass the salt?” subjects were asked to make sentence/nonsentence judgments on target sentences. When the target was a paraphrase of the literal meaning of the indirect request (e.g., a target about ability to pass the salt), there was no facilitation of this judgment task from the indirect request preceding it. However, in the reverse situation, facilitation did occur: subjects were faster at making sentence/nonsentence judgments on paraphrases of the indirect request meaning even if the request was meant literally. Gibbs concluded from this asymmetry of priming that literal meaning of indirect requests was not computed, but that the nonliteral meaning always was.

Glucksberg, Gildea, & Bookin (1982) addressed the question of when literal and nonliteral meanings or metaphors are computed. Subjects were asked to make judgments about the literal truth of sentences of the form “All/Some X are Y.” Subjects took significantly longer to judge the truth of sentences which had a reasonable metaphorical meaning than those which did not. That is, judgment about the literal truth of sentences such as “All men are wolves” was slower than for “All men are telephones.” Since the task only required subjects to make judgments based on literal meaning, the results suggest that metaphor comprehension is an automatic process; i.e., metaphorical interpretations are computed even when literal meaning is intended.

The evidence against the Standard Model has led some to conclude that literal meaning is not computed during comprehension of nonliteral constructions. Perhaps the strongest proponent of this view is Gibbs, who goes so far as to question the validity of the notion of literal meaning (Gibbs, 1984). However, as we will argue in the next section, it is not the construction of literal meaning and the use of mapping rules that causes the Standard Model to make these incorrect predictions, it is the way in which mapping rules are applied in the model.

Metaphor Processing in LINK

We now present our alternative theory of metaphor processing, which is implemented in the LINK system. While our model also utilizes mapping rules, it differs from the Standard Pragmatic Model in that the rules are applied automatically, rather than in a failure-driven manner. As we will see later, automatic application of mapping rules leads to correct predictions about reading times and activations.

In LINK, metaphor processing follows these general steps:

1. Incrementally construct the (literal) interpretation of constituents of the utterance.
2. As they are constructed, try to match the interpretations against mapping rules. If any rules match, also construct the result of the mapping (the metaphorical interpretation) as an alternative interpretation of the constituent.
3. Use contextual information to determine which of the candidate interpretations is (are) preferred.

This model is a refinement of Martin’s (1990) theory of metaphor processing. It differs from Martin’s approach in three ways: first, the construction of metaphorical interpretations in LINK is an incremental process, which proceeds during the comprehension of a sentence. In Martin’s system, all metaphorical interpretations were constructed after completion of parsing. Second, mapping rules in LINK are used to process other nonliteral constructions, such as metonymies and figures of speech. Finally, because metaphor processing is integrated with parsing, LINK’s mapping rules can contain a mixture of syntactic and semantic information. This is important in the processing of other types of nonliteral constructions, such as metonymies and figures of speech.

Before we can explain the details of LINK’s mapping rules, we must give a general overview of LINK.

LINK’s unification grammar

All syntactic, semantic, and pragmatic knowledge is encoded in LINK in unification constraint rules. These rules are very similar in form to other unification grammars, such as PATR-II (Shieber, 1986). Each constraint rule consists of a set of equations, each of which constrains the interpretation which the parser can build in some way, by limiting, for a class of nodes (i.e., any node with a particular label), the set of arcs that can lead from a node of that class, as well as the types of nodes that arcs can lead to. Here are two simplified examples of constraint rules:

\[
S: \quad (1) = NP \\
(2) = VP \\
(head) = (2 head) \\
(head subj) = (1 head) \\
V: \quad (1) = ate \\
(head rep) = EAT FOOD \\
(head subj rep) = (head rep actor) \\
(head dobj rep) = (head rep object) \\
\]

\[2\text{In Martin’s theory, the initial representation constructed was called the “primal representation” of the utterance. We view this as equivalent to literal meaning, although it is not clear that Martin would agree.}\]
Each equation in the first rule (eqs. 1-4) specifies a property which any node labeled S must have. Equations constrain properties that nodes may have by specifying the label of the node to be found at the end of a path, or sequence of arcs (equations 1-2); or by specifying that two paths must lead to the identical node (equations 3-4). Identity here is defined by the unification operation. Unification merges the properties of two nodes; thus, two paths can unify if their values have no properties which explicitly contradict each other.

