A corpus-based analysis of context effects on metaphor comprehension

James H. Martin

Abstract

This article describes our attempts to shed light on the relationship between results from psycholinguistic research on the effects of context on metaphor comprehension and the nature of metaphor as it occurs in naturally occurring text.

The hypothesis underlying this work is that the facilitation and inhibition effects observed in laboratory subjects reflect the patterns of co-occurrence of various kinds of contexts with metaphoric language in the environment.

We take a three part approach to exploring this hypothesis: a rational analysis of the notion of a context effect, an empirical corpus-based effort to fill out that rational analysis, and a reconsideration of the pertinent psycholinguistic results with respect to that analysis. Finally, we present a proposal for a mechanistic model that is in accord with the results of this analysis.

1. Introduction

A wide variety of results from psycholinguistic research over the last several decades have shown that context has a strong effect on the processing of metaphoric language (Gernsbacher et al. 2001, Gerrig and Healy 1983, Gibbs 1984, Gildea and Glucksberg 1983, Glucksberg 1982, Kemper 1989, Inhoff et al. 1984, Keysar 1989, Ortony et al. 1978). These results have been used to both support and refute a bewildering array of computational accounts of metaphor processing (Fass 1991, Fass 1988, Martin 1990, Martin 1992, Martin 1994, Gentner et al. 1988, Gildea and Glucksberg 1983, Russell 1976, Wilks 1978, Hobbs 1979, Carbonell 1981, Indurkhya 1987, Narayanan 1999). Perhaps the most well-known result from this research is that appropriate contexts facilitate the processing of metaphor to the extent that there is no significant timing difference from equivalent literal language. These results has been primarily used to argue against the stage model (Searle 1979) of metaphor processing where the literal meaning of an utterance is first computed, found to be lacking, and then reanalyzed as a metaphor.

Lost in this debate is the fact that a considerable amount of information has been amassed showing how specific *kinds of context* facilitate, fail to

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 207/228

facilitate, or actually inhibit the process of metaphor comprehension. These results provide constraints on computational models of metaphor above and beyond basic timing constraints (Gerrig 1989). Specifically, if the presence or absence of a particular type of information has predictable effects on metaphor comprehension, then computational theories of metaphor must show how they can transparently admit such influences.

A second issue that has received little attention is the exact relationship between the psycholinguistic effects observed in the laboratory and the environment of naturally occurring text. Of particular interest is the question of whether these effects are artifacts of a specialized metaphor processing mechanism, or whether they are simply reflections of the way that metaphor naturally occurs in real text. If the latter situation holds then that provides further evidence that the capacity to interpret metaphor is similar to, if not the same as, our ability to process other kinds of language.

The bulk of this article is an attempt to shed further light on these two issues. We are interested in studying the nature of these contextual influences by taking a detailed look out at the environment. More specifically, we are concerned with the relationship between observed patterns of co-occurrence of context and metaphor in naturally occurring text and known psycholinguistic results on these kinds of contexts.

The hypothesis underlying this work is that the facilitation and inhibition effects observed in laboratory subjects reflect the patterns of co-occurrence of these contexts with metaphoric language in the environment. We take a three part approach to exploring this hypothesis: a rational analysis of the notion of a context effect, an empirical corpus-based effort to provide real data for that rational analysis, and a reconsideration of the pertinent psycholinguistic results with respect to the analysis.

2. A Minimal Rational Analysis of Context Effects on Metaphor

We start by taking a step back to consider the effect of context on metaphor comprehension from a rather minimalist rational point of view. By this, we have in mind an account derived from minimal assumptions about cognitive processing, combined with a detailed corpus-based observation of the phenomenon in question. Under this view, the first step in analyzing some phenomenon is to analyze the task and make certain minimal assumptions about the information needed to perform it. The second step is to look to the environment to see how such information is present-

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 208/228

ed in situations where the task presumably has to be addressed. The final step is to juxtapose measures of human performance on appropriate tasks with the analysis of the environment. This can suggest likely hypotheses for computational models and impose constraints on such models.

The analysis presented here is loosely inspired by Anderson's theory of the adaptive nature of human cognition (Anderson 1990, Anderson and Schooler 1991, Oaksford and Chater 1998). The basic notion in this approach is that an efficient language processor can be seen as having molded itself to the regularities in its environment. When faced with a particular language problem, the processor takes action based on an implicit or explicit encoding of these regularities. Such an approach makes the testable claim that observable regularities in the environment will lead to predictions about processor performance under similar conditions.

We begin by defining the task as comprehending a sentence containing a metaphor after having already processed a short span of text. The preceding span of text and the ensuing metaphor will be referred to as the *context* and the *test sentence*, respectively. By comprehension, we have in mind the simple notion commonly used in various chronometric psycholinguistic studies. In these studies, a subject typically reads some context and is then presented with a test sentence. They are told to perform some simple physical task when they feel that they have adequately understood the sentence. In these experiments, the subject is typically not asked to study and deeply appreciate the test sentence. Rather, they are being asked to process the text in a normal automatic fashion. The time taken to perform the physical recognition task is measured and compared across the various conditions of interest.

Having sketched out the task, the next step is to determine the various sources of information needed to perform it. While there is wide disagreement about the mechanisms underlying metaphor comprehension, nearly all current theories postulate that successful metaphor comprehension results in, or involves at some stage, representations involving a *source* concept, a *target* concept and a set of correspondences or associations between them, often referred to as a *ground*.

To make these notions concrete consider the following example extracted from the Wall Street Journal.

