Metaphor Understanding – A Cognitive Science Perspective

Raluca Vasilescu
Computer Science Department, CMU

September 2, 1997

Abstract

This paper surveys some of the important trends of the research on metaphor comprehension in the cognitive science area. We examine the semantic theories of the metaphor, some pros and (mostly) cons of the very challenged error-recovery model of metaphor comprehension. We present Lakoff's view that metaphors are conceptual rather than linguistic phenomena; we discuss how theories of other tropes are consistent with metaphor theories; and finally we talk about some attempts to develop a computational theory of the metaphor.

1 Motivation

A legitimate question the reader of this paper might ask is: why metaphors? Aren't they the exclusive appanage of poets and writers? Is this a study about the psychology of people involved in comprehending literary works?

The answer to these both questions is "not necessarily." Indeed, every day language is full of dead or novel metaphors. Metaphors are one of the most valuable sources of new words for a language; and the all English (or other language) speakers make use of metaphors on a regular base.

Moreover, metaphors are not only an ubiquitous phenomenon in the natural language, but researchers claim that they actually influence our way of thinking. There is a famous paper of Reddy [Reddy 1979] that shows how our beliefs about language and communication are a consequence of the conduit metaphor which is often used by the English speakers when they talk about language. Lakoff [Lakoff 1992, Lakoff et al. 1980] extends the class of "cultural metaphors" and supports this idea with lots of evidence.

It is often the case that metaphors help us grasp abstract concepts by expressing them in terms of more concrete notions. A lot of the words denoting abstract nouns were initially metaphors.

The main questions we are trying to answer in this paper are: how do people understand metaphors? Is there any special process that makes metaphor understanding different than literal word understanding? Is a new meaning created when a metaphor is understood? What are our chances to develop a computational theory of the metaphor and which constraints should it satisfy? What makes a good metaphor good?

Equally interesting questions are: why do people use metaphors? What makes them choose a metaphor instead of one or more literal words? Or maybe there is no equivalent literal translation of the metaphor? Although episodically we might touch this second class of questions in our survey, they are not our main focus, partly because the literature on these topics is less rich.

2 The Structure of the Metaphor

A few words should be said about the terminology of this paper. We shall use the term **metaphor** to refer to a figure of speech (or trope) that expresses a domain by making reference to another (unrelated) domain. Examples are *Juliet is the sun*, *The ship ploughed the sea*, *Time is money* etc. The **topic (tenor, target)** of the metaphor is the domain we are trying to explain (i.e *Juliet, the ship, time* in the previous examples). The **vehicle (source)** is that to which we make reference (i.e. *the sun, plow, money*). The **ground** is made of the common properties shared by both the source and the target and which are emphasized by the metaphor.

3 Semantic Theories of the Metaphor

The semantic theories of metaphor seek for answers to questions like: what are metaphors, how are they understood, how literal meaning is discerned from the figurative meaning.

3.1 Comparison Theories

The comparison theories regard metaphor as a non-explicit comparison between the target and the source. They are based on the classical view on metaphors whose roots can be traced back to Aristotle.

Some of the comparison theories consider objects represented as sets of features.

Tversky's contrast theory [Tversky 1977, Glucksberg et al. 1990] regards a comparison of the type *a is like b* as concerning only some relevant features of the two objects. Since the numbers of features potentially attributable to an object is infinite, Tversky postulates a prior extraction and compilation of the features of a and b which are relevant to a given context. He defines the following measure for the similarity between a and b:

$$s(a, b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$$

where θ, α and β are some weights, A, B are the sets of features of a, respectively b and f is some scale function. Since in most metaphorical comparisons only some common features of the target and the source are relevant, the other particular features of each of the terms being non-essential, we can expect to have $\alpha = \beta = 0$. However, this assumption implies symmetry for the similarity relation. Symmetry is not a property of the metaphors, as you can see from the following example. Indeed, *Some jobs are jails* can be a meaningful metaphor, but *Some jails are jobs* is not (at least in an obvious way). To account for the asymmetry of the metaphors, Tversky assumes that the comparison is directional and therefore the features of the target are weighted more heavily than the features of the source (i.e. $\alpha > \beta$). The directionality comes from a difference in salience between the features of the target and those of the source (the features of the target being more salient).

People seem to be able to distinguish good metaphors from bad ones. For instance, they think that *Her cheeks were red like a rose* is less metaphorical (or a worse metaphor) than *Her cheeks were like a rose*. The problem with the contrast theory is that it doesn't explain for the difference in the degrees of metaphoricity between.