Functionally, the above rule encodes information about English sentences as follows. Equations 1 and 2 specify that a sentence is made up of two subconstituents: NP and VP, in that order. Equation 3 assigns the HEAD of the sentence to be the same as the HEAD of the VP, by unifying the VP's HEAD path with the HEAD path of the S. Finally, equation 4 assigns the NP to be the subject of the sentence.

The HEAD property is used to bundle information together, so that it can be shared across rules. Because of equation 3, any information on the HEAD of the VP is accessible from the S node. Similar equations would assign the head of the verb (V) to be the HEAD of the VP, and a particular lexical item to be the HEAD of the V.

The second rule (eqs. 5-8) is an example of a lexical entry. These rules typically provide the values which are propagated by HEAD links. Equation 6 specifies the semantic representation of the word “ate.” By convention, the semantic representation of a constituent appears as the value of the (head rep) path. Equations 7-8 specify mappings from syntactic to semantic dependencies. Whatever constituent fills the SUBJ role in the sentence will also be assigned as the ACTOR of the EAT-FOOD, and the syntactic direct object (DOBJ) will be assigned as the semantic OBJECT. Thus, in conjunction with equation 4, the sentence “John ate the apple” is interpreted to mean that “John” is the ACTOR of the action EAT-FOOD.

Equations 7 and 8 are used in conjunction with the system’s domain knowledge, to impose restrictions on the semantic properties (i.e., the values of the REP path) of the subject and direct object of “ate” (i.e., the ACTOR and OBJECT of EAT-FOOD). Domain knowledge is also encoded in constraint rules. In this particular case, the relevant rule is the following:

EAT FOOD: (actor) = HUMAN \hspace{1cm} [9]
(object) = FOOD \hspace{1cm} [10]
(instrument) = UTENSIL \hspace{1cm} [11]

Because of the mapping provided by “ate” between its subject and the ACTOR of EAT-FOOD, the restriction that this constituent’s representation must be HUMAN is propagated up to the NP which fills the SUBJ role specified by equation 5. Similarly, the FOOD restriction on the object of an EAT-FOOD would propagate to the NP assigned as the direct object (DOBJ) of “ate”, because of equations 11 and 13.

Semantic or domain knowledge is often used to eliminate inappropriate interpretations of ambiguous lexical items or constructions. For example, in the sentence “John ate the chips,” the POKER-CHIP sense of “chips” would be eliminated, since it is not a type of FOOD. This mechanism for eliminating contextually inappropriate interpretations is important in the selection of literal or metaphorical interpretations of potentially metaphorical expressions.

**LINK’s mapping rules**

The definition of a mapping rule in LINK is shown below. A $<$label$>$ is any syntactic or semantic category used in the grammar. A $<$var$>$ (variable) by convention begins with a “?”, and the same set of variables appear in the left and right hand sides of a rule.

\[
\text{<mapping-rule>} \ ::= \ <\text{spec}> \Rightarrow <\text{spec}>
\]

\[
\text{<spec>} \ ::= \ <\text{label}> \text{<eqn>} \ldots \text{<eqn>}
\]

\[
\text{<eqn>} \ ::= \ <\text{path}> = <\text{label}> | \<\text{path}> = <\text{var}> | <\text{path}> = (<\text{var}> <\text{label}>)| <\text{path}>
\]

Mapping rules are interpreted to mean the following: whenever the parser builds a node with the appropriate label, and that node explicitly satisfies all the constraints specified in the left-hand-side constraint list, then an alternate interpretation can be built; namely, a node satisfying the description on the right hand side. Variables indicate mappings between the values of properties of the original node and properties of the alternate node.

LINK is implemented as a chart parser. Thus, alternative interpretations of a constituent can co-exist, as competing links in the chart. Selection of an interpretation is in effect performed when a complete parse can be found which uses one of the competing links.