(1) Spain Fund *tumbled* 23% in *turbulent* trading, *dragging down* the shares of other so-called country funds.

This example has three instantiations of the NUMERICAL-VALUE-AS-LOCATION metaphor (Hobbs 1979). In this metaphor, the core source concept is

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 209/228

the notion of a location in physical space, often an altitude. The target domain is the notion of some abstract state that has an associated numerical value. The ground, or set of structured correspondences, in this case stipulates that the numerical value is to be viewed as a physical location and that a change in value is to be viewed as motion along some dimension.

In our current study, we make the minimal assumption that language that serves to introduce the source or target concepts directly into the context may have an effect on the processing of subsequent metaphors involving these concepts. A further assumption is that text that introduces a metaphor of the same basic type as the test metaphor into the context can be seen as introducing the ground of the metaphor.

Having sketched out the task and its potential information needs, the next step is to analyze the ways that the environment actually presents this information to readers in naturally occurring texts. Specifically, we need to be able to locate and identify specific metaphors in text and to locate their corresponding source, target and metaphorical uses in contexts immediately preceding the metaphor. These various contexts can then be analyzed in terms of how well they *predict the occurrence of the subsequent metaphor*. Specifically, we are interested in whether particular contexts can be seen as making a subsequent metaphor more or less likely. Of course, to be able to assess the predictive value of these contexts, information about the overall frequency of occurrence of these contextual cues and the metaphors themselves needs to be gathered.

The comparison to human performance is based on data from metaphor comprehension experiments measuring the time needed to perform various experimental tasks that shed light on comprehension difficulty. Such experiments have shown that the rate at which sentences are comprehended is affected by the context within which the sentences are presented. Three types of contexts are of interest here:

– Contexts containing literal expressions of the source concepts of the test metaphor
– Contexts containing relevant literal expressions of the target concepts of the test metaphor
– And finally, contexts containing metaphoric expressions with the same basic structure, or ground, as the test metaphor.

The rational account presented here is ultimately based on the juxtaposition of this human performance data with the data gleaned from the environment. The specific hypothesis is that it is the degree to which the various types of context *predict* the future occurrence of a metaphor that determines the degree of facilitation observed in the human perfor-

Verlag: Walter de Gruyter | **P-Nr.:** B12-421394 | **Pfad/Job:** W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | **P-Anfang:** 07.09.2005 **P-Aktuell:** 23.09.2005 | **ID:** int01; int04 – AGB4 – 09.05 | **Printjob:** 23.09., 07:36, Seiten: 210/228

mance data. We will refer to this as the *Metaphor Prediction Hypothesis* (MPH).

3. Corpus Analysis

There have been large number empirical studies of naturally occurring metaphor in both spoken and written forms. However, none of these studies provide the specific contextual data needed to test the validity of the MPH. Fortunately, various factors gleaned from these studies point towards a way to gather the appropriate information.

The first result comes from exhaustive analyses that have been performed to determine the overall frequency of metaphor and metonymy in a wide variety of texts (Pollio et al. 1990). These studies show that these phenomena appear quite frequently, averaging around 5 uses per 100 words of text. The second result comes from a wide variety of research that indicates that there is a relatively small core set of important conceptual metaphors underlying most of the metaphors actually observed (Talmy 1975, Talmy 1988, Lakoff and Johnson 1980, Lakoff 1987, Lakoff and Turner 1988, Johnson 1987).

Taken together, these considerations indicate that an analysis of a relatively small number of sentences randomly sampled from a coherent corpus would provide useful information. Specifically, by sampling a relatively small amount of text from a coherent collection of text it should be possible to produce an accurate characterization of the important conceptual metaphors that occur in a given collection. In such an analysis metaphors in the sampled text are identified by hand and clustered together based on the conceptual similarity of their source, target and ground components. Such an analysis has been performed on a random sample of sentences from the Wall Street Journal (Martin 1994).

The results from this study provided us with the basis to collect the specific contextual information we require; the collected examples of the more frequently occurring metaphor types give us the means to search for further examples of those specific types *in context*.

3.1. Methodology

The need to gather information specific to particular types of metaphor led us to focus on a small number of conceptual metaphors known to frequent the Wall Street Journal corpus. Based on the sampling and cluster-

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 211/228

ing work already completed, it was possible to develop simple and accurate lexical profiles of some of these metaphors. These profiles made it possible to filter large amounts of text for the occurrence of specific metaphor types.

To be concrete, consider the NUMERICAL-VALUE-AS-LOCATION metaphor discussed above. It is the most frequently occurring metaphor in the Wall Street Journal and is used to express a wide range of economic and commercial concepts. By examining the sampled metaphors of this type, we were able to produce two lists of words that provided a lexical profile of this metaphor. The first contained words used to express the source concept of location or change of location. The second list contained words used to express the target concepts of the sampled metaphors. In this case, the source list included words such as *fall, tumble, sink, downhill, slide, drop, top, climb, rise, boost,* and *plunge*. The target list included *recession, inflation, bid, rate, borrow, priced, earnings, trading, price, pay, costs, income,* and *earnings.*

Taken together these two lists provide an extremely crude, but effective, way to find metaphors of a particular type; any sentence containing words from both the source and target list for a particular metaphor type is a candidate for containing the metaphor used to produce the lists. Used separately they can be used to find candidate sentences containing expressions of either the source or target domains of particular metaphors.