Ortony [Ortony 1979a] overcomes such difficulties by modifying Tversky's theory to include the notion of **salience imbalance**. Ortony considers that the ground of the comparison (or metaphor) is indeed made by some feature(s) shared by both the source and the target, but hypothesizes that the common feature is more salient in the source than in the target. This assumption confirms our intuition that metaphor (like any other utterance) should have a high informative value. If the common feature were very salient in the definition of target, then the metaphor would seem redundant or senseless (for instance, saying *"roses are like cherries"* to transmit the information *"roses are red"* doesn't make much sense since the color red is a salient property of roses). The ratio between the salience of the shared feature in b and the salience of the same feature in a for a statement *a is like b* defines the degree of metaphoricity of that statement.

The problem with Ortony's theory is that every comparison (literal or metaphoric) should obey the salience imbalance law in order to be informative. Therefore, the theory has no means to distinguish between literal and metaphoric comparisons.

Another difficulty (common to an entire class of comparison theories known as matching models) is that many metaphoric comparisons seem to involve properties that are not a priori present in the listener's representation of the target. The metaphoric utterance emphasizes new properties of that representation; therefore a metaphoric comparison cannot involve some matching between the properties of the source and those of the target. Indeed, as Ortony himself points out, one can distinguish between attribute-promoting metaphors and attribute-introducing metaphors. The attribute promoting happens when the metaphor increases the salience of an existent feature of the topic (like in *Sermons are like sleeping pills*). The attribute introducing occurs when the hearer is informed of a new, unknown property of the topic (like in *Roger is a tiger in our board meetings*).

3.2 Interaction Theory

The interaction theory is due to Black [Black 1979] (with origins in the work of I. A. Richards. Black argues that any criterion for "metaphorhood" (i.e. for distinguishing metaphors from other forms of language) is doomed to failure. He believes that some metaphors help us to understand new aspects of the reality; and that these aspects are created by the metaphor itself.

Black uses the terms "primary subject" for topic and "secondary subject" for vehicle. Each of them should be seen as an "implicative complex", made by beliefs and properties commonly held to be true. Then

> The metaphorical utterance works by "projecting upon" the primary subject a set of "associated implications" comprised in the implicative complex, that are predicable of the secondary subject ... [Black 1979]

The metaphorical utterance actually reorganizes and emphasizes the features of the primary subject by "applying to it statements isomorphic with the members of the secondary subject's implicative complex." [Black 1979]. It is central for the interaction view to remark that such a projection depends strongly on the primary subject, as well as on the secondary subject. The primary subject determines which features of the secondary subject are to be projected by the metaphoric statement. The metaphor induces a change in the secondary subject, too. Indeed, to follow one of Max Black's examples, in the metaphor *Man is a wolf*, the interaction between the two subjects produces no only an "animalization" of men, but also a "antropomorphization" of wolves. This is due to selecting from the "commonplaces" related to wolves those that can apply to men and in this way impoverishing the implicative concept of wolves.

A final quote from Black will allow us to introduce the class inclusion theory of the metaphors, due to Glucksberg and Keysar [Glucksberg et al. 1990].

> It would be more illuminating in some of these cases (i.e. of metaphors imputing similarities difficult to discern otherwise) to say that the metaphor creates the similarity than to say that it formulates some similarity antecedently existing. [Black 1962], cf. [Black 1979].

We can locate the attempt to regard metaphors as class inclusion statements [Glucksberg et al. 1990] in same the line as Black's interaction theory. According to these authors, the source of the metaphor stands for the prototype of an ad-hoc created superordinate category that includes both the source and the target and whose features these terms of the metaphor determined. This is in agreement with Black's claim that a metaphor builds a new structure (in which both source and target play equal parts), inexprimable otherwise. The practice of naming categories by their prototype is met in a lot of cultures (for instance, the word for "eagle" is used by Shoshoni speakers to refer to large birds in general [Glucksberg et al. 1990]). The centrality of the prototype chosen as vehicle would dictate the degree of metaphoricity of a metaphor and also its comprehensibility.

The context, together with the topic, plays a very important part in the definition of the new superordinate category. It may determine which features should be included in its representation. Indeed, it has been shown that the similarity of the objects is modified by the context in which they are put.

In this light, Glucksberg and Keysar argue that any similarity (or comparison) should be actually regarded as a class inclusion statement, and that we should explain the comparison as a metaphor rather than viceversa. The categorization is the one that produces the similarity.