To illustrate, let us consider three different examples of mapping rules. First, here is a mapping rule for the highly productive CHANGE-ALTITUDE is CHANGE-VALUE metaphor:

\[
\Delta\text{-ALTITUDE}: \Rightarrow \Delta\text{-VALUE}:
\]

\[
(\text{change}) = ?X \quad (\text{change}) = ?X \quad (\text{object}) = \text{NUMBER}
\]

The simplicity of this rule reflects the productivity of the metaphor. Rather than just providing an ambiguous definition for one word, such as “rose”, it automatically creates an alternate interpretation whenever the semantic content of any word or group of words is constructed that fits the LHS. Thus sentences as disparate as “the temperature fell” and “The population of NYC is slowly creeping upward” are handled by the same rule. Additional semantic constraints in other rules would ensure selection of the appropriate interpretation, depending on the semantic category of the
1. "the plane is rising" would be interpreted metaphorically, while "the temperature is rising" would be interpreted literally, while "the temperature is rising" would be interpreted metaphorically.

Mapping rules in LINK are used to process other nonliteral constructions, such as metonymies and figures of speech. Here is a rule for a common metonymy used in conjunction with the CHANGE-ALTITUDE is CHANGE-VALUE metaphor:

\[ \text{PHYS-OBJ: } \Rightarrow (?)X \]
\[ (\text{assoc-num}) = (?X \text{ NUMBER}) \]

Many physical objects are used to refer to closely associated numbers, e.g., the stock market for the Dow Jones Industrial Average, or the thermostat for its temperature setting. This rule means that "stock market" can be interpreted as either the building or the DJIA.

Finally, here is a rule for the idiom "to go through the roof", meaning to increase (in height or some other dimension) rapidly:

\[ \text{VP: } \Rightarrow \text{VP:} \]
\[ (\text{head rep}) = \text{PIERCER} \]
\[ (\text{head obj} \text{ rep}) = \text{ROOF} \]
\[ (\text{head rep change}) = \text{\(\Delta\)-ALTITUDE} \]
\[ (\text{head rep change rate}) = \text{FAST} \]

In this example, as with many idioms, none of the literal constructs of the LHS are relevant to the nonliteral interpretation, so there is no need for variables. This rule also demonstrates that the syntactic content of the construct may be used as easily as the semantic. It is also possible for rules to refer to exact words and their order.

Now consider the following example:

The stock market went through the roof.

In parsing this sentence, LINK creates literal and nonliteral interpretations (links on the chart) for both phrases ("the stock market" and "went through the roof") before trying to combine them to form a complete sentence. Given the semantic constraint that the stock market building is IMMOVABLE, and thus cannot be a (semantic) object of verbs requiring motion, then only the triply nonliteral interpretation will unify to create a complete interpretation. If we make no such restriction, then there will be three S-nodes created, and the sentence is truly ambiguous. This example also demonstrates the flexibility of the first mapping rule, as alternate word definitions would need to be quite convoluted to arrive at the CHANGE-VALUE in this sentence.

LINK's model and the psychological data

Now that we have described LINK's model of metaphor processing in detail, let us return to the psychological data discussed earlier to see if our model is compatible with this data. Summarizing the experiments discussed earlier, there are three main findings:

1. In the appropriate context, reading times for metaphors are no slower than reading times for literal constructions.

2. Common metaphorical meanings appear to be computed even when expressions are intended to be taken literally.

3. Literal meanings are not primed after comprehension of a sentence containing a nonliteral construction.

Recall that all three of these findings contradict the Standard Pragmatic Model. But do they necessarily imply that literal meaning is not computed during metaphor comprehension, as some have suggested? We would argue that the first two findings are problematic for the Standard Model not because it states that literal meaning is computed first, but rather because of its assertion that metaphor processing is failure-driven. According to the Standard Model a metaphorical interpretation is not constructed unless some difficulty is encountered in constructing a literal interpretation. Waiting for a semantic (or pragmatic) constraint violation before invoking a mapping rule results in the incorrect prediction of additional processing time for metaphors, as well as the incorrect prediction that metaphorical meanings will not be computed during comprehension of literal utterances. Thus, it seems to be step 2 of the Standard Model that results in the incorrect predictions.