Such lists were produced for four metaphors, were known to occur with considerable regularity in the WSJ corpus: NUMERICAL-VALUE-AS-LOCATION, COMMERCIAL-ACTIVITY-AS-CONTAINER, COMMERCIAL-ACTIVITY-AS-PATH-FOLLOWING and COMMERICIAL-ACTIVITY-AS-WAR. Representative examples of each of these metaphors from the WSJ are shown in Table 1.

To broaden the reach of these lists, they were each augmented through the use of a thesaurus, introspection, and further sampling. Next a simple program was written to find example sentences of each type; given a source list, a target list, and a metaphor name, this program tags all sentences in a corpus containing at least one term from each list. An iterative process was used refine the lists to remove obvious and frequent false positives.

Using this simple approach, it was possible to quickly filter large amounts text from the Wall Street Journal at the sentence level for the metaphors of interest. For the purposes of this study approximately 600 instances of each of the four metaphor types shown in Table 1 were identified. Due to the differing relative frequencies of the metaphors this necessitated examining widely different amounts of text for each metaphor. This ranged from

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 212/228

130,000 words for the NUMERICAL-VALUE-AS-LOCATION metaphor to nearly 2,000,000 for the COMMERCIAL-ACTIVITY-AS-WAR metaphor.

Of course, while this simple technique is extremely effective in this text type it is by no means perfect. Therefore, each sentence tagged was verified by hand to make certain that it actually contained the correct metaphor. It is also highly likely that the approach will miss some metaphors of each type. We believe that this situation is acceptable in the current study as long as we are sure that it will find a large percentage of the metaphors that do occur. Fortunately, one of the findings from our earlier work (Martin 1994) is that each individual metaphor type displays a Zipf-like distribution for the lexical items that are used to express that metaphor. In other words, while these metaphors are wildly productive, a relatively small number of lexical items account for the bulk of metaphor instances.

Table 1. Sample metaphors known to frequent the WSJ

NUMERICAL-VALUE-AS-LOCATION
Barge rates on the Mississippi River *sank* yesterday on speculation that widespread rain this week in the Midwest might temporarily alleviate the situation.
At the same time, an increase of land under cultivation after the drought has *boosted* production of corn, soybeans and other commodities, causing a *fall* in prices that has been only partly *cushioned* by heavy grain buying by the Soviets
COMMERCIAL-ACTIVITY-AS-CONTAINER
Four Brazilian fruit-juice makers are planning to enter the Japanese market and to build a huge juice-storage tank complex, a Japanese trading-house official said.
The situation is that the bankruptcy court will *get out of* the shipbuilding business.
COMMERCIAL-ACTIVITY-AS-PATH-FOLLOWING
So *where* does the IMF *go from here*?
Swissair, which signed a marketing agreement with Delta Air Lines early this year, *took the next step* this summer by buying Delta.
COMMERCIAL-ACTIVITY-AS-WAR
While the *two camps* have competed *aggressively* for years, lately the *fight has turned ferocious* – and often damaging to the public.
What triggered the latest *clash* was a *skirmish* over the timing of a New Zealand government bond issue.

Finally, this corpus was tagged with the source and target lists separately to identify all expressions of these concepts in context. The resulting corpus thus contains sentences tagged as containing expressions of the source concepts, target concepts, or metaphors, for each of the four metaphors embedded in their original contexts. This tagged corpus then

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 213/228

served as the basis for obtaining the predictive measures needed to test the MPH.

3.2. Results Bearing on the MPH

The following sections detail the results of these analyses for each of the three context types, for each of the four metaphors. Specifically, they present the following information.

- Base frequency rates for each of the metaphors.
- Predictiveness of contexts containing literal expressions of the source concept of subsequent test metaphors.
- Predictiveness of contexts containing expressions of concepts that serve as targets for subsequent metaphors.
- Predictiveness of contexts containing a metaphor for subsequent metaphors of the same type.

At this point in the analysis, the notion of context has to be made more concrete. The term context has been used in a wide variety of ways in the psycholinguistics literature. It has been used to refer to everything from clauses preceding a metaphor within a sentence, to multiple paragraph-length preceding texts. Based on an informal examination of this literature, we decided to focus on five sentence context windows. Mapping this directly to the psycholinguistic literature we have contexts consisting of four sentences, followed by a sentence containing a metaphor. However, since we are interested in determining the predictive power of the various types of contexts it is more useful to think of single sentence contexts, followed by a four sentence window. In this framework, the ability of a given kind of context to predict a subsequent metaphor will be estimated as the conditional probability of seeing a metaphor in a subsequent four sentence context given a particular type of contextual cue.

3.2.1. Prior Probabilities

To assess the predictive power of the four types of context, the prior probability of finding an instance of one of the four metaphors in a random context had to be measured. These probabilities were estimated by considering the amount of text that had to be searched to gather the approximately 600 instances of each of the four metaphors. The base rate for each metaphor was estimated as the ratio of the number of metaphors found to sentences examined.

Rather than making any independence assumptions, the probability of one or more instances of these metaphors occurring in a random 4 sen-

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 214/228

tence sequence was determined directly by sliding a four sentence window across the tagged texts and counting the number of windows containing at least one instance of the metaphor. Again the probability measure is estimated as the ratio of 4 sentence windows containing a metaphor to the number of windows. Table 2 gives both the individual and 4 sentence window frequencies for each metaphor expressed as probabilities.