3.3 The Incoherence (3-Stage, Error-Recovery) Theory

The incoherence view is due to Searle [Searle 1979]. Since it is easier to be psychologically tested than the other semantic theories, it has actually been contested by a wide range of experiments in the 80s. We shall talk about some of these experiments in section 4.

Searle assumes that, when a metaphoric utterance is encountered, the hearer first tries to find a literal meaning for this utterance, and, only failing to do that, he/she looks for a metaphorical meaning. The appropriateness of the literal meaning is dictated by the context. The recognition of a metaphor is a 3-step process: first, the literal interpretation of the sentence is built; second, this interpretation is matched against the context; third, if no consistent match can be found, the hearer decides for a metaphorical interpretation.

This view implies that the metaphoric meaning is deviant: it needs a special process of understanding. It also assesses the priority of literal meaning over the nonliteral meaning; metaphoric utterances thus require a richer and highly specified context in order to be understood.

The process of metaphor understanding is also (like the process of detecting a metaphor) a 3-step process. First, the hearer must decide for a metaphorical interpretation. Second, he/she must have some strategies for deriving the possible features of the topic emphasized by the metaphor. And third, he must have a set of principles for restricting the range of the features discovered in the second step. Searle asserts that the vehicle plays the essential part in the second step and the topic is the one that is used to make the proper selection in the third step. The participation of both the topic and the vehicle to the metaphorical meaning makes Searle's theory resemblant to the interaction theory.

4 Is The "Incoherence View" a Plausible Psychological Assumption?

Since Searle's paper has been published in 1979, a lot of psycholinguistic experiments have been carried to refute his hypothesis. Most of them rely on the following corollary of Searle's view: if the metaphoric processing were initiated only when after the literal meaning is found defective, then metaphorical utterances would take longer to understand. This last statement seems not to hold for humans.

Indeed, even in 1979, D. E. Rumelhart claimed that the distinction between the literal and the figurative meaning is not reflected in a quantitative change in the psychological process

[Rumelhart 1979]. His principal argument comes from looking at how children acquire and use metaphoric language. It seems that this process parallels the one of literal language acquisition; no special difficulties are encountered. Rumelhart emphasizes that figurative language understanding, like literal language understanding, is dependent on general knowledge of the world rather than on the acquisition of some special skill.

This claim is repeated in the Ortony's paper [Ortony 1978]. Ortony built an experiment in which subjects were presented metaphorical sentences preceded by either a context encouraging a literal interpretation or a context encouraging a metaphorical one. He varied also the length of the context. His experimental data showed that, in cases where there is a lot of contextual support, metaphors are processed as quickly and easy as literals. However, in cases where little contextual support had been provided, metaphors take longer to understand that literal sentences. Ortony concludes that metaphor comprehension involves the same sort of process as literal sentence comprehension. Starting from this type of result, Inhoff [Inhoff et al. 1984] speculated that the context encourages a processing strategy, not a given meaning. According to this theory, there are processing strategies, not separate processing stages.

Another consequence of the 3-stage model is explored in a set of experiments conducted by Glucksberg and his colleagues [Glucksberg et al. 1982]. The 3-stage model implies that the literal meaning of a sentence is always computed, while the metaphorical meaning is "optional." Glucksberg argues that it seems difficult to inhibit our metaphoric understanding of simple utterances as *Some hearts are closets*, independent on how rich or appropriate the context is. His claim is supported by the experimental data provided by his experiments. Indeed, the data shows that whenever a metaphorical interpretation is available, people take longer to judge the literal meaning of a sentence (out of any context) as false. Therefore, it seems reasonable to infer that neither the literal nor the figurative meaning have unconditional priority one over another.

The studies that support the 3-stage theory are not very many and for most of them the design of the experiment seems questionable. One more recent study, however, is particularly interesting [Gibbs 1990]. It shows that for cases where the topic is not explicitly mentioned in a sentence, but it is referred exclusively by the vehicle (i.e. for metaphoric referential descriptions; e.g. "The pig didn't want to lend him money."), the understanding takes longer than for similar literal sentences (containing the topic itself). The experiment has been done with sentences presented in metaphor-supporting contexts.

Metaphoric referential description are very often encountered in every day language and they represent one of the main sources of language enrichment. The question one may ask in such situation is: is there a new meaning created for the metaphoric description (i.e the vehicle)? If yes, when exactly is this meaning created? Is an error-recovery model (similar to Searle's) appropriate? Do people first try to retrieve a literal meaning for the word, and, after failing to find one, create a new meaning for it?