Our theory differs from the Standard Model in several ways. First, in our theory the rules are applied automatically, regardless of whether or not semantic or pragmatic failures are encountered while constructing literal meaning. Second, in our theory mapping rules never eliminate possible interpretations, they instead add other interpretations as alternatives. This differs from the Standard Model, in which literal meaning is transformed into (i.e. replaced by) a metaphorical interpretation.

Because of these differences, our model can also account for findings 1 and 2 above. First, literal expressions and metaphors will take the same amount of time to compute, because metaphorical interpretations are always computed, even in contexts in which the literal meaning is the one which is intended. Second, the fact that metaphorical interpretation is an automatic process means that even in literal contexts, metaphorical meaning is constructed. This accounts for the findings of (Glucksberg et al., 1982) and (Gibbs, 1983), in which evidence was found that nonliteral interpretations were constructed in literal contexts.

We still have to explain the third finding, namely that literal meaning is not primed after comprehension of a sentence containing a metaphorical expression. This would seem to suggest that literal meaning has not been computed, which would contradict our theory. How can we account for this discrepancy? To answer this question, consider the situation in which an ambiguous word is encountered in an utterance. Swinney (1979) and Tanenhaus et al. (1979) have shown that all senses of an ambiguous word (even if it is syntactically ambiguous) are activated briefly (between 200 and 600 msecs), regardless of context, when the word is encountered, after which contextual informa-
tion suppresses inappropriate senses. We envision that this same selection process is used in metaphor comprehension. After a mapping rule has been applied, the situation is analogous to one in which all senses of an ambiguous word have been activated. As possible alternative interpretations are computed, contextual information constrains which interpretations are acceptable in that context. Thus, by the end of a sentence containing a metaphor, it is quite likely that the literal sense of the metaphorical expression has already been suppressed by contextual constraints, just as alternate meanings of an ambiguous word are suppressed by context. This suggests that Gibbs' (1983) results indicating that literal meanings were not primed at the end of a sentence may have been due to the target appearing after suppression of the literal meaning had already taken place.

Conclusion

The psychological evidence that we have cited above clearly indicates that metaphorical and other nonliteral constructions are processed as quickly as comparable literal constructions, given an appropriate context. This evidence has shown the Standard Pragmatic Model to be an inadequate model of human metaphor processing. Since the Standard Model stipulates that the first step in understanding an utterance is to compute its literal meaning, and since the Standard Pragmatic Model's failure to explain the psychological data is directly tied to this first step in understanding, the role of literal meaning in understanding has come into question. Many researchers, Gibbs among them, have claimed that the evidence shows that literal meaning need not be computed on the way to computing metaphorical meaning.

We have argued that the problem with the Standard Pragmatic Model lies not in the fact that it computes literal meaning, but in the process by which literal meaning is computed and used to compute intended meaning. We have presented an alternative model of metaphor comprehension which computes literal meaning but which is consistent with the existing psychological data.

Although we feel that there are strong computational motivations for preferring an approach such as ours that uses mapping rules, it appears that the existing empirical evidence is not adequate to distinguish between Gibbs' model, in which literal meaning is not computed, and our alternate model. The data are inadequate because, in general, they only tell us about processing at the sentence level, rather than the word or phrase level. The reading time comparisons which we have discussed have been done on entire sentences rather than on individual words or phrases, and priming experiments have tested facilitation only at the end of sentences. Thus, existing results say little about the process of metaphor comprehension at the word or phrase level. It is at this level that we could distinguish between the competing theories.

References


3We still need to explain Gibbs' finding that nonliteral meanings were still primed at the end of a sentence which was interpreted literally. It is possible that the asymmetry of these results could be explained by frequency effects: if the indirect request sense of a construction occurs much more frequently than the literal sense, this might overpower the ability of context to suppress this meaning.