Table 2. Base rates for metaphors studied

Metaphor	P(Metaphor)	P(Metaphor in Window)
NUMERICAL-VALUE -AS-LOCATION	.104	.265
COMMERCIAL-ACTIVITY-AS-PATH	.016	.046
COMMERCIAL-ACTIVITY-AS-CONTAINER	.017	.044
COMMERICIAL-ACTIVITY-AS-WAR	.006	.011

It should be noted that the window results reflect a clear tendency on the part of these metaphors to cluster. Specifically, for each metaphor, the number of non-metaphor four sentence windows observed was higher than would be expected if the metaphors were uniformly distributed throughout the text. As a consequence the probability of running into one of these metaphors in a random four sentence span is less than would be expected if they were uniformly distributed across the texts.

These probabilities will be used as a baseline to measure the predictive power of the various contexts for ensuing metaphors. If the probability of seeing a metaphor in the next four sentences, given the current context type, is higher than these base rates, then that evidence should facilitate comprehension. Inhibition should be signaled by a lower than base rate probability given a particular context.

3.2.2. Literal Source Concept

Our first results concern the predictive power of contexts containing literal expressions of a source concept of the test metaphor. To be more concrete, we are interested in contexts like the following:

(2) Dirk was out climbing mountains all last week.
(3) Back at the office, his chances for promotion had plunged.

The first sentence contains a literal expression of the concept change of location in the vertical direction. The second sentence contains a metaphor where the source concept is from the same conceptual domain. The

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 215/228

statistic of interest is the conditional probability of seeing a metaphor in the next four sentences, given that a literal use of the source of that metaphor has already been seen.

To obtain these statistics, the tagged corpus was first examined by hand. Literal uses of the already tagged source terms were marked as such. The probability of encountering a metaphor given a literal source encounter was estimated by computing the ratio of the number of literal source uses followed by a metaphor in a subsequent four sentence context, to the number of literal uses found overall. Table 3 presents the results for each of our four metaphors.

Table 3. Predictive Power of Literal Source Contexts

Metaphor	P(Metaphor|Source)	Change from Base Rate
NUMERICAL-VALUE -AS-LOCATION	.069	.259
COMMERCIAL-ACTIVITY-AS-PATH	.022	.491
COMMERCIAL-ACTIVITY-AS-CONTAINER	.038	.844
COMMERCIAL-ACTIVITY-AS-WAR	< .004	< .418

These results indicate that literal concepts are poor predictors of specific future metaphors with that concept as a source. Each of the four metaphors occurs below its base frequency in contexts containing literal language from their source domain. In the case of the COMMERCIAL-ACTIVITY-AS-WAR metaphor, no contexts containing literal war expressions followed by this metaphor were found after examining all sentences containing literal war expressions extracted from approximately 2,000,000 words of text.

3.2.3. Target Concept Expressions

The next set of results concern how well expressions involving the target concept of a metaphor predict future expressions of a specific metaphor. Again, to be concrete, we are interested in contexts like the following.

(4) Dirk's opportunities for a promotion had been improving for several months.
(5) With his latest success they skyrocketed.

The first sentence introduces the notion of "opportunity for promotion". The second sentence follows up by metaphorically structuring that notion as a change of altitude. To formalize such situations we are interested in the conditional probability of seeing a specific metaphor given that some expression of the target concept has already been seen.

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 216/228

The following results were computed in a fashion similar to the previous literal source result. The ratio of target uses followed by a metaphor to the number of target uses found overall was computed by scanning the tagged corpus.

Table 4. Predictive power of target contexts

Metaphor	P(Metaphor|Target)	Change from Base Rate
NUMERICAL-VALUE -AS-LOCATION	.677	2.55
COMMERCIAL-ACTIVITY-AS-PATH	.073	1.59
COMMERCIAL-ACTIVITY-AS-CONTAINER	.087	1.95
COMMERCIAL-ACTIVITY-AS-WAR	.031	2.84

As shown in Table 4, these metaphors are more likely to occur in contexts containing expressions of their target concepts. As with the literal source result, the numbers varied from metaphor to metaphor but all displayed some effect of increasing the probability of the metaphor from its base rate.

3.2.4. Metaphoric Expressions

Our final result concerns the predictive power of a particular metaphor for subsequent instances of the same metaphor. The following example illustrates this situation.

(6) Dirk's opportunities for promotion had been *falling* for months.
(7) With his latest boondoggle, they really *plummeted*.

The first sentence contains a metaphor that structures the notion of changing opportunity as a change in altitude. The ensuing context follows up with an expression with the same basic metaphorical structure.

Table 5. Predictive power of metaphorical contexts

Metaphor	P(Metaphor|Metaphor)	Change from Base Rate
NUMERICAL-VALUE -AS-LOCATION	.703	2.65
COMMERCIAL-ACTIVITY-AS-PATH	.196	4.29
COMMERCIAL-ACTIVITY-AS-CONTAINER	.267	6.00
COMMERCIAL-ACTIVITY-AS-WAR	.277	25.20

As indicated in Table 5, this predictor is obviously the best from among those studied, with the clearest advantage in the case of the COMMERCIAL-ACTIVITY-AS-WAR metaphor.

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 217/228

4. Psycholinguistic Results

Our next task was to take these corpus-based results and juxtapose them with relevant data from the psycholinguistic literature. As might be expected, none of the relevant experiments presented here has results that directly correspond to our predictive measures. Nevertheless, there are results that can with some massaging be readily interpreted as relevant to our measures.