Gibbs' data supports the idea that no sense creation actually takes place. There exists some kind of reference to the antecedent (i.e. to the topic) referred by the metaphoric description. The process is similar with translating the metaphoric reference into a literal one, and then understanding the sentence. Gibbs argues that an association between the vehicle and the topic is created and that it is this association that makes possible the understanding of the sentence.

In the same line of research devoted to metaphoric referential description, Gerrig [Gerrig 1989a] takes the alternative position: he considers that there is no evidence for supporting the claim of an error-recovery based understanding of the metaphoric referential descriptions. He claims that only when the context is unavailable people attempt error recovery; if context is available, then the process of sense selection and the process of creation of a new meaning (for the vehicle) are concurrent. Although the time course and the components of the process of meaning creation are far from being fully specified, Gerrig has shown that people are reliably good at forming new interpretations of conventional expressions.

5 Metonimy, Irony, Idioms and Indirect Speech

Like metaphors, idioms, proverbs and indirect speech acts are frequent nonliteral phenomena of natural language. Consider, for instance the following common instance of indirect speech: *Can you pass the salt?*. The hearer readily understands that this utterance is not a question about one's ability, but rather a polite request. The same happens with idioms: people seem to ignore the literal meanings of *He kicked the bucket*, *He spilled the beans* or *It rains with cats and dogs*. Metonimy is sometimes considered a special kind of metaphor: an object is expressed by making reference to some related domain (the relation can be part-to-whole – and the figure is then called a *synecdoche* – like in *We need a new glove in our team*, or other – like in *The White House refused to talk with the rebels*).

We won't dwell too much on metonimy. Some of the experiments we have discussed about in section 4 were carried for metonimies too (e.g. [Gibbs 1990]). In general, the data for metonimies parallel the data for metaphors. There is no evidence that processing of this trope is different from the processing of the metaphor.

Searle tries to account for irony and indirect speech in the same way he accounts for metaphor [Searle 1979]. He assumes he who utters an ironical sentence means the opposite of what the literal interpretation of that sentence would say. For instance saying to someone who just broke a precious vase *That was a brilliant thing to do* means actually saying to him/her *That was a stupid thing to do*. Thus, when an ironical sentence is encountered, people look first for a literal meaning. If no literal interpretation adequate for the context can be found, they negate the sentence and, in this way they reach the interpretation of the ironical sentence.

Searle claims that in the case of indirect speech acts the speaker means what he/she says and even more. The hearer first detects the nonliteral meaning of the sentence (for instance, by realizing that a question about one's ability to pass the salt lacks any conversational point). Then the hearer uses its general knowledge about speech acts to infer that what the speaker actually means is to have the salt passed.

However, as in the case of metaphor, Searle's hypotheses have not found much experimental ground. Indeed, it has been shown that indirect sentences that fit the context (e.g. *Can you close the window?* in a context where the window is open) do not take longer to understand than literal requests [Gibbs 1979]. If the context is not supporting a nonliteral interpretation (i.e the window is closed, for instance), the comprehension is the longest. If no context is presented at all, indirect requests are longer to understand than literal ones. It appears that people are biased to comprehend the nonliteral meaning of indirect speech, if it occurs in appropriate context. It is not the case that both the literal and nonliteral meanings are retrieved in parallel [Gibbs 1983].

As for idioms, it seems there is a bias to perceive their nonliteral meanings. The reason might be the existence of a special entry in our mental lexicon for each idiom. Such a claim is supported by the difficulty second language learners experience in learning idioms: most of the time they must give up a logical, translatable and inferable interpretation and learn them by heart. Ortony's 1978 influential experiment for metaphors was done for idiomatic expressions, too [Ortony 1978]. Subjects took longer to understand the literal meaning of an idiom when preceded by a context that supported this meaning (e.g. *Let the cat out of the bag* in a context about a cat and a bag) than to understand the nonliteral meaning when preceded by a context encouraging it (e.g. *Let the cat out of the bag* in a context about a surprise party).

To conclude, research on idioms and indirect requests has shown there is no evidence for the hypothesis that people retrieve the literal meaning of such utterances. The literal meaning may not be processed at all during comprehension

6 Metaphors As Conceptual Phenomena

Starting from an idea introduced by Reddy's 1979 paper [Reddy 1979], Lakoff [Lakoff 1992, Lakoff et al. 1980] claims that metaphors happen not at the language level, but rather at the thought level – "they are general mappings across conceptual domains." A metaphorical expression is a "surface realization" of a cross-domain mapping. Lakoff uses small caps to denote such a (metalinguistic) mapping. He gathers an notable number of such cross-domain mappings in his books (e.g. MORE IS UP, LESS IS DOWN; EVENTS ARE ACTIONS, PURPOSES ARE DESTINATIONS, PEOPLE ARE PLANTS etc.).