4.1. Literal Source Contexts

The primary result for the case of literal source contexts comes from a series of studies reported by Inhoff (1984). In these experiments, subjects were first presented with a single context sentence followed by a metaphor test sentence. The three types of context that were investigated correspond nicely to our contexts: a literal source context, a literal target context, and a closely related metaphor context. Reaction times across the three contexts were then compared to assess the relative effects of these contexts.

The major difference between these results and our corpus analysis is that their literal source context was constructed by employing a literal use of the *same word* as is used metaphorically in the test sentence. Our corpus results measured *any* literal use from the same conceptual field as the source of the metaphorical use.

The basic results of this study were that recognition time was shortest with metaphorical contexts, longer with relevant literal target contexts, and much longer still with literal source contexts. These results reflect the same pattern seen in our corpus based study. Specifically, the corpus indicates that when compared to target and metaphor contexts, literal source contexts are the least predictive.

This study was silent on the stronger prediction made by the MPH. The MPH predicts that since literal source texts reduce the likelihood of subsequent metaphors with the same source, they should have an *inhibitory* effect on recognition rates. Gernsbacher et al. (2001) report just such an effect. Contexts involving a literal use from the source domain of a subsequent metaphor suppress the reading time of the metaphor as compared to a baseline context prime.

4.2. Target Contexts

The most widely replicated result in the literature on context effects shows that appropriate contexts strongly facilitate the interpretation of

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 218/228

relevant metaphors. In these experiments, literal expressions of various concepts serve to facilitate subsequent metaphors involving those same concepts as targets (Ortony et al. 1978, Gerrig and Healy 1983, Inhoff et al. 1984, Kemper 1989). Note that as with the previous literal source result, what the facilitation is with respect to varies from study to study.

As already discussed, one of the results reported in Inhoff et al (1984) shows that target contexts facilitate metaphor comprehension as compared to literal source contexts. In a separate experiment, they also showed that target contexts facilitate subsequent metaphor comprehension as compared to unrelated contexts. The results from Kemper (1989) show a similar pattern for both the literal source and target conditions.

The work reported by Gerrig and Healy (1983) is somewhat different from these other context studies. Rather than manipulate prior sentential context, they manipulated the presentation of material within metaphorical sentences. Consider the following example:

(8) The train followed the parallel ribbons.
(9) The parallel ribbons were followed by the train.

In the active example, the word "train" provides a prior target context within which "parallel ribbons" is interpreted. In the passive example, no such context is available at the point that the referent to "ribbons" is introduced. Their results show that a context providing arrangement of clauses facilitates recognition over the null context clause arrangement.

To summarize, the pattern of results observed in these studies is consistent with the patterns of predictiveness observed in our corpus and is consistent with the MPH. Contexts containing expressions of target concepts of subsequent metaphors facilitate comprehension of those metaphors.

4.3. Metaphorical Contexts

Results concerning our final context type, metaphorical contexts, come from both the Inhoff et al (1984) and Kemper (1989) studies. In both cases, contexts were created containing instances of metaphors generated from constructs that correspond to the notion of conventional conceptual metaphors. Materials were prepared so that test sentences containing metaphors were preceded by contexts containing an instance of the same metaphor type.

In these studies, metaphorical contexts displayed the highest degree of facilitation. In both sets of experiments a pattern of facilitation was

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 219/228

shown that had source concepts with the least facilitation, followed by target concepts, followed by metaphors with the highest facilitation. As with the results on source and target contexts, these results are in accord with our corpus-based findings.

5. Discussion

This section examines the implications of our corpus-based results from three perspectives that are somewhat broader than the narrow focus taken thus far. We present a discourse oriented discussion of the results, followed by a review of the results along with other known constraints on metaphor processing, and finally a proposal for a mechanistic account that is consistent with these constraints.

5.1. Discourse

If one combines the view of metaphor advanced by Lakoff and Johnson (1980) with more discourse oriented notions of what makes a text coherent, then none of our corpus-based results are particularly surprising. Coherent text tends to be about some topic, or set of topics that display a high degree of semantic overlap and interconnectedness. This notion of coherence combined with the fact that the topics will be metaphorically structured in systematic ways can be used to account for all of our results.

First consider the literal-source expression in context situation. Our results show that these contexts predict that future metaphors with that source are fairly unlikely. The source and target domains of metaphors are by definition about different kinds or types of concepts. It is, therefore, not at all surprising to find that literal expressions of source concepts rarely co-occur with metaphors using that concept as a source. This follows since coherent texts don't typically mix completely disjoint topics within the kind of short spans we looked at in our study.

The results observed for contexts containing expressions of target concepts can be accounted for in a similar fashion. Again, if a coherent text is about some topic then it is likely that if that topic is even partially structured with some metaphor, then that metaphor will eventually occur in the context. Of course, conceptual domains differ both in the degree to which they metaphorically structured and in the number of distinct metaphors used to structure them. These factors will tend to mitigate the predictive power of target language for any particular metaphor.

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 220/228

Finally, consider the observed predictive power of metaphors for future metaphors of the same type. Given a context where a concept has already been introduced and metaphorically structured, one would expect repetitions of that metaphor as long as the discourse continued to focus on relevant aspects of that concept.