Lakoff distinguishes a hierarchy of conceptual mappings: some of them are instantiations of more general ones (e.g. A PURPOSEFUL LIFE IS A JOURNEY is an instantiation of PURPOSES ARE DESTINATIONS; DEATH IS DEPARTURE is an instantiation of STATES ARE LOCATIONS etc.).

Lakoff claims that part of the cognitive apparatus shared by members of the same culture is a very large set of basic conceptual metaphors. These metaphors are systematic mappings from one domain to another (that is the structure of the source domain is mapped onto the structure of the target domain). Evidence for this hypothesis comes mainly from polysemy and transfers of inference patterns from one domain to another. Such metaphorical mappings are largely unconscious and widely conventionalized in language (there is a large number of idioms whose interpretation relies on these metaphors – e.g. *The prices rose* as a realization of MORE IS UP etc.). They obey the Invariance Principle: "The image-schema structure of the source domain is projected onto the target domain in a way that is consistent with the inherent target domain structure" [Lakoff 1992]. The metaphorical mappings originate in the every day knowledge and experience. They vary in universality: some are met in most cultures, other are rarer.

Although conceptual metaphors are pervasive, Lakoff argues that this does not imply that everything is metaphorical: "Metaphorical understanding is grounded in nonmetaphorical understanding." [Lakoff 1992]. For instance, our knowledge of the three dimensional space is nonmetaphorical. This concrete conception about space helps us represent and understand the more abstract concept of time (e.g. the past is behind us, the future in front etc.). In this light, metaphors can be regarded as a cognitive instrument.

There are several problems with Lakoff's theory on metaphor. Some of them are discussed in [Cacciari et al. 1992]. First, there is the question of novel metaphors – Lakoff argues that they are either reinstantiations and elaborations of the conceptual ones or they are image mappings (for instance, a visual image from one domain is moved to another, based on our knowledge about both domains – like in Rabelais' *His toes were like the keyboard of a spinet*). Metaphors like the one from a newspaper headline referring to a flood in Chicago: *City of Broad Shoulders Has Wet Feet* [Cacciari et al. 1992] is hard to explain. Also, it is not clear how the conceptual metaphoric system is acquired initially or how one realization of a metaphor is chosen over another. Why would people prefer *My job is a jail* to *My job is a traffic jam*? [Cacciari et al. 1992].

7 Computational Models

The computational models of metaphor understanding seem to have evolved from the need of the natural language processing systems to integrate the ubiquitous phenomenon of metaphor. That is why most of such systems don't observe the psycholinguistic results we have discussed in the previous sections. Therefore, rather than models of the metaphor understanding by humans, they offer ways of dealing with this problem in computer based systems. Their merit for the cognitive science shouldn't be underestimated, since they provide a good starting point for true computational models of the metaphor.

Before discussing these models (we will continue to call them so in lack of a better term), we shall make a quick overview of the constraints a true computational model of metaphor understanding should satisfy as they were stated by Gerrig [Gerrig 1989b].

As we have seen in section 4, there is strong evidence that the process of metaphor understanding is not an error-recovery one: people do not seem to try metaphoric understanding of an utterance only when its literal meaning is defective. This result has been reflected in the similar processing times for metaphorical and literal understanding. Therefore, "the consequence for the authors of computational models of metaphor is that they are not licensed to hypothesize processes that would require elongated comprehension times for metaphors" [Gerrig 1989b] (*the total time constraint*). As Gerrig points out, this constraint does not imply that the two processes of comprehension are identical – they might be different but still take the same total time.

Second, a computational model of the metaphor should account for the comprehension of all types of metaphors, be they conventional, every day language metaphors or sophisticated poetic metaphors. It is possible however that two different theories might explain the time-limited comprehension (of conventional, every day metaphors) and the leisurely comprehension (of particularly elegant poetic metaphors). "What becomes critical in developing a computational theory is for the author to make a claim about which of these two types of comprehension are being modelled" [Gerrig 1989b].

The attempts to model computationally the metaphor understanding can be split in three classes: the knowledge intensive approaches, which make use of separate lexicon entries for words in the conventional metaphors ([Hirst 1987, Wilenky et al. 1980] cf. [Martin 1992]), the analogy based approaches [Carbonell 1982, Gentner et al. 1988, Indurkhya 1987], which try to use analogy mappings to comprehend metaphors, and the mixed approaches which make use of knowledge of conventional metaphors and try to extend this knowledge to novel metaphors [Martin 1992].