5.2. Empirical Constraints

The task of turning a high level discourse-oriented account into a plausible mechanism is a non-trivial one. The following section presents a sketch of one such proposal. Before considering that proposal, this section will review the various constraints that we now have for such a model. Among these constraints are ones that have been culled from the psycholinguistic literature and ones that follow the results of our corpus study:

- Total Time Constraint
- Non-Optionality Constraint
- On-Line Constraint
- Differential Behavior Constraint
- Contextual Influence Constraint

The *Total Time Constraint* (Gerrig 1989) states that when supported by appropriate context the time needed to process various kinds of non-literal language does not differ significantly from the time taken to interpret direct literal language. Minimally, this constraint argues that whatever mechanism is proposed for processing metaphoric language it can not have a markedly different time complexity from those mechanisms proposed for literal language processing. However, it does not by itself require that identical processing mechanisms be employed for metaphorical and literal language.

The *Non-Optionality Constraint* (Glucksberg et al. 1982, Keysar 1989) stems from research that shows that possible metaphorical interpretations are activated even in contexts where the literal meaning is both well-formed and preferred by those contexts. These results are based on demonstrations that show parallel metaphorical readings interfering with correct and plausible literal readings resulting in longer processing times on various tasks. These results parallel those in lexical access that show initial activation of all of a words senses followed by a rapid pruning based on local and global context.

The *On-Line Constraint* (Gerrig and Healy 1983) states that metaphorical interpretations, like other interpretations, are constructed in an on-line incremental fashion. Gerrig and Healy's research shows that the or-

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 221/228

der of presentation of material providing evidence for or against a metaphorical reading has strong effects on on-line processing. In effect, they show that subjects display a non-monotonic behavior, where the activation of particular interpretations is continuously updated based on available evidence. This update procedure may result in what they call *truncation,* where an interpretation that had been viable is eliminated when evidence becomes available. In the case of metaphor, they show that placing evidence for a metaphorical interpretation early in the sentence leads to faster reaction times. This arises from an earlier truncation of the parallel literal meaning. When disambiguating information is delayed to later in the sentence reaction times are longer because of interference from the competing interpretations.

It should be noted that these two constraints are really two sides of the same coin. They both point toward a model of processing where parallel interpretations are created and pruned based on currently available evidence. In this regard, they provide further evidence that metaphorical processing is subject to the same kind of constraints observed for literal language.

The *Contextual Influence Constraint* refines the total time constraint by specifying how various *kinds* of contexts can effect the processing time of subsequent metaphors. The basic pattern observed both in experimental settings and our corpus is that prior metaphors of the same type have the strongest effect on subsequent processing, followed by target concepts, and finally literal source concepts. The corpus-based results further predict an inhibitory effect on subsequent metaphor processing. Therefore, any mechanism that is proposed to account for metaphor processing must in some straightforward manner display this pattern.

Finally, the *Differential Behavior Constraint* stems from our results that show that metaphors, like lexical items, display differing frequency and recency patterns. Therefore, while various kinds of context do influence the time needed to process metaphor, the specific amount of facilitation varies with the specific type of metaphor. Like the other constraints, this one is consistent with findings from both the memory and lexical access literatures. Individual memory traces items display idiosyncratic patterns of access based on frequency and recency.

Taken together, these constraints argue for an approach to metaphor that is in large part based on the kind of generic evidential memory access and working memory constraints that have been independently proposed for processing ordinary language.

Verlag: Walter de Gruyter | **P-Nr.:** B12-421394 | **Pfad/Job:** W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | **P-Anfang:** 07.09.2005 **P-Aktuell:** 23.09.2005 | **ID:** int01; int04 – AGB4 – 09.05 | **Printjob:** 23.09., 07:36, Seiten: 222/228

5.3. A Construction Based Account

There are far too many mechanistic accounts of metaphor processing to assess them all with respect to our list of constraints. Rather than attempt such a survey, we will sketch an instantiation of one position based on work from current research on computational modeling on semantic interpretation and current work on Construction Grammar (Fillmore et al. 1988).

The account presented here is based on three interrelated research efforts: the notion of conventional conceptual metaphors as first articulated by Lakoff and Johnson (1980), direct computational implementations of this notion as in (Martin 1990), and finally more fine-grained, on-line, computational implementations found in Jurafsky (1992) and Jurafsky and Narayanan (1998). These implementations are based in part on a broadened notion of a construction that accounts for frequency based evidential access in a manner that is consistent with the data presented in this article.

In a construction-based account, knowledge of language is equated with knowledge of a large repository of constructions ranging from individual lexical items to rather abstract constructions like the Subject–Predicate Construction. In between, there exist a wide range of constructions encompassing both frozen and productive idioms and more traditional syntactic configurations. At the core of this approach is the idea that individual constructions consist of a structural alignment of specific grammatical, semantic and pragmatic facts about the language. Grammatical constructs are, therefore, coupled directly with their specific semantic and pragmatic content.

Under the accounts given in Jurafsky (1992) and Jurafsky and Narayanan (1998), interpretation is seen as a process of accumulating evidence in an incremental on-line fashion for the instantiation of a construction. Once instantiated, an interpretation is created by combining the particular semantic and pragmatic content of a construction with the corresponding content from other constructions in a working store. Multiple interpretations may be pursued in parallel as long as there is sufficient space in working memory and the individual interpretations are sufficiently well-formed. Pruning of competing interpretations occurs when an interpretation becomes too ill-formed when compared against its competitors in working memory.

Extending these notions to the case of conventional conceptual metaphors is fairly straightforward. Considered as constructions, conventional metaphors consist of a bundle of associations that directly encode the se-

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 223/228

mantic and pragmatic constraints on the source and target concepts that make up the metaphor. In context, the source and target components of a metaphor provide constraints on what can serve as possible evidence for the presence of a given metaphor. In the case of wholly conceptual metaphors this evidence is based on two factors: the presence of concepts placed in working memory by other constructions that match either the source and target parts of the metaphor, and the combined predictive power of these concepts based on statistical information particular to each individual metaphor. Finally, the set of conventional conceptual associations, or ground, provides the semantic and pragmatic constraints by which the intended meaning of the metaphor is constructed from the other conceptual content in working memory.