We shall overview some of the analogy-based methods and a mixed approach due to Martin.

7.1 Analogy-Based Models

Carbonell proposes a model of metaphor understanding starting from Lakoff's idea of a basic system of conceptual metaphors. He assumes a set of such metaphors exist in the memory of a computer system, together with the tools to recognize them and with mappings from the source to the target domain. The process of understanding follows the 3-stage model: first the system tries to retrieve a literal meaning for an utterance; if it cannot find any semantically consistent literal meaning, it checks to see whether any of the existent representation structures for conventional metaphors recognizes the input utterance. Once a metaphor is detected, the system uses its previous knowledge about the metaphor to map the components of the source domain into the components of the target domain. It also uses a "locality principle" – it remembers the current instantiation of the metaphor, considering it likely to reoccur in the present context.

What happens when the system encounters a new metaphor, i.e. there is no already built-in structure to recognize and analyze this metaphor? One way of learning metaphors and the corresponding mappings from source to target domains is to extract these mappings from the interpretation of several (previously-met) analogies. Such a procedure is supported by Carbonell's claim that "the computational process used in understanding analogies is the same as that used in understanding metaphors. The difference is one of recognition and universality of acceptance of the underlying mappings" [Carbonell 1982]. That is to say that the metaphor is an analogy which uses a non-explicit mapping of the two domains (because such a mapping is either known or obvious).

Carbonell also discusses the comprehension of new metaphors [Carbonell 1982, Carbonell 1981]. He considers (in the Lakoff's line) that metaphors are a powerful source of learning. Carbonell investigates the laws that allow us to transfer methods of inference from one domain to another. He postulates the existence of a "invariance hierarchy" of mapping criteria; the items high in this hierarchy are more likely to be transfered from one domain to another. Thus, goals, planning strategies, causal structures, functional attributes and temporal orderings are (in this order) most likely to be kept invariant by the metaphorical mapping. Social roles, structural relations, physical attributes or object identity are items least likely to be preserved, and therefore they occupy low positions in the invariance hierarchy. A system that comprehends new metaphor should first look for mappings at levels higher in the hierarchy.

Gentner also exploits the similarity between metaphor and analogy. Her structure mapping theory [Gentner et al. 1988, Gentner 1983] for analogy claims that, in making an analogy, the correspondence between target and source domain is made on the basis of high order relation matching rather than of particular object description matching. This is somehow similar to Carbonell's invariance hierarchy: the concepts more likely to be transfered in an analogy are relations rather than attributes. Gentner demonstrated how the same analogy analysis works for metaphors using a system called Structure Mapping Engine (SME). SME starts finding local matchings between single items in the source and target domains, based on relations with the same names in the two domains. Then it orders this matches according to some degree of matching (some matches may be more complex than others if, for instance, several parts of the two domains may match) and collects them into global matches. The largest of these matches is considered as a possible interpretation of the analogy (metaphor).

Indurkhya formalizes mathematically a constrained semantic transference theory starting from Gentner's structure mapping ideas [Indurkhya 1987]. He shows how it is possible to associate the T-MAPS from his theory with one metaphorical interpretation of a set of sentences. The T-MAPS are basically mathematical mappings that preserve the structure of the source domain onto the target domain. They are able to capture some important properties of metaphor: the fact that they highlight certain aspects of the two domains and downplay others, their asymmetry, their power to transfer inferences from one domain to another etc. He also gives some algorithms for generating the T-MAPS attached to a metaphorical interpretation.

7.2 Mixed Models

Martin's MIDAS system [Martin 1992] is an attempt to fuse the advantages of the knowledge base approach and of the analogy based approach. Lakoff's claim of a basic set of conventional metaphors leads Martin to build a knowledge base of such metaphors to be used by MIDAS. However only "prototypes" of the metaphors are kept by the system; when new realizations of a conventional metaphor are encountered, MIDAS is able to recognize them.

The knowledge representation in MIDAS is based on semantic nets. A special kind of relation denotes a metaphoric association between the target and the source domains of the metaphor. These domains are represented at the conceptual and not at the lexical level.