Note that under this model, the success or failure of a given metaphorical interpretation is not based in any direct way on the well-formedness of a literal interpretation. They are both merely possible candidates created from constructions that have been simultaneously activated by context. Successful interpretations are those that are most well-formed based on the constraints from their constituent constructions, and the degree of support they receive from context.

To summarize, the proposed construction-based framework is based on the following notions.

- An extension of the notion of a construction from traditional form-meaning pairings to conventionalized concept-concept pairings.
- Generic memory access notions such as priming, recency and frequency to control the activation of particular constructions.
- Working memory constraints to constrain the number of possible parallel interpretations.

5.4. Plausibility of the Metaphorical Construction Account

We now move on to consider how such a model might fare in light of the five constraints given above. In keeping with the spirit of the model, our intent is to show that the model is in accord with the constraints because it treats metaphor as a normal part of language processing, making use of generic processing capabilities.

5.4.1. Total-Time Constraint

Under this model, knowledge of conventional conceptual metaphor is represented within the same type of framework, and is subject to the same kind of processing constraints, as other forms of linguistic knowl-

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 224/228

edge including syntactic, idiomatic, lexical and pragmatic knowledge. The observed processing time for all types of interpretations is simply based on the amount of contextual evidence for the correct combination of constructions. Specifically, metaphorical utterances will be processed quickly and effectively when they provide sufficient evidence for the activation and integration of the appropriate conventional metaphors. This is precisely the same kind of processing required for all other literal and non-literal language.

5.4.2. On-Line Constraint and Non-Optionality Constraint

As discussed above, these two constraints are two sides of the same coin and will be discussed together. Under the framework sketched here, interpretation occurs via the activation and integration of constructions into working memory based on evidence from both within the utterance and prior context. Simultaneous interpretations can be pursued in parallel as long as there is sufficient evidence to activate their constituent constructions and the resulting interpretations are sufficiently well-formed. As more evidence becomes available the well-formedess of any interpretation may drop, causing it to be pruned from working memory. Therefore, metaphorical interpretations may be built up incrementally and eventually survive or be pruned based on the available evidence. The On-Line constraint is satisfied since partial metaphorical interpretations will always be built as long as there is minimal sufficient evidence to activate the required metaphorical constructions. At the same time these partial results can be either confirmed or short-circuited by subsequent evidence as it becomes available on-line.

5.4.3. Differential Behavior Constraint

In our suggested framework, individual conceptual metaphors are retrieved from long-term memory in the same fashion as other memory traces based on what Anderson (1990) calls their *need probability*. This is simply the probability that a memory trace will be needed given the evidence currently available in the context. In our model, this is achieved by augmenting constructions with two types of frequency information: the overall frequency of individual constructions, and the predictive power of the presence of their parts in context as evidence. The differential behavior of various metaphors arises both from their differing frequencies and the differing ability of their source and target parts to predict future uses.

Verlag: Walter de Gruyter | **P-Nr.:** B12-421394 | **Pfad/Job:** W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | **P-Anfang:** 07.09.2005 **P-Aktuell:** 23.09.2005 | **ID:** int01; int04 – AGB4 – 09.05 | **Printjob:** 23.09., 07:36, Seiten: 225/228

5.4.4. Contextual Influence

As with the previous constraints, contextual influences on metaphor are accounted for by assuming that they result from behavior that is a known part of the human memory system. Specifically, constructions are considered to be structured memory traces that are susceptible to all the various priming, recency and frequency effects observed in both the memory literature and the literature on lexical access and access to idioms. In particular, as suggested by Anderson's model, metaphorical constructions can be primed based either on their prior activation or the prior activation of their parts. The degree of activation is simply based on how predictive the evidence is of future uses of the metaphor. Therefore, the observed vast difference in ability of source concepts, target concepts and whole metaphors to predict future metaphors arises directly from the differing patterns of occurrence of these cues in real texts.

6. Conclusions

Our results indicate that the various experimental results concerning the effects of context on metaphor processing are neither artifacts of the laboratory nor artifacts of a special purpose metaphor processing mechanism. Rather, they are reflections of the environment in the language comprehension mechanism. Specifically, the various inhibition and facilitation effects of context on metaphor comprehension are consistent with corpus-based results concerning the predictive value of contextual cues for future metaphors. These results, when combined with converging evidence from other relevant studies, paint a picture of the on-line metaphor comprehension process as a normal part of our cognitive language capacity.

References

Anderson, John R. 1990. *The Adaptive Character of Thought*. Hillsdale, NJ: Lawrence Erlbaum.

Anderson, John R. and Lael J. Schooler 1991 Reflections of the environment in memory. *Psychological Science* 2: 396–408.

Carbonell, Jaime 1981 Invariance hierarchies in metaphor interpretation. *Proceedings of the Third Meeting of the Cognitive Science Society*, 292–295. Hillsdale, NJ: Lawrence Erlbaum.

Fass, Dan 1988 *Collative Semantics: A Semantics for Natural Language.* PhD dissertation, New Mexico State University, Las Cruces, New Mexico.