The interpretation process builds all consistent interpretations for a metaphoric sentence. That is how it is able to satisfy Gerrig's time constraint. The interpretation in MIDAS uses two mechanisms: concretion and metaphoric unviewing. Concretion replaces an abstract concept by a more specific one. It helps in recognizing metaphors that are realizations of conventional, more general metaphors, already known to the system (for instance, interpreting the metaphor *John has a cold* as an instance of the more general metaphor BEING IN A STATE IS HAVING). Metaphoric unviewing replaces the source concept in a metaphor with the corresponding (assumedly existent) target concept.

If no consistent representation can be found, MIDAS tries to extend one of the existing metaphors; this is done based on a process of analogy not between the source and the target of the metaphor, but between an old metaphor and the new one.

If two consistent interpretation can be found, a context analysis must be pursued in order to decide which of the two (or more) is correct.

7.3 Metaphor Banks

Both Carbonell's and Martin's approaches could make use of a metaphor bank – i.e. of a knowledge base containing most of the conventional metaphors. This bank should constitute the corpus of "cultural metaphors" (and their instantiations). Actually, there is an attempt to build such a bank, due to Lakoff himself and to his students (Berkeley Metaphor List). The entries are compiled by hand from the published metaphor literature and from a series of graduate seminars held at University of California at Berkeley. They contain the metaphorical mapping and metaphorical realizations for each metaphor.

Martin [Martin 1994] proposes "an empirically derived and theoretically motivated knowledge base of English metaphorical conventions." In his paper, he shows how such a knowledge base could be actually built with metaphors described in the Ortolingua representation language. He proposes two methods for gathering information: one based on a directed search for known metaphors, another based on random sampling of texts.

The directed search method might start from the Berkeley Metaphor List (or from some other list of metaphor). It should consider some of the most frequent words met in the realization of a certain metaphor and then search for such words in large amounts of text. For each hit, a human expert should judge if it is a realization of the metaphor or not. In this way, it is possible to get information about the frequency of a certain metaphor and about its universality (in which domains it is usually applied). Also, it is possible to find other conventional metaphors not current in the data base, frequency information about these other metaphors or about the probe word with other meanings etc.

The random sampling method randomly extracts sentences from a text; a human observer must judge them as metaphorical or not. The metaphors found in this way could be clustered in groups; some information about the relative frequency of each group, about the realizations possible within each group and about the nature of the target/source domain of this groups can be thus obtained.

8 Some Final Words

We have seen that, while there is no general agreement for one semantic theory of metaphor, psycholinguists seem to have reached the consensus that metaphoric and literal understanding are similar processes, and that metaphor understanding is not triggered by some defective literal meaning. This hypothesis is supported by phenomena related to other figures (like idioms, indirect speech or metonimy) commonly met in every day language. The research directed towards a computational theory of metaphor is not yet perfectly coordinated with the psycholinguistic results, but still the interesting ideas that have evolved in this area can be fruitfully used in artificial intelligence systems.

This survey has no pretense of exhaustivity; it just tries to scan what the author believes that are the most important trends in the research on metaphor understanding.

References

[Black 1962] Black M., 1962. Metaphor. In M. Black, *Models and Metaphors*. Cornell University Press.

[Black 1979] Black M., 1979. More About Metaphor. In A. Ortony (ed.), *Metaphor and Thought*. Cambridge University Press.

[Cacciari et al. 1992] Cacciari C., Glucksberg S., 1992. Understanding Figurative Language. In *Handbook of Psycholinguistics*.

[Carbonell 1981] Carbonell J., 1981. Invariance Hierarchies in Metaphor Interpretation. In *Proceedings of the Third Meeting of the Cognitive science Society*. Cognitive Science Society.

[Carbonell 1982] Carbonell J., 1982. Metaphor: an Inescapable Phenomenon in Natural Language Comprehension. In W. Lehnert, M. Ringle (eds.). *Strategies for Natural Language Processing*. Laurence Erlbaum.

[Cohen 1979] Cohen J., 1979. The Semantics of Metaphor. In A. Ortony (ed.).*Metaphor and Thought*. Cambridge University Press.

[Gentner 1983] Gentner D., 1983. Structure Mapping: A Theoretical framework For Analogy. *Cognitive Science, 7*, pg. 155-170.

[Gentner et al. 1988] Gentner D., Fakenhainer B., Skorstad J., 1988. Viewing Metaphor As Analogy: The Good, The Bad and The Ugly. In D. Helman (ed.). *Analogical Reasoning*. Kluwer Academic Publishers.

[Gerrig 1989a] Gerrig R., 1989. The Time Course of Sense Creation. *Memory and Cognition, 17(2)*, pg. 197-207.

[Gerrig 1989b] Gerrig R., 1989. Empirical Constraints on Computational Theories of Metaphor: Comments on Indurkhya. *Cognitive Science, 13*, pg. 235-241.

[Gibbs 1979] Gibbs R., 1979. Contextual Effects in Understanding Indirect Requests. *Discourse Processes, 2*, pg. 1-10.

[Gibbs 1990] Gibbs R., 1990. Comprehending Figurative Referential Descriptions. *Journal of Experimental Psychology: learning, Memory and Cognition, 16*, pg. 56-66.

[Gibbs 1983] Gibbs R., 1983. Do People Always Comprehend The Literal Meaning of Indirect Requests? *Journal of Experimental Psychology: Learning, Memory and Cognition, 9*, pg. 524-533.

[Gibbs 1984] Gibbs R., 1984. Literal Meaning and Psychological theory.*Cognitive Science 8*, pg. 275-304.

[Gibbs 1993] Gibbs R., 1993. Process and products in making sense of tropes. In A. Ortony (ed).*Metaphor and thought*, Cambridge University Press.

[Glucksberg et al. 1990] Glucksberg S., Keysar B., 1990.Understanding Metaphorical Comparisons: Beyond Similarity. *Psychological Review, 97*, pg. 3-18.

[Glucksberg et al. 1982] Glucksberg S., Glidea P., Bookin H., 1982. On Understanding Literal Speech: Can People Ignore Metaphors? *Journal of Verbal Learning an Verbal Behavior, 21*, pg. 85-98.

[Hirst 1987] Hirst G., 1987. Semantic Interpretation and the Resolution of Ambiguity. Cambridege University Press.

[Hoffman et al. 1987] Hoffman R., Kemper S., 1987. What Could Reaction-Time Studies Be telling us About Metaphor?*Metaphor and Symbolic Activity, 2(3)*, pg. 187-199.

[Indurkhya 1987] Indurkhya B., 1987. Approximate Semantic Transference: A Computational Theory of Metaphors and Analogies. *Cognitive Science, 11*, pg. 445-480.

[Inhoff et al. 1984] Inhoff A., Lima S., Carroll P., 1984. Contextual Effects on Metaphor Comprehension and Reading. *Memory and Cognition, 12*, pg. 558-567.

[Lakoff 1992] Lakoff G., 1992. The Contemporary Theory Of Metaphor. In A. Ortony (ed.), 1992. *Metaphor and Thought.* Cambridge University Press.

[Lakoff et al. 1980] Lakoff G., Johnson M., 1980. Metaphors We Live By. Chicago University Press.

[Martin 1992] Martin J., 1992. Computer Understanding of Conventional Metaphoric Language. *Cognitive Science, 16*, pg. 233-270.

[Martin 1994] Martin J., 1992. Metabank: a Knowledge Base of Metaphoric Language Conventions. *Computational Intelligence, 10*, pg. 134-149.

[Ortony 1979a] Ortony A., 1979. Beyond Literal Similarity. *Psychological Review, 86*, pg. 161-180.

[Ortony 1978] Ortony A., Schallert D., Reynolds R., Antos S., 1978. Interpreting Metaphors and Idioms: some Effects on Comprehension. In *Journal of Verbal Learning and Verbal Behavior, 17.* pg. 465-477.

[Ortony 1979b] Ortony A., 1979. The role of similarity in similes and metaphors. In A. Ortony (ed.). *Metaphor and thought.* Cambridge University Press.

[Ortony 1993] Ortony A., 1993. Metaphors, Language and Thought. In A. Ortony (ed.). *Metaphor and thought.* Cambridge University Press.

[Reddy 1979] Reddy M., 1979. The conduit Metaphor: A Case Of Frame Conflict In Our Language About Language. In A. Ortony (ed.), 1992. *Metaphor and Thought.* Cambridge University Press.

[Rumelhart 1979] Rumelhart D.E., 1979. Some Problems With The Notion Of Literal Meaning. In A. Ortony (ed). *Metaphor and Thought.* Cambridge University Press.

[Searle 1979] Searle J. R., 1979. Metaphor. In A. Ortony (ed.),*Metaphor and Thought.* Cambridge University Press.

[Tversky 1977] Tversky A., 1977. Features of Similarity. *Psychological Review, 84*, pg. 327-352.

[Wilenky et al. 1980] Wilensky R., Arens Y., 1980. Phran – A Knowledge-Based Approach To Natural Language Analysis. (Tech. Rep. UCB/ERL/M80/34). University of California, Berkeley.