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 226/228

— 1991 met*: A method for discriminating metaphor and metonymy by computer. *Computational Linguistics* 17: 49–90.
Fillmore, Charles, Paul Kay and Mary O'Connor 1988 Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64: 510–538.
Gentner, Dedre, Brian Falkenhainer and Janice Skorstad 1988 Viewing metaphor as analogy. In: David H. Helman (ed.), *Analogical Reasoning*, 171–177. Dordrecht: Kluwer Academic Publishers.
Gernsbacher, Morton Ann, Boaz Keysar, Rachel Robertson and Necia Werner 2001 The Role of Suppression and Enhancement in Understanding Metaphors. *Journal of Memory and Language* 45: 433–450.
Gerrig, Richard J. 1989 Empirical constraints on computational theories of metaphor: Comments on Indurkhya. *Cognitive Science* 13: 235–241.
Gerrig, Richard J. and Alice F. Healy 1983 Dual processes in metaphor understanding: comprehension and appreciation. *Journal of Experimental Psychology: Learning, Memory and Cognition* 9: 667–675.
Gibbs, Raymond W. 1984 Literal meaning and psychological theory. *Cognitive Science* 8: 275–304.
— 1992 Categorization and metaphor understanding. *Psychological Review* 99: 572–577.
Gildea, Patricia and Sam Glucksberg 1983 On understanding metaphor: The role of context. *Journal of Verbal Learning and Verbal Behavior* 22: 577–590.
Glucksberg, Sam, Patricia Gildea and Howard Bookin 1982. On understanding nonliteral speech: Can people ignore metaphors?. *Journal of Verbal Learning and Verbal Behavior* 21: 85–98.
Hobbs, Jerry 1979 Metaphor, metaphor schemata, and selective inferencing. Technical Note 204, San Mateo, CA: SRI.
Indurkhya, Bipin 1987 Approximate semantic transference: A computational theory of metaphors and analogy. *Cognitive Science* 11: 445–480.
Inhoff, Alber W., Susan D Lima and Patrick J Carrol 1984 Contextual effects on metaphor comprehension in reading. *Memory and Cognition* 12: 558–567.
Johnson, Mark 1987 *The Body in the Mind: The Bodily Basis of Meaning, Imagination and Reason*. Chicago: The University of Chicago Press.
Jurafsky, Dan 1996 A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science* 20: 137–194.
Kemper, Susan 1989 Priming the comprehension of metaphors. *Metaphor and Symbolic Activity* 4: 1–17.
Keysar, Boaz 1989 On the functional equivalence of literal and metaphorical interpretations in discourse. *Journal of Language and Memory* 28: 375–385.
Lakoff, George 1987 *Women, Fire and Dangerous Things*. Chicago: The University of Chicago Press.
Lakoff, George and Mark Johnson 1980 *Metaphors We Live By*. Chicago: The University of Chicago Press.
Lakoff, George and Mark Turner 1988 *More Than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago: The University of Chicago Press.
Martin, James H. 1990 *A Computational Model of Metaphor Interpretation*. Cambridge, MA: Academic Press.
— 1992 Computer understanding of conventional metaphoric language. *Cognitive Science* 16: 233–270.
— 1994 Metabank: A knowledge-base of metaphoric language conventions. *Computational Intelligence* 10: 134–149.

Verlag: Walter de Gruyter | **P-Nr.:** B12-421394 | **Pfad/Job:** W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | **P-Anfang:** 07.09.2005 **P-Aktuell:** 23.09.2005 | **ID:** int01; int04 – AGB4 – 09.05 | **Printjob:** 23.09., 07:36, Seiten: 227/228

Narayanan, Srini 1999 Moving right along: A computational model of metaphoric reasoning about events. *Proceedings of the National Conference on Artificial Intelligence*, 121–128. Menlo Park, CA: AAAI Press.

Narayanan, Srini and Dan Jurafsky 1998 Bayesian models of human sentence processing. *Proceedings of the Cognitive Science Society*, 752–757. Hillsdale, NJ: Lawrence Erlbaum.

Oaksford, Mike and Nicholas Chater 1998 An introduction to rational models of cognition. In: *Rational Models of Cognition*, 1–17. Oxford: Oxford University Press.

Ortony, Andrew, Diane L. Schallert, Ralph E. Reynolds and Stephen J. Antos 1978 Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior* 17: 465–477.

Pollio, Howard R., Michael K. Smith, and Marilyn R. Pollio 1990 Figurative language and cognitive psychology. *Language and Cognitive Processes* 5: 141–167.

Russell, Sylvia W. 1976. Computer understanding of metaphorically used verbs. *American Journal of Computational Linguistics*, Microfiche 44.

Searle, John R. 1979 Metaphor. In: Andrew Ortony (ed), *Metaphor and Thought*, 92–123. Cambridge University Press.

Talmy, Leonard 1975 Semantics and syntax of motion. In: John Kimball (ed), *Syntax and Semantics* 4, 181–238. New York: Academic Press.

— 1988 Force dynamics in language and cognition. *Cognitive Science* 12: 49–100.

Wilks, Yorick 1978 Making preferences more active. *Artificial Intelligence* 11: 197–223.

Verlag: Walter de Gruyter | P-Nr.: B12-421394 | Pfad/Job: W:\DeGruyter\TiLSM\Band_171_Stefanowitsch\11_Martin.fm | P-Anfang: 07.09.2005 P-Aktuell: 23.09.2005 | ID: int01; int04 – AGB4 – 09.05 | Printjob: 23.09., 07:36, Seiten: 228